aws

Guida per gli sviluppatori

OpenSearch Servizio Amazon

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

OpenSearch Servizio Amazon: Guida per gli sviluppatori

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

I marchi e il trade dress di Amazon non possono essere utilizzati in relazione ad alcun prodotto o servizio che non sia di Amazon, in alcun modo che possa causare confusione tra i clienti, né in alcun modo che possa denigrare o screditare Amazon. Tutti gli altri marchi non di proprietà di Amazon sono di proprietà dei rispettivi proprietari, che possono o meno essere affiliati, collegati o sponsorizzati da Amazon.

Table of Contents

Che cos'è Amazon OpenSearch Service? 1
Caratteristiche di Amazon OpenSearch Service 2
Quando utilizzare 3
Versioni supportate di Elasticsearch e OpenSearch 4
Supporto standard ed esteso 4
Versioni standard supportate ed estese 7
Calcolo dei costi di assistenza estesa 7
Prezzi 8
Servizi correlati 9
Configurazione 11
Concessione delle autorizzazioni 11
Concessione dell'accesso programmatico 11
Configura il AWS CLI 13
Aprire la console 14
Nozioni di base 15
Creare un dominio 16
Caricamento dei dati per l'indicizzazione 17
Opzione 1: Caricamento di un singolo documento 18
Opzione 2: Caricamento di più documenti 18
Ricerca di documenti 19
Come eseguire la ricerca di documenti dalla riga di comando 19
Cerca documenti utilizzando le dashboard OpenSearch 20
Eliminazione di un dominio 21
OpenSearch Ingestione di Amazon 22
Concetti chiave 22
Vantaggi 24
Limitazioni 24
Versioni di Data Prepper supportate 25
Scalabilità delle pipeline 26
Prezzi 27
Supportato Regioni AWS 27
Configurazione di ruoli e utenti 28
Ruolo della pipeline 29
Ruolo di ingestione 32

Concedere alle pipeline l'accesso ai domini ... 34
Concedere alle pipeline l'accesso alle raccolte ... 38
Guida introduttiva a Ingestion OpenSearch ... 43
Tutorial: inserisci dati in un dominio ... 44
Tutorial: inserisci dati in una raccolta ... 52
Caratteristiche della pipeline ... 58
Buffering persistente ... 59
Divisione ... 61
Concatenamento ... 62
Code di lettere morte ... 63
Gestione degli indici ... 65
End-to-end riconoscimento ... 69
Contropressione alla fonte ... 69
Creazione di pipeline ... 70
Prerequisiti e ruolo IAM richiesto ... 71
Autorizzazioni IAM richieste ... 72
Specificare la versione della pipeline ... 73
Specificare il percorso di ingestione ... 74
Creazione di pipeline ... 75
Monitoraggio dello stato della creazione della pipeline ... 80
Lavorare con i progetti ... 81
Visualizzazione delle pipeline ... 83
Aggiornamento delle pipeline ... 85
Considerazioni ... 86
Autorizzazioni richieste ... 86
Aggiornamento delle pipeline ... 87
Implementazioni blu/verdi per gli aggiornamenti della pipeline ... 88
Gestione dei costi della pipeline ... 89
Arresto di una pipeline ... 90
Avvio di una pipeline ... 91
Eliminazione delle pipeline ... 92
Plugin e opzioni supportati ... 92
Plugin supportati ... 93
Processori stateless e processori stateful ... 95
Requisiti e vincoli di configurazione ... 96
Integrazione delle pipeline ... 101

Costruzione dell'endpoint di ingestione ... 101
Creazione di un ruolo di importazione ... 102
Servizi Atlassian ... 104
Amazon Aurora ... 122
Amazon DynamoDB ... 151
Amazon DocumentDB ... 167
Confluent Cloud Kafka ... 186
MSK Amazon ... 196
Amazon RDS ... 204
Simple Storage Service (Amazon S3) ... 233
Amazon Security Lake ... 243
Fluent Bit ... 247
Fluentd ... 249
Inferenza batch offline ML ... 251
OpenTelemetry Collettore ... 267
Amazon Managed Service per Prometheus ... 274
Kafka autogestito ... 282
Cluster autogestiti OpenSearch ... 290
Flusso di dati Amazon Kinesis ... 298
Fasi successive ... 310
AWS Lambda ... 310
Migrazione dei dati tra domini e raccolte ... 314
Limitazioni ... 314
OpenSearch Il servizio come fonte ... 315
Specificazione di più sink di dominio di servizio OpenSearch ... 317
Migrazione dei dati verso una raccolta OpenSearch VPC serverless ... 318
Gestione delle pipeline con AWS SDKs ... 319
Python ... 319
Sicurezza in Ingestion OpenSearch ... 323
Configurazione dell'accesso VPC per le pipeline ... 324
Configurazione delle pipeline per l'ingestione tra account ... 329
Identity and Access Management ... 336
Monitoraggio con CloudTrail ... 344
Endpoint VPC (AWS PrivateLink) ... 348
Etichettatura delle pipeline ... 351
Autorizzazioni richieste ... 352

Utilizzo dei tag (console) ... 352
Utilizzo dei tag (AWS CLI) ... 352
Registrazione di log e monitoraggio ... 353
Monitoraggio dei log della pipeline ... 353
Monitoraggio delle metriche della pipeline ... 355
Best practice ... 387
Best practice generali ... 388
Allarmi consigliati CloudWatch ... 388
Amazon OpenSearch Serverless ... 395
Vantaggi ... 395
Cos'è Amazon OpenSearch Serverless? ... 396
Casi d'uso per Serverless OpenSearch ... 397
Come funziona ... 397
Scelta di un tipo di raccolta ... 398
Prezzi ... 400
Supportato Regioni AWS ... 401
Limitazioni ... 401
Service e Serverless a confronto OpenSearch OpenSearch ... 402
Tutorial: Guida introduttiva a Serverless OpenSearch ... 405
Fase 1: configurazione delle autorizzazioni ... 406
Fase 2: creazione di una raccolta ... 407
Fase 3: Caricamento e ricerca dei dati ... 408
Gestione degli errori ... 409
Fase 4: Eliminazione della raccolta ... 410
Fasi successive ... 410
Raccolte ... 411
Gestione delle raccolte ... 411
Lavorare con le raccolte di ricerca vettoriale ... 435
Utilizzo di politiche relative al ciclo di vita dei dati ... 443
Gestire le raccolte con AWS SDKs ... 452
Creazione di raccolte con CloudFormation ... 463
Backup delle raccolte tramite istantanee ... 465
Supporto codec standard in Amazon Serverless OpenSearch ... 472
Salva spazio di archiviazione utilizzando una fonte derivata ... 474
Gruppi di raccolta ... 476
Concetti chiave ... 477

Limiti di capacità del gruppo di raccolta 477
Crittografia e chiavi KMS nei gruppi di raccolta 481
Crea gruppi di raccolta 482
Gestisci i gruppi di raccolta 484
Gestione dei limiti di capacità 488
Configurazione delle impostazioni di capacità 490
Limiti di capacità massima 491
Monitoraggio dell'utilizzo della capacità 491
Importazione dei dati nelle raccolte 491
Autorizzazioni minime richieste 492
OpenSearch Ingestione 493
Fluent Bit 493
Amazon Data Firehose 494
Go 494
Java 497
JavaScript 498
Logstash 501
Python 503
Ruby 505
Altri clienti 506
Configurazione del Machine Learning 507
Machine Learning 507
Connettori 507
Modelli 508
Configurare le autorizzazioni per il Machine Learning 508
Funzionalità e funzionalità non supportate APIs 510
Configura la ricerca neurale e ibrida 511
Ricerca neurale 511
Ricerca ibrida 511
Interrogazioni neurali e ibride 512
Configurazione delle autorizzazioni 513
Configura i flussi di lavoro 516
Flussi di lavoro 516
Configurazione delle autorizzazioni 516
Sicurezza in modalità serverless OpenSearch 518
Policy di crittografia 520

Policy di rete ... 520
Policy di accesso ai dati ... 521
Autenticazione IAM e SAML ... 522
Sicurezza dell'infrastruttura ... 523
Nozioni di base sulla sicurezza ... 523
Identity and Access Management ... 537
Encryption (Crittografia) ... 555
Accesso alla rete ... 566
Conformità a FIPS ... 578
Controllo dell'accesso ai dati ... 583
Endpoint VPC del piano dati ... 594
Endpoint VPC del piano di controllo ... 606
Autenticazione SAML ... 608
Convalida della conformità ... 618
Assegnazione di tag alle raccolte ... 618
Autorizzazioni richieste ... 619
Etichettatura delle raccolte (console) ... 619
Etichettatura delle raccolte ()AWS CLI ... 619
Operazioni e plug-in supportati ... 620
Operazioni e autorizzazioni API supportate OpenSearch ... 620
OpenSearch Plugin supportati ... 630
Monitoraggio senza server OpenSearch ... 631
Monitoraggio con CloudWatch ... 632
Monitoraggio con CloudTrail ... 639
Monitoraggio con EventBridge ... 644
Creazione e gestione dei domini ... 648
Creazione OpenSearch di domini di servizio ... 648
Creazione OpenSearch di domini di servizio (console) ... 648
Creazione OpenSearch di domini di servizio ()AWS CLI ... 655
Creazione OpenSearch di domini di servizio ()AWS SDKs ... 657
Creazione OpenSearch di domini di servizio ()AWS CloudFormation ... 657
Configurazione delle policy di accesso ... 657
Migrazione al OpenSearch servizio con Migration Assistant ... 658
Panoramica di ... 658
Scenari di migrazione ... 659
Guida introduttiva a Migration Assistant ... 660

Architettura della soluzione ... 660
Impostazioni avanzate del cluster ... 661
Arricchimento semantico automatico ... 661
Introduzione ... 661
Dettagli del modello e benchmark delle prestazioni ... 662
Lingue supportate ... 663
Imposta un indice di arricchimento semantico automatico per i domini ... 663
Inserimento e ricerca di dati ... 665
Configurazione delle autorizzazioni per l'arricchimento semantico automatico ... 666
Riscritture delle query ... 669
Limitazioni dell'arricchimento semantico automatico ... 669
Prezzi ... 670
Modifiche di configurazione ... 670
Modifiche che di solito causano implementazioni blue/green ... 671
Modifiche che di solito non causano blue/green distribuzioni ... 672
Opzioni di implementazione blu/verdi ... 673
Determinare se una modifica provocherà una blue/green distribuzione ... 675
Monitoraggio di una modifica alla configurazione ... 680
Fasi di una modifica della configurazione ... 682
Impatto delle implementazioni sulle prestazioni blue/green ... 685
Costi per le modifiche di configurazione ... 686
Risoluzione degli errori di convalida ... 686
Aggiornamenti del software del servizio ... 692
Aggiornamenti facoltativi e obbligatori ... 693
Aggiornamenti delle patch ... 694
Considerazioni ... 694
Avvio di un aggiornamento ... 695
Finestre non di punta ... 698
Aggiornamenti di monitoraggio ... 699
Quando i domini non sono idonei per un aggiornamento ... 699
Finestre non di punta ... 700
Aggiornamenti software di servizio non di punta ... 701
Ottimizzazioni Auto-Tune in caso di picco ... 702
Attivazione della finestra non di punta ... 702
Configurazione di una finestra personalizzata non di punta ... 703
Visualizzazione delle azioni pianificate ... 704

Riprogrammazione delle azioni ... 706
Migrazione dalle finestre di manutenzione di Auto-Tune ... 708
Notifications ... 709
Nozioni di base sulle notifiche ... 709
Gravità delle notifiche ... 710
Evento di esempio EventBridge ... 711
Configurazione di un dominio Multi-AZ ... 711
Multi-AZ con Standby ... 712
Multi-AZ senza Standby ... 714
Interruzioni delle zone di disponibilità ... 717
Supporto per VPC ... 719
VPC e domini pubblici ... 720
Limitazioni ... 720
Architecture ... 721
Creazione di snapshot di indici ... 728
Prerequisiti ... 729
Registrazione di un repository di snapshot manuali ... 733
Acquisizione di snapshot manuali ... 737
Ripristino di snapshot ... 739
Eliminazione degli snapshot manuali ... 742
Automatizzazione delle istantanee con Snapshot Management ... 742
Automazione di snapshot con Index State Management ... 745
Utilizzo di Curator per gli snapshot ... 745
Aggiornamento dei domini ... 746
Percorsi di aggiornamento supportati ... 746
Aggiornamento di un dominio (console) ... 751
Aggiornamento di un dominio (CLI) ... 751
Aggiornamento di un dominio (SDK) ... 752
Errori nella risoluzione dei problemi di convalida ... 754
Risoluzione dei problemi relativi all'aggiornamento ... 754
Utilizzo di uno snapshot per migrare i dati ... 756
Creazione di un endpoint personalizzato ... 763
Endpoint personalizzati per nuovi domini ... 764
Endpoint personalizzati per domini esistenti ... 765
Mappatura CNAME ... 765
Regolazione automatica ... 765

Tipi di modifiche 766
Abilitazione o disabilitazione della regolazione automatica 768
Pianificazione dei miglioramenti di Auto-Tune 768
Monitoraggio delle modifiche Auto-Tune 769
Assegnazione di tag ai domini 770
Esempi di assegnazione di tag 770
Etichettatura dei domini (console) 771
Taggare i domini ()AWS CLI 772
Etichettatura dei domini ()AWS SDKs 773
Esecuzione di azioni amministrative 775
Riavvio del OpenSearch processo su un nodo di dati 775
Riavvio di un nodo di dati 775
Riavvio del processo Dashboards 776
Limitazioni 777
Lavorare con le interrogazioni dirette 778
Interrogazioni dirette in S3 778
Creazione di una fonte di dati S3 779
Configurazione di un'origine dati S3 787
Prezzi 789
Limitazioni 790
Raccomandazioni 791
Quote 791
Supportato Regioni AWS 792
Interrogazioni dirette nei log CloudWatch 792
Creazione di una fonte di CloudWatch dati Logs 793
Configurazione di un'origine dati Logs CloudWatch 799
Prezzi 801
Limitazioni 802
Raccomandazioni 802
Quote 804
Supportato Regioni AWS 805
Interrogazioni dirette in Security Lake 806
Creazione di una fonte di dati Security Lake 806
Configurazione di un'origine dati Security Lake 813
Prezzi 815
Limitazioni 816

Raccomandazioni 816
Quote 817
Supportato Regioni AWS 818
Interrogazioni dirette in Managed Prometheus 819
Creazione di un'origine dati Prometheus gestita 820
Configurazione e interrogazione di un'origine dati Prometheus gestita 824
Prezzi 826
Limitazioni 826
Raccomandazioni 826
Quote 827
Supportato Regioni AWS 827
Gestione di una fonte di dati 828
Monitoraggio con CloudWatch fonti di dati metriche 829
Abilitazione e disabilitazione delle fonti di dati 832
Monitoraggio con budget AWS 832
Eliminazione di un'origine dati 833
Comandi SQL e PPL supportati 834
Comandi SQL supportati 835
Comandi PPL supportati 1044
Monitoraggio dei domini 1228
Monitoraggio dei parametri del cluster 1229
Visualizzazione delle metriche in CloudWatch 1230
Interpretazione delle cartelle cliniche in Service OpenSearch 1230
Parametri cluster 1231
Parametri nodo master dedicato 1239
Metriche del nodo di coordinatore dedicato 1241
Parametri volume EBS 1242
Parametri dell'istanza 1245
Metriche calde 1257
Parametri di archiviazione a freddo 1264
OpenSearch Metriche ottimizzate per le istanze () OR1 1265
Parametri di avvisi 1266
Parametri di rilevamento delle anomalie 1267
Parametri di ricerca asincrona 1269
Metriche Auto-Tune 1271
Multi-AZ con metriche Standby 1272

Metriche puntuali ... 1274
Parametri SQL ... 1275
Parametri k-NN ... 1276
Parametri di ricerca tra cluster ... 1280
Parametri di replica tra cluster ... 1280
Parametri di Learning to Rank ... 1282
Parametri Piped Processing Language (PPL) ... 1283
Monitoraggio dei log ... 1283
Abilitazione della pubblicazione di log (console) ... 1285
Abilitazione della pubblicazione di log (AWS CLI) ... 1287
Abilitazione della pubblicazione di log (AWS SDKs) ... 1289
Abilitazione della pubblicazione di log (CloudFormation) ... 1290
Impostazione delle soglie di slow log delle richieste di ricerca ... 1292
Impostazione delle soglie di slow log degli shard ... 1293
Test degli slow log ... 1293
Visualizzazione dei log di ... 1294
Monitoraggio dei log di verifica ... 1294
Limitazioni ... 1295
Abilitazione dei log di verifica ... 1295
Abilita la registrazione di controllo utilizzando il AWS CLI ... 1297
Abilitare la registrazione di controllo tramite l'API di configurazione ... 1298
Livelli e categorie dei log di verifica ... 1298
Impostazioni dei log di verifica ... 1301
Esempi di log di verifica ... 1304
Configurazione dei log di verifica tramite la REST API ... 1307
Monitoraggio degli eventi ... 1308
Eventi di aggiornamento del software di servizio ... 1309
Eventi di regolazione automatica ... 1316
Eventi sull'integrità del cluster ... 1321
Eventi di endpoint VPC ... 1334
Eventi di ritiro dei nodi ... 1336
Eventi di ritiro dei nodi degradati ... 1338
Eventi di errore di dominio ... 1340
Tutorial: ascolto degli eventi OpenSearch di servizio ... 1343
Tutorial: Invio di avvisi SNS per gli aggiornamenti disponibili ... 1345
Monitoraggio con CloudTrail ... 1346

Informazioni sul OpenSearch servizio Amazon in CloudTrail 639
Informazioni sulle voci dei file OpenSearch di registro di Amazon Service 643
Cluster Insights 1351
Vantaggi 1351
Accedi a Cluster Insights tramite Console 1352
Accedi a Cluster Insights e alle metriche dettagliate tramite l'interfaccia utente OpenSearch .. 1352
Crea e configura un'applicazione per visualizzare Cluster Insights 1353
Comprendere Cluster Insights 1354
Dashboard di panoramica 1355
Stato attuale del cluster 1355
Andamento degli approfondimenti 1355
Informazioni aperte attuali 1356
OpenSearch Cluster di servizi 1356
Le migliori informazioni in base alla gravità 1356
Dettagli degli approfondimenti 1356
Dettagli del cluster 1356
Panoramica di 1356
Best practice per la configurazione 1357
Approfondimenti 1357
Sezione Metriche 1358
Visualizzazioni di nodi, indici e frammenti 1358
Visualizzazione delle interrogazioni 1359
Accedi a Insights tramite EventBridge eventi Amazon 1361
Catalogo Insights 1361
Sicurezza 1365
Protezione dei dati 1366
Crittografia dei dati a riposo 1367
Node-to-node crittografia 1371
Identity and Access Management 1372
Tipi di policy 1372
Effettuazione e firma di richieste OpenSearch di assistenza 1381
Quando le policy entrano in collisione 1383
Riferimenti agli elementi della policy 1384
Opzioni avanzate e considerazioni sulle API 1390
Configurazione delle policy di accesso 1394
Altre policy di esempio 1394

Riferimento alle autorizzazioni API ... 1394
AWS politiche gestite ... 1394
Prevenzione del problema "confused deputy" tra servizi ... 1405
Controllo granulare degli accessi ... 1407
Il quadro generale: controllo granulare degli accessi e sicurezza dei servizi OpenSearch ... 1408
Concetti chiave ... 1411
Informazioni sull'utente principale ... 1412
Abilitazione del controllo granulare degli accessi ... 1413
Accesso alle OpenSearch dashboard come utente principale ... 1417
Gestione delle autorizzazioni ... 1418
Configurazioni consigliate ... 1424
Limitazioni ... 1426
Modifica dell'utente principale ... 1428
Utenti principali aggiuntivi ... 1428
Snapshot manuali ... 1430
Integrazioni ... 1430
Differenze della REST API ... 1431
Tutorial: controllo granulare degli accessi con autenticazione Cognito ... 1433
Tutorial: Database utente interno e autenticazione di base ... 1438
Convalida della conformità ... 1441
Resilienza ... 1441
Token Web JSON ... 1442
Considerazioni ... 1442
Modifica della policy di accesso al dominio ... 1443
Configurazione dell'autenticazione e dell'autorizzazione JWT ... 1443
Utilizzo di un JWT per inviare una richiesta di test ... 1444
Sicurezza dell'infrastruttura ... 1446
Utilizzo degli endpoint OpenSearch VPC gestiti dal servizio ... 1446
Autenticazione SAML per dashboard OpenSearch ... 1450
Panoramica della configurazione SAML ... 1451
Considerazioni ... 1451
Autenticazione SAML per domini VPC ... 1452
Modifica della policy di accesso al dominio ... 1452
Configurazione dell'autenticazione avviata da SP o da IdP ... 1454
Configurazione dell'autenticazione avviata da SP e avviata da IdP ... 1460
Configurazione dell'autenticazione SAML (AWS CLI) ... 1461

Configurazione dell'autenticazione SAML (API di configurazione) 1461
Risoluzione dei problemi SAML 1462
Disabilitazione dell'autenticazione SAML 1465
Supporto IAM Identity Center per OpenSearch 1466
Considerazioni 1466
Modifica della policy di accesso al dominio 1467
Configurazione dell'autenticazione e dell'autorizzazione di IAM Identity Center (console) ... 1467
Configurazione del controllo granulare degli accessi 1468
Configurazione dell'autenticazione e dell'autorizzazione (CLI) di IAM Identity Center 1468
Disabilitazione dell'autenticazione IAM Identity Center sul dominio 1469
Autenticazione Amazon Cognito per dashboard OpenSearch 1469
Prerequisiti 1470
Configurazione di un dominio per l'uso dell'autenticazione Amazon Cognito 1473
Concessione del ruolo autenticato 1477
Configurazione dei provider di identità 1478
(Facoltativo) Configurazione dell'accesso granulare 1479
(Facoltativo) Personalizzazione della pagina di accesso 1480
(Facoltativo) Configurazione della sicurezza avanzata 1480
Test in corso 1480
Quote 1481
Problemi di configurazione comuni 1481
Disattivazione dell'autenticazione Amazon Cognito per dashboard OpenSearch 1485
Eliminazione di domini che utilizzano l'autenticazione Amazon Cognito per dashboard OpenSearch 1486
Uso di ruoli collegati ai servizi 1486
Ruolo di creazione del dominio VPC e dell'origine dati 1487
Ruolo di creazione della raccolta 1490
Ruolo di creazione della pipeline 1493
Codice di esempio 1497
Compatibilità con i client Elasticsearch 1497
Compressione delle richieste HTTP 1498
Abilitazione della compressione gzip 1498
Intestazioni richieste 1499
Codice di esempio (Python 3) 1499
Usando il AWS SDKs 1501
Java 1501

Python 1512
Nodo 1515
Indicizzazione dei dati 1518
Limitazioni di denominazione per gli indici 1518
Riduzione delle dimensioni della risposta 1519
Codec indicizzati 1521
Caricamento di dati di streaming in OpenSearch Service 1521
Caricamento di dati di streaming da Ingestion OpenSearch 1522
Caricamento di dati in streaming da Amazon S3 1522
Caricamento dei dati in streaming in Amazon Kinesis Data Streams 1528
Caricamento di dati in streaming da una tabella Amazon DynamoDB 1532
Caricamento di dati di streaming da Amazon Data Firehose 1537
Caricamento di dati di streaming da Amazon CloudWatch 1538
Caricamento di dati di streaming da AWS IoT 1538
Caricamento dei dati con Logstash 1538
Configurazione 1538
Ricerca di dati 1542
Ricerche negli URI 1543
Ricerche nel corpo della richiesta 1544
Campi di boosting 1546
Evidenziazione dei risultati della ricerca 1546
API conteggio 1548
Paginazione dei risultati della ricerca 1549
Point in time 1549
I `size` parametri `from` e 1549
Linguaggio query dashboard 1550
Pacchetti 1552
Autorizzazioni richieste 1553
Caricamento di pacchetti in Amazon S3 1553
Importazione e associazione di pacchetti 1554
Utilizzo di pacchetti con OpenSearch 1555
Aggiornamento dei pacchetti 1559
Aggiornamento manuale degli indici con un nuovo dizionario 1563
Dissociazione e rimozione dei pacchetti 1565
Plugin personalizzati 1566
Plugin di terze parti 1583

Supporto per SQL 1587
Chiamata di esempio 1589
Note e differenze 1589
SQL Workbench 1590
SQL CLI 1444
Driver JDBC 1590
Driver ODBC 1591
Funzionalità di ricerca tra cluster 1591
Limitazioni 1592
Prerequisiti di ricerca tra cluster 1592
Prezzi della funzionalità di ricerca tra cluster 1593
Configurazione di una connessione 1593
Rimozione di una connessione 1594
Configurazione della sicurezza e spiegazione passo per passo di esempio 1595
OpenSearch Pannelli di controllo 1601
Learning to Rank 1601
Nozioni di base su Learning to Rank 1602
API Learning to Rank 1623
Ricerca asincrona 1630
Chiamata di ricerca di esempio 1630
Autorizzazioni di ricerca asincrona 1632
Impostazioni della ricerca asincrona 1633
Funzionalità di ricerca tra cluster 1633
UltraWarm 1635
Point in time 1635
Considerazioni 1635
Crea un PIT 1636
Autorizzazioni temporanee 1638
Impostazioni PIT 1639
Funzionalità di ricerca tra cluster 1639
UltraWarm 1639
Ricerca agentica 1639
Ricerca semantica 1640
Ricerca simultanea di segmenti 1640
Generazione di query in linguaggio naturale 1641
Prerequisiti 1641

Nozioni di base 1642
Configurazione delle autorizzazioni 1642
Automazione della configurazione 1643
Ricerca vettoriale 1644
Importazione da Amazon S3 Vectors a Serverless OpenSearch 1645
Architettura di integrazione 1646
Autorizzazioni IAM richieste 1648
Configurazione dell'integrazione di Amazon S3 Vectors con OpenSearch 1651
Funzionalità di ricerca avanzate con un motore vettoriale Amazon S3 1653
Abilitazione di S3 Vectors 1653
Creazione di indici con il motore vettoriale S3 1654
Limitazioni funzionali 1656
Indicizzazione dei documenti 1658
Ricerca di documenti 1658
Parametri di mappatura supportati 1660
Misurazione e fatturazione 1661
Disattivazione del motore s3vector 1661
Ricerca k-NN 1661
Nozioni di base su k-NN 1662
Differenze, regolazione e limitazioni di k-NN 1664
Ingestione vettoriale 1665
Come funziona 1665
OpenSearch Prezzi di importazione vettoriale 1665
Prerequisiti 1665
Crea un database vettoriale 1666
Funzionalità correlate 1667
Esporta l'indice vettoriale di Amazon S3 nel motore vettoriale Service OpenSearch 1668
Importa lo spazio dei nomi vettoriali Amazon S3 nel motore vettoriale Service OpenSearch 1669
Visualizza i processi di inserimento vettoriale e la cronologia delle importazioni 1670
Ottimizzazione automatica 1672
Vantaggi 1672
Come funziona 1673
Prerequisiti 1673
Casi d'uso per l'ottimizzazione automatica 1674
Limitazioni 1675

Fatturazione e costi 1676
Formati di dati supportati 1677
Utilizzo dell'ottimizzazione automatica nella console 1680
Funzionalità correlate 1686
Accelerazione GPU per l'indicizzazione vettoriale 1686
Prerequisiti 1686
Come funziona 1686
Prezzi dell'accelerazione GPU 1687
Accelerazione tramite GPU e operazioni di scrittura 1687
Configurazioni degli indici supportate 1688
Supportato Regioni AWS 1688
Abilita l'accelerazione GPU 1689
Crea indici vettoriali accelerati dalla GPU 1692
Indicizzazione e unione forzata 1693
Best practice 1694
OpenSearch Pannelli di controllo 1695
Controllo dell'accesso ai pannelli di controllo 1695
Utilizzo di un proxy per accedere al servizio da dashboard OpenSearch 1696
Configurazione delle dashboard per l'utilizzo di un server di mappe WMS 1700
Connessione di un server Dashboards locale al servizio OpenSearch 1701
Gestione degli indici nelle dashboard 1702
Funzionalità aggiuntive 1703
OpenSearch Interfaccia utente 1704
Cronologia delle versioni 1705
Nozioni di base 1707
Autorizzazioni richieste per la creazione di applicazioni Amazon OpenSearch Service 1708
Creazione di un'applicazione 1711
Gestione degli amministratori delle applicazioni 1719
AI agentica 1723
Funzionalità chiave 1724
Limiti di prezzo e utilizzo 1724
Sicurezza 1725
Abilitazione e disabilitazione delle funzionalità di intelligenza artificiale agentica 1725
Disponibilità nelle regioni 1726
Chat agentica 1727
Agente investigativo 1731

Memoria agentica 1733
Crittografia e chiave gestita dal cliente 1734
Prerequisiti per l'utilizzo delle chiavi gestite dal cliente 1735
Creazione di un'applicazione con crittografia a chiave gestita dal cliente utilizzando la console 1737
Creazione di un'applicazione con crittografia a chiave gestita dal cliente utilizzando AWS CLI 1737
Monitoraggio dell'utilizzo delle chiavi gestite dal cliente 1738
Aggiornamento delle impostazioni di crittografia 1738
Abilitazione della federazione SAML con IAM 1739
Passaggio 1: configurare l'applicazione del provider di identità (Okta) 1740
Passaggio 2: configura la configurazione per Okta AWS 1744
Fase 3: Creare la policy di accesso ad Amazon OpenSearch Service in IAM 1746
Passaggio 4: verifica l'esperienza Single Sign-On avviata dal provider di identità con SAML 1749
Passaggio 5: configura il controllo granulare degli accessi basato sugli attributi SAML 1753
Gestione delle associazioni delle fonti di dati e delle autorizzazioni di accesso VPC 1757
Associazione di una fonte di dati a un'applicazione UI OpenSearch 1757
Gestione dell'accesso ai domini in un VPC 1758
Configurazione dell'accesso alle raccolte OpenSearch Serverless in un VPC 1759
Utilizzo delle aree di lavoro OpenSearch di Amazon Service 1763
OpenSearch Creazione di aree di lavoro per applicazioni UI 1763
Privacy dello spazio di lavoro e collaboratori 1764
Tipi di aree di lavoro 1764
Accesso ai dati tra regioni e account diversi 1765
Accesso ai dati di più account ai domini OpenSearch 1767
Funzionalità di ricerca tra cluster 1789
Gestione dell'accesso all' OpenSearch interfaccia utente da un endpoint VPC 1800
Creazione di una connessione privata tra un VPC e un'interfaccia utente OpenSearch 1800
Aggiornamento della policy degli endpoint VPC per consentire l'accesso all'applicazione UI OpenSearch 1802
Revoca dell'accesso all' OpenSearch interfaccia utente in una policy di endpoint VPC 1802
Migrazione da Dashboard OpenSearch 1803
Panoramica sulla migrazione 1804
Prerequisiti 1804
Fase 1: Esportazione degli oggetti salvati dai pannelli di controllo OpenSearch 1805

Passaggio 2: Importa gli oggetti salvati nell'interfaccia utente OpenSearch 1808
Endpoint e quote 1810
OpenSearch Endpoint dell'interfaccia utente 1810
OpenSearch quote di servizio UI 1813
Gestione degli indici 1814
Istanze ottimizzate per OpenSearch 1814
Limitazioni 1815
Ottimizzazione per una migliore velocità di ingestione 1815
In che modo le istanze OpenSearch ottimizzate differiscono dalle altre istanze 1816
In che modo le istanze OpenSearch ottimizzate differiscono dalle istanze UltraWarm 1816
Fornitura di istanze OpenSearch ottimizzate a un dominio 1817
Storage a più livelli 1819
Architettura di storage a più livelli 1820
UltraWarm archiviazione 1826
Prerequisiti 1827
UltraWarm requisiti di archiviazione e considerazioni sulle prestazioni 1828
UltraWarm prezzi 1829
Abilitazione UltraWarm 1830
Migrazione degli indici verso lo storage UltraWarm 1832
Automazione delle migrazioni 1835
Regolazione della migrazione 1835
Annullamento di migrazioni 1836
Elenco degli indici ad accesso frequente e degli indici a caldo 1836
Ritorno di indici a caldo all'archiviazione ad accesso frequente 1837
Ripristino degli indici caldi dalle istantanee 1837
Snapshot manuali di indici a caldo 1839
Migrazione degli indici a caldo all'archiviazione a freddo 1840
Procedure consigliate per gli indici KNN 1840
Disabilitazione UltraWarm 1841
Archiviazione a freddo 1841
Prerequisiti 1842
Requisiti di archiviazione a freddo e considerazioni sulle prestazioni 1844
Prezzi dell'archiviazione a freddo 1844
Abilitazione dell'archiviazione a freddo 1844
Gestione degli indici freddi nelle dashboard OpenSearch 1846
Migrazione degli indici all'archiviazione a freddo 1846

Automatizzazione delle migrazioni all'archiviazione a freddo 1848
Annullamento delle migrazioni all'archiviazione a freddo 1848
Elencare gli indici freddi 1849
Migrazione degli indici a freddo all'archiviazione a caldo 1853
Ripristino di indici a freddo da snapshot 1854
Annullamento delle migrazioni dall'archiviazione a freddo a quella a caldo 1854
Aggiornamento dei metadati dell'indice a freddo 1855
Eliminazione di indici freddi 1855
Disabilitazione dell'archiviazione a freddo 1856
Index State Management 1856
Creazione di una policy ISM 1857
Policy di esempio 1858
Modelli ISM 1862
Differenze 1862
Tutorial: Automazione dei processi ISM 1864
Rollup di indici 1869
Creazione di un processo di rollup dell'indice 1869
Trasformazioni degli indici 1871
Creazione di un processo di trasformazione dell'indice 1871
Replica tra cluster 1872
Limitazioni 1873
Prerequisiti 1874
Requisiti per le autorizzazioni 1875
Configurazione di una connessione tra cluster 1876
Avvio della replica 1877
Conferma della replica 1878
Sospensione e ripristino della replica 1879
Arresto della replica 1880
Auto-follow 1880
Aggiornamento dei domini connessi 1882
Reindicizzazione remota 1882
Prerequisiti 1883
Reindicizza i dati tra i domini Internet del Servizio OpenSearch 1883
Reindicizza i dati quando il dominio remoto si trova in un VPC 1885
Reindicizza i dati tra domini non OpenSearch di servizio 1889
Reindicizzazione di set di dati di grandi dimensioni 1890

Impostazioni di reindicizzazione remota ... 1892
Flussi dei dati ... 1892
Nozioni di base sui flussi di dati ... 1893
Monitoraggio dei dati ... 1896
Avviso ... 1896
Autorizzazioni per gli avvisi ... 1897
Nozioni di base sugli avvisi ... 1897
Notifications ... 1898
Differenze ... 1898
Rilevamento anomalie ... 1900
Prerequisiti ... 1901
Nozioni di base sul rilevamento di anomalie ... 1901
Tutorial: Rileva un elevato utilizzo della CPU con il rilevamento delle anomalie ... 1904
Assistente AI ... 1908
Supportata Regioni AWS ... 1908
Configurazione di AI Assistant for OpenSearch Service ... 1909
Genera visualizzazioni utilizzando il linguaggio naturale ... 1909
Visualizza riepiloghi e approfondimenti sugli avvisi ... 1910
Prima di iniziare ... 1912
Visualizzazione di riepiloghi e approfondimenti sugli avvisi ... 1914
Visualizza i riepiloghi dei risultati delle query generati da AI Assistant nella pagina Discover ... 1914
Visualizza i rilevatori di anomalie consigliati ... 1915
Accedi alla chat di AI Assistant per le domande OpenSearch di servizio ... 1916
Machine learning ... 1918
Connettori per Servizi AWS ... 1918
Prerequisiti ... 1919
Crea un connettore di servizio OpenSearch ... 1922
Connettori per piattaforme esterne ... 1924
Prerequisiti ... 1925
Crea un connettore di servizio OpenSearch ... 1928
CloudFormation integrazioni di modelli ... 1930
Modelli disponibili CloudFormation ... 1931
Prerequisiti ... 1932
Modelli Amazon Bedrock ... 1933
Configurazione di Agentic Search con Bedrock Claude ... 1934
Modelli di integrazione con server MCP ... 1936

SageMaker Modelli Amazon ... 1939
Inferenza remota per modelli di evidenziazione semantica ... 1940
Impostazioni ML Commons non supportate ... 1941
Plugin del framework Flow ... 1941
Creazione di connettori ML in Service OpenSearch ... 1942
Configurazione delle autorizzazioni ... 1949
Osservabilità ... 1951
Cosa offre Amazon OpenSearch Service ... 1951
L'approccio OpenSearch di Amazon Service all'osservabilità ... 1951
Nozioni di base ... 1952
Avvio rapido ... 1952
Implementazione CDK ... 1953
Invio di telemetria ... 1953
Ulteriori informazioni ... 1954
Acquisizione della telemetria delle applicazioni ... 1954
OpenTelemetry Configurazione del Collector ... 1955
Configurazione delle pipeline OpenSearch di ingestione ... 1956
Verifica dell'ingestione ... 1960
Fasi successive ... 1960
Set di dati ... 1960
Tipi di set di dati ... 1961
Per creare un set di dati di log ... 1961
Per creare un set di dati Traces ... 1963
Per visualizzare i set di dati ... 1964
Analisi dei set di dati in Discover ... 1964
Scopri i log ... 1965
Per accedere alla pagina dei registri ... 1965
Esplorazione dei dati di registro ... 1965
Interrogazione dei log tramite PPL ... 1966
Creazione di visualizzazioni dai log ... 1967
Aggiungere visualizzazioni alle dashboard ... 1970
Scopri le tracce ... 1970
Per accedere alla pagina Traces ... 1970
Configurazione dei set di dati di traccia ... 1971
Esplorazione dei dati di traccia ... 1972
Visualizzazione di un intervallo specifico ... 1973

Pagina dei dettagli di Trace 1973
Correlazione delle tracce con i log 1974
Interrogazione delle tracce tramite PPL 1976
Scopri le metriche 1977
Configurazione di un'origine dati Prometheus 1977
pannello Query 1978
Filtro temporale 1978
Visualizzazione dei risultati 1978
Configurazione delle visualizzazioni 1979
Correlazioni 1979
Requisiti di correlazione 1979
Per creare una correlazione trace-to-logs 1980
Visualizzazione delle correlazioni nei log, nei set di dati 1983
Utilizzo delle correlazioni nella pagina Traces 1983
Gestione delle correlazioni 1984
Pannelli di controllo 1984
Struttura del pannello di controllo 1986
Creazione di dashboard da Discover 1986
Filtri del dashboard 1987
Creazione di dashboard 1988
Condivisione dei pannelli di controllo 1991
Dashboard di risoluzione dei problemi 1992
Monitoraggio dell'applicazione 1994
Prerequisiti 1994
Come funziona 1995
Servizi 1995
Mappa dell'applicazione 2002
Osservabilità dell'IA 2006
Nozioni di base 2006
Interfaccia utente di Agent Tracing 2008
Analisi di sicurezza 2015
Componenti e concetti di analisi della sicurezza 2015
Tipi di log 2016
Rilevatori 2016
Regole 2016
Esiti 2017

Avvisi ... 2017
Esplorazione dell'analisi della sicurezza ... 2017
Configurazione delle autorizzazioni ... 2019
Risoluzione dei problemi ... 2021
Nessun errore di indice di questo tipo ... 2021
Best practice ... 2022
Monitoraggio e avvisi ... 2022
Configura CloudWatch gli allarmi ... 2022
Abilitazione della pubblicazione dei log ... 2022
Strategia di partizione ... 2023
Determinazione del numero di partizioni e di nodi di dati ... 2024
Evitare l'asimmetria di storage ... 2025
Stabilità ... 2025
Resta aggiornato con OpenSearch ... 2025
Migliora le prestazioni delle istantanee ... 2026
Abilitare nodi principali dedicati ... 2026
Esecuzione dell'implementazione in più zone di disponibilità ... 2027
Controllo del flusso di importazione e del buffering ... 2027
Creare mappature per i carichi di lavoro di ricerca ... 2028
Utilizzare modelli di indici ... 2028
Gestire gli indici con Index State Management ... 2030
Rimuovere gli indici inutilizzati ... 2030
Utilizzare più domini per un'elevata disponibilità ... 2030
Performance ... 2031
Ottimizzare le dimensioni e la compressione delle richieste in blocco ... 2031
Ridurre le dimensioni delle risposte alle richieste in blocco ... 2031
Ottimizzare gli intervalli di aggiornamento ... 2031
Abilitare la regolazione automatica ... 2032
Sicurezza ... 2032
Abilitare il controllo granulare degli accessi ... 2032
Distribuire domini all'interno di un VPC ... 2033
Applicare una policy di accesso restrittiva ... 2033
Abilitare la crittografia dei dati a riposo ... 2033
Abilita la crittografia node-to-node ... 2033
Monitora con AWS Security Hub CSPM ... 2034
Ottimizzazione dei costi ... 2034

Utilizzare tipi di istanza di ultima generazione 2034
Utilizzo dei volumi gp3 di Amazon EBS più recenti 2034
Utilizzo UltraWarm e conservazione a freddo dei dati di registro delle serie temporali 2035
Rivedere i suggerimenti per le istanze riservate 2035
Allarmi consigliati CloudWatch 2036
Altri allarmi che potresti prendere in considerazione 2040
Dimensionamento dei domini 2043
Calcolo dei requisiti di archiviazione 2044
Scelta del numero di partizioni 2046
Scelta del tipo di istanza e test 2047
Scala in petabyte 2049
Nodi coordinatori dedicati 2051
Quando utilizzare i nodi coordinatori dedicati 2051
Architettura e comportamento 2051
Requisiti e limitazioni 2052
Fornitura di nodi coordinatori dedicati 2052
Best practice 2053
Nodi master dedicati 2055
Scelta del numero di nodi principali dedicati 2056
Scelta dei tipi di istanza per nodi principali dedicati 2057
Riferimenti generali 2059
Tipi di istanze supportati 2059
Tipi di istanza della generazione attuale 2059
Tipi di istanza di generazioni precedenti 2077
Funzionalità per versione di motore 2088
Plug-in per versione motore 2094
Plugin opzionali 2098
Operazioni supportate 2099
Differenze significative tra le API 2099
Quote 2154
Quote di archiviazione Warm Node 2154
Numero di nodi dati per AZ 2057
Limite totale di nodi per famiglia di istanze 2156
Quote delle dimensioni dei volumi EBS 2157
Quote di rete 2170
Quote di dimensione condivise 2180

Quote di conteggio condivise ... 2181
Quota dei processi Java ... 2182
Quota della policy di dominio ... 2182
Istanze riservate ... 2182
Acquisto di istanze riservate (console) ... 2183
Acquisto di istanze riservate (AWS CLI) ... 2184
Acquisto di istanze riservate ()AWS SDKs ... 2186
Analisi dei costi ... 2188
Ottimizzazione dei costi ... 2188
Ottimizzazione dei costi per i cluster gestiti da Amazon OpenSearch Service ... 2189
Ottimizzazione dei costi per Amazon OpenSearch Service Serverless ... 2194
Altre risorse supportate ... 2199
Esercitazioni ... 2200
Creazione e ricerca di documenti ... 2200
Prerequisiti ... 2200
Aggiunta di un documento a un indice ... 2201
Creazione generata automaticamente IDs ... 2202
Aggiornamento di un documento con un comando POST ... 2203
Esecuzione di operazioni in blocco ... 2204
Ricerca di documenti ... 2205
Risorse correlate ... 2207
Migrazione al servizio OpenSearch ... 2207
Acquisizione e caricamento dello snapshot ... 2207
Creare un dominio ... 2209
Fornire le autorizzazioni al bucket S3. ... 2210
Ripristino dello snapshot ... 2212
Creazione di un'applicazione di ricerca ... 2214
Prerequisiti ... 2215
Fase 1: Indicizzazione dei dati di esempio ... 2216
Fase 2: Creare e distribuire la funzione Lambda ... 2216
Fase 3: Creare l'API in API Gateway ... 2219
Fase 4: (facoltativa) Modifica della policy di accesso al dominio ... 2222
Mappatura del ruolo Lambda (se si utilizza il controllo granulare degli accessi) ... 2225
Fase 5: Test dell'applicazione Web ... 2225
Fasi successive ... 2227
Visualizzazione delle chiamate di supporto ... 2228

Fase 1: Configurazione dei prerequisiti 2229
Fase 2: Copia del codice di esempio 2230
(Facoltativo) Fase 3: Indicizzazione dei dati di esempio 2234
Fase 4: Analisi e visualizzazione dei dati 2236
Fase 5: Pulizia delle risorse e fasi successive 2240
Rinomina OpenSearch del servizio Amazon 2242
Nuova versione dell'API 2242
Tipi di istanza rinominati 2242
Modifiche delle policy di accesso 2243
Policy IAM 2243
Policy SCP 2243
Nuovi tipi di risorsa 2244
Kibana è stato rinominato in Dashboards OpenSearch 2245
CloudWatch Metriche rinominate 2245
Modifiche della console Gestione fatturazione e costi 2247
Nuovo formato evento 2247
Cosa rimane lo stesso? 2248
Inizia: aggiorna i tuoi domini alla OpenSearch versione 1.x 2248
Risoluzione dei problemi 2250
Impossibile accedere alle OpenSearch dashboard 2250
Impossibile accedere al dominio VPC 2250
Cluster in stato di sola lettura 2250
Cluster in stato rosso 2252
Correzione automatico di cluster rossi 2253
Ripristino da un carico di elaborazione costantemente elevato 2254
Stato giallo del cluster 2256
ClusterBlockException 2256
Mancanza di spazio di archiviazione disponibile 2256
Pressione di memoria JVM elevata 2257
Errore durante la migrazione a Multi-AZ con Standby 2258
Creazione di un indice, di un modello di indice o di una politica ISM durante la migrazione da domini senza standby a domini con standby 2021
Numero errato di copie dei dati 2258
JVM OutOfMemoryError 2258
Nodi cluster con errori 2259
Limite massimo di partizioni superato 2260

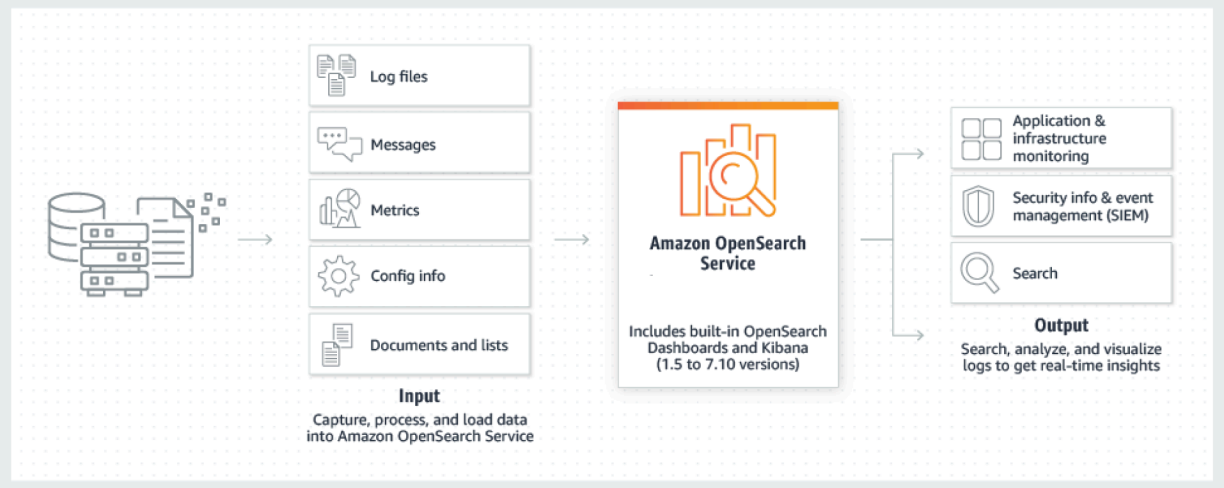
Dominio bloccato nello stato di elaborazione ... 2260
Saldo di burst EBS basso ... 2261
La metrica EBS aumenta durante il ridimensionamento del volume ... 2261
Impossibile abilitare i log di verifica ... 2261
Impossibile chiudere l'indice ... 2262
Controlli delle licenze client ... 2262
Limitazione delle richieste ... 2262
Impossibile eseguire SSH nel nodo ... 2262
Errore snapshot "Non valido per la classe di archiviazione dell'oggetto" ... 2263
Intestazione dell'host non valida ... 2263
Tipo di istanza M3 non valido ... 2263
Le hot query smettono di funzionare dopo l'attivazione UltraWarm ... 2264
Impossibile eseguire il downgrade dopo l'aggiornamento ... 2264
Hai bisogno di un riepilogo dei domini per tutti Regioni AWS ... 2264
Errore del browser durante l'utilizzo OpenSearch delle dashboard ... 2265
Asimmetria di partizioni e storage di nodi ... 2265
Asimmetria di partizioni e storage di indici ... 2266
Operazione non autorizzata dopo la selezione dell'accesso VPC ... 2267
Caricamento bloccato dopo la creazione di un dominio VPC ... 2267
Richieste rifiutate all'API OpenSearch ... 2267
Impossibile connettersi da Alpine Linux ... 2268
Troppe richieste per Search Backpressure ... 2269
Errore di certificato quando si utilizza un SDK ... 2269
L'installazione del plug-in personalizzato non riesce a causa della compatibilità della versione ... 2270
Cronologia dei documenti ... 2273
Aggiornamenti precedenti ... 2340
AWS Glossario ... 2344
... mmcccxlv

Che cos'è Amazon OpenSearch Service?

Amazon OpenSearch Service è un servizio gestito che semplifica l'implementazione, il funzionamento e la scalabilità OpenSearch dei cluster nel AWS cloud. Un dominio OpenSearch di servizio è sinonimo di cluster. OpenSearch I domini sono cluster con le impostazioni, i tipi di istanza, il numero di istanze e le risorse di archiviazione specificate. Amazon OpenSearch Service supporta OpenSearch un sistema operativo Elasticsearch legacy (fino alla versione 7.10, l'ultima versione open source del software). Quando crei un dominio, puoi scegliere quale motore di ricerca utilizzare.

OpenSearchè un motore di ricerca e analisi completamente open source per casi d'uso come l'analisi dei log, il monitoraggio delle applicazioni in tempo reale e l'analisi del clickstream. Per ulteriori informazioni, consulta la documentazione relativa ad OpenSearch .

Amazon OpenSearch Service fornisce tutte le risorse per il OpenSearch cluster e lo avvia. Inoltre, rileva e sostituisce automaticamente i nodi di OpenSearch servizio guasti, riducendo il sovraccarico associato alle infrastrutture autogestite. Puoi dimensionare il cluster con un'unica chiamata API o con pochi clic nella console.



Per iniziare a utilizzare OpenSearch Service, è necessario creare un dominio di OpenSearch servizio, equivalente a un cluster. OpenSearch Ogni istanza EC2 nel cluster funge da unico nodo OpenSearch di servizio.

Puoi utilizzare la console OpenSearch di servizio per configurare e configurare un dominio in pochi minuti. Se preferisci l'accesso programmatico, puoi usare il AWS CLIAWS SDKs, the o Terraform.

Caratteristiche di Amazon OpenSearch Service

OpenSearch Il servizio include le seguenti funzionalità:

Dimensionamento

- Numerose configurazioni di CPU, memoria e capacità di archiviazione, note come tipi di istanza, comprese le istanze Graviton a costi ridotti
- Supporta fino a 1002 nodi di dati
- Fino a 25 PB di storage collegato
- Archiviazione a freddo economica per UltraWarmdati di sola lettura

Sicurezza

- AWS Identity and Access Management (IAM) controllo degli accessi
- Semplice integrazione con Amazon VPC e i gruppi di sicurezza VPC
- Crittografia dei dati inattivi e node-to-node crittografia
- Autenticazione Amazon Cognito, HTTP basic o SAML per dashboard OpenSearch
- Sicurezza a livello di indice, a livello di documento e a livello di campo
- Log di verifica
- Multi-tenancy di Dashboards

Stabilità

- Più posizioni geografiche per le risorse, note come regioni e zone di disponibilità
- Allocazione dei nodi tra due o tre zone di disponibilità nella stessa regione AWS , nota come Multi-AZ
- Nodi master dedicati per l'offload delle attività di gestione del cluster
- Istantanee automatizzate per il backup e il ripristino dei domini di servizio OpenSearch

Flessibilità

- Supporto SQL per l'integrazione con applicazioni di Business Intelligence (BI)
- Pacchetti personalizzati per migliorare i risultati della ricerca

Integrazione con i servizi più diffusi

- Visualizzazione dei dati tramite dashboard OpenSearch
- Integrazione con Amazon CloudWatch per il monitoraggio delle metriche OpenSearch del dominio del servizio e l'impostazione degli allarmi
- Integrazione con AWS CloudTrail per il controllo delle chiamate API di configurazione ai domini di servizio OpenSearch
- Integrazione con Amazon S3, Amazon Kinesis e Amazon DynamoDB per caricare dati di streaming in Service OpenSearch
- Avvisi da Amazon SNS quando i dati superano determinate soglie

Quando utilizzare OpenSearch rispetto ad Amazon Service OpenSearch

Utilizza la tabella seguente per decidere se il OpenSearch servizio Amazon con provisioning o con gestione automatica OpenSearch è la scelta giusta per te.

OpenSearch	OpenSearch Servizio Amazon
• La tua organizzazione è disposta a monitorare e gestire manualmen te i cluster forniti autonomamente e dispone di persone con le competenze adeguate. • Desiderate il controllo completo a livello di compilazione del codice. • La tua organizzazione preferisce, o utilizza esclusivamente, il software open source. • Hai una strategia multi-cloud, che richiede tecnologie che non sono specifiche del fornitore.	• Non vuoi gestire, monitorare e mantenere manualmente la tua infrastruttura. • Desideri modi semplici per gestire i crescenti costi di analisi distribuendo i dati su più livelli di storage, sfruttando la durabilità e il basso costo di Amazon S3. • Vuoi sfruttare le integrazioni con altri sistemi Servizi AWS come DynamoDB, Amazon DocumentDB (con compatibilità MongoDB), IAM e. CloudWatch CloudFormation • Desideri accedere facilmente all'assistenza fornita per la manutenzione preventiva e durante i Supporto problemi di produzione.

OpenSearch	OpenSearch Servizio Amazon
• Il tuo team è in grado di risolvere qualsiasi problema critico di produzione. • Desiderate la flessibilità necessaria per utilizzare, modificare ed estendere il prodotto come preferite. • Desideri accedere immediatamente alle nuove funzionalità non appena vengono rilasciate.	• Desiderate sfruttare funzionalità come la riparazione automatica, la manutenzione proattiva, la resilienza e i backup.

Versioni supportate di Elasticsearch e OpenSearch

OpenSearch Il servizio supporta le seguenti versioni di: OpenSearch

- 3.5, 3.3, 3.1, 2.19, 2.17, 2.15, 2.13, 2.11, 2.9, 2.7, 2.5, 2.3, 1.3, 1.2, 1.1 e 1.0

OpenSearch Il servizio supporta le seguenti versioni della versione precedente di Elasticsearch:

- 7.10, 7.9, 7.8, 7.7, 7.4, 7.1, 6.8, 6.7, 6.5, 6.4, 6.3, 6.2, 6.0, 5.6, 5.5, 5.3, 5.1, 2.3 e 1.5

Si consiglia di eseguire l'aggiornamento alla OpenSearch versione più recente disponibile per utilizzare al meglio il Servizio, in termini di rapporto prezzo/prestazioni, ricchezza di funzionalità e miglioramenti della sicurezza. OpenSearch

Supporto standard ed esteso

AWS fornisce correzioni di bug e aggiornamenti di sicurezza per le versioni con supporto standard. Per le versioni con supporto esteso, AWS offre correzioni di sicurezza critiche per almeno 12 mesi dopo la fine del supporto standard, a una tariffa fissa per Normalized Instance Hour (NIH). NIH si basa sulle dimensioni dell'istanza e sulle ore di utilizzo.

I costi di assistenza estesa si applicano automaticamente quando un dominio esegue una versione che non rientra più nel supporto standard. Per evitare questi costi, esegui l'upgrade a una versione supportata.

Le tabelle seguenti mostrano la pianificazione della fine del supporto per le versioni precedenti di Elasticsearch OpenSearch e per le versioni precedenti.

OpenSearch Il servizio supporta più versioni OpenSearch e versioni legacy open source di Elasticsearch. Per alcune versioni, abbiamo già pubblicato le date di fine del supporto standard e del supporto esteso. Ti consigliamo di eseguire l'aggiornamento alla OpenSearch versione più recente disponibile per utilizzare al meglio il OpenSearch Servizio in termini di rapporto prezzo/prestazioni, ricchezza di funzionalità e miglioramenti della sicurezza. Le tabelle seguenti forniscono elenchi di Elasticsearch e delle relative OpenSearch versioni e le relative pianificazioni di supporto.

La pianificazione di fine del supporto per le versioni di Elasticsearch è la seguente:

Versione del software	Fine del supporto standard	Fine dell'Extended Support
Versioni di Elasticsearch 1.5 e 2.3	7 novembre 2025	7 novembre 2026
Versioni da 5.1 a 5.5 di Elasticsearch	7 novembre 2025	7 novembre 2026
Versioni di Elasticsearch 5.6	7 novembre 2025	7 novembre 2028
Versioni di Elasticsearch da 6.0 a 6.7	7 novembre 2025	7 novembre 2026
Versioni di Elasticsearch 6.8	Non annunciato	Non annunciato

Versione del software	Fine del supporto standard	Fine dell'Extended Support
Versioni di Elasticsearch da 7.1 a 7.8	7 novembre 2025	7 novembre 2026
Versioni di Elasticsearch 7.9	Non annunciato	Non annunciato
Versioni di Elasticsearch 7.10	Non annunciato	Non annunciato

La pianificazione della fine del supporto per OpenSearch le versioni è la seguente:

Versione del software	Fine del supporto standard	Fine dell'Extended Support
OpenSearch versioni da 1.0 a 1.2	7 novembre 2025	7 novembre 2026
OpenSearch versioni 1.3	Non annunciato	Non annunciato
OpenSearch versioni da 2.3 a 2.9	7 novembre 2025	7 novembre 2026
OpenSearch versioni 2.11 e	Non annunciato	Non annunciato

Versione del software	Fine del supporto standard	Fine dell'Extended Support
versioni successive		

Supporto standard e supporto esteso di OpenSearch Elasticsearch

AWS fornisce correzioni di bug e aggiornamenti di sicurezza regolari per le versioni coperte da Standard Support. Per le versioni con supporto esteso, AWS fornisce correzioni di sicurezza critiche per un periodo di almeno 12 mesi dopo la fine del supporto standard, a una tariffa fissa aggiuntiva per ogni Normalized Instance Hour (NIH). NIH viene calcolato come fattore della dimensione dell'istanza (ad esempio media, grande) e del numero di ore dell'istanza (per un esempio, vedi la sezione Calcolo dei costi di supporto esteso di seguito). I costi di supporto esteso vengono applicati automaticamente quando un dominio esegue una versione per la quale il supporto standard è terminato. È possibile eseguire l'aggiornamento a una versione recente ancora coperta dal supporto standard per evitare costi di supporto estesi. Per ulteriori informazioni sui costi del supporto esteso, consulta la pagina dei prezzi. Per informazioni generali sul supporto esteso, consulta le domande frequenti sul supporto esteso.

Calcolo dei costi di assistenza estesa

Ai domini che eseguono versioni con supporto esteso verrà addebitato un importo fisso aggiuntivo per ora di fee/Normalized istanza (NIH), ad esempio 0,0065 USD nella regione Stati Uniti orientali (Virginia del Nord). Il NIH viene calcolato come fattore tra la dimensione dell'istanza (ad esempio, media, grande) e il numero di ore dell'istanza. Ad esempio, se utilizzi un'istanza m7g.medium.search per 24 ore nella regione Stati Uniti orientali (Virginia del Nord), al prezzo di 0,068 USD/ora di istanza (su richiesta), in genere pagherai 1,632 USD (0,068 USD/24 USD). Se utilizzi una versione con supporto esteso, pagherai 0,0065 USD/NIH aggiuntivi, calcolati come 0,0065 USD x 24 (numero di ore di istanza) x 2 (fattore di normalizzazione delle dimensioni; 2 per istanze di medie dimensioni), che equivale a 0,312 USD per il supporto esteso per 24 ore. L'importo totale che pagherai per 24 ore sarà la somma del costo di utilizzo standard dell'istanza e del costo del supporto esteso, pari a 1,944 USD (1,632 USD+0,312 USD). La tabella seguente mostra il fattore di normalizzazione per istanze di varie dimensioni in Service. OpenSearch

Dimensioni istanza	Fattore di normalizzazione	
nano	0,25	
micro	0,5	
small	1	
medium	2	
large	4	
xlarge	8	
2xlarge	16	
4xlarge	32	
8xlarge	64	
9xlarge	72	
10xlarge	80	
12xlarge	96	
16xlarge	128	
18xlarge	144	
24xlarge	192	
32xlarge	256	

Prezzi per Amazon OpenSearch Service

Per il OpenSearch servizio, paghi per ogni ora di utilizzo di un'istanza EC2 e per la dimensione cumulativa di tutti i volumi di storage EBS collegati alle tue istanze. Si applicano anche le tariffe standard per il trasferimento AWS dei dati.

Tuttavia, esistono alcune eccezioni di trasferimento dati notevoli. Se un dominio utilizza più zone di disponibilità, il OpenSearch servizio non fattura il traffico tra le zone di disponibilità. Un significativo trasferimento di dati avviene all'interno di un dominio durante l'allocazione e il ribilanciamento degli shard. OpenSearch Non servono né contatori né bollette per questo traffico. Allo stesso modo, il OpenSearch Servizio non fattura il trasferimento di dati tra UltraWarm/cold nodes e Amazon S3.

Per i dettagli completi sui prezzi, consulta i prezzi OpenSearch di Amazon Service. Per informazioni sui costi addebitati durante le modifiche di configurazione, consulta the section called “Costi per le modifiche di configurazione”.

Servizi correlati

OpenSearch Il servizio viene comunemente utilizzato con i seguenti servizi:

Amazon CloudWatch

OpenSearch I domini di servizio inviano automaticamente le metriche a CloudWatch in modo da poter monitorare lo stato e le prestazioni del dominio. Per ulteriori informazioni, consulta Monitoraggio delle metriche dei OpenSearch cluster con Amazon CloudWatch.

CloudWatch I log possono anche andare nella direzione opposta. È possibile configurare CloudWatch i registri per lo streaming dei dati al OpenSearch Servizio per l'analisi. Per ulteriori informazioni, consulta the section called “Caricamento di dati di streaming da Amazon CloudWatch”.

AWS CloudTrail

Utilizzalo AWS CloudTrail per ottenere una cronologia delle chiamate all'API di configurazione del OpenSearch servizio e degli eventi correlati per il tuo account. Per ulteriori informazioni, consulta Monitoraggio delle chiamate API di Amazon OpenSearch Service con AWS CloudTrail.

Amazon Kinesis

Kinesis è un servizio gestito per l'elaborazione in tempo reale dei dati di streaming su vasta scala. Per ulteriori informazioni, consultare the section called “Caricamento dei dati in streaming in Amazon Kinesis Data Streams” e the section called “Caricamento di dati di streaming da Amazon Data Firehose”.

Amazon S3

Amazon Simple Storage Service (Amazon S3) è un servizio di archiviazione su Internet. Questa guida fornisce il codice di esempio Lambda per l'integrazione con Amazon S3. Per ulteriori informazioni, consultare the section called “Caricamento di dati in streaming da Amazon S3”.

AWS IAM

AWS Identity and Access Management (IAM) è un servizio web che puoi utilizzare per gestire l'accesso ai tuoi domini OpenSearch di servizio. Per ulteriori informazioni, consulta the section called “Identity and Access Management”.

AWS Lambda

AWS Lambda è un servizio di elaborazione che consente di eseguire codice senza effettuare il provisioning o la gestione di server. Questa guida fornisce il codice di esempio Lambda per lo streaming dei dati da DynamoDB, Amazon S3 e Kinesis. Per ulteriori informazioni, consultare the section called “Caricamento di dati di streaming in OpenSearch Service”.

Amazon DynamoDB

Amazon DynamoDB è un servizio di database NoSQL interamente gestito che combina prestazioni elevate e prevedibili con una scalabilità ottimale. Per ulteriori informazioni sullo streaming di dati su OpenSearch Service, consulta. the section called “Caricamento di dati in streaming da una tabella Amazon DynamoDB”

Amazon Quick

Puoi visualizzare i dati di OpenSearch Service utilizzando i dashboard Quick. Per ulteriori informazioni, consulta Using Amazon OpenSearch Service with Quick nella Guida rapida per l'utente.

> **Note**
>
> OpenSearch include alcuni codici Elasticsearch con licenza Apache di Elasticsearch B.V. e altro codice sorgente. Elasticsearch B.V. non è la fonte di quell'altro codice sorgente. ELASTICSEARCH è un marchio registrato di Elasticsearch B.V.

Configurazione di Amazon OpenSearch Service

Concessione delle autorizzazioni

Negli ambienti di produzione, si consiglia di utilizzare politiche più dettagliate. Per ulteriori informazioni sulla gestione degli accessi, consulta Access management for AWS resources nella IAM User Guide.

Per fornire l'accesso, aggiungi autorizzazioni agli utenti, gruppi o ruoli:

- Utenti e gruppi in AWS IAM Identity Center:

  Crea un set di autorizzazioni. Segui le istruzioni riportate nella pagina Create a permission set (Creazione di un set di autorizzazioni) nella Guida per l'utente di AWS IAM Identity Center .
- Utenti gestiti in IAM tramite un provider di identità:

  Crea un ruolo per la federazione delle identità. Segui le istruzioni riportate nella pagina Create a role for a third-party identity provider (federation) della Guida per l'utente IAM.
- Utenti IAM:
  - Crea un ruolo che l'utente possa assumere. Segui le istruzioni riportate nella pagina Create a role for an IAM user della Guida per l'utente IAM.
  - (Non consigliato) Collega una policy direttamente a un utente o aggiungi un utente a un gruppo di utenti. Segui le istruzioni riportate nella pagina Aggiunta di autorizzazioni a un utente (console) nella Guida per l'utente IAM.

Concessione dell'accesso programmatico

Gli utenti necessitano di un accesso programmatico se desiderano interagire con utenti AWS esterni a. Console di gestione AWS Il modo per concedere l'accesso programmatico dipende dal tipo di utente che accede. AWS

Per fornire agli utenti l'accesso programmatico, scegli una delle seguenti opzioni.

Quale utente necessita dell'accesso programmatico?	Per	Come
IAM	(Consigliato) Utilizza le credenziali della console come credenziali temporanee per firmare le richieste programmatiche a,, o. AWS CLI AWS SDKs AWS APIs	Segui le istruzioni per l'interfaccia che desideri utilizzare. • Per la AWS CLI, consulta Login for AWS local development nella Guida per l'AWS Command Line Interface utente. • Per AWS SDKs, consulta Login for AWS local development nella AWS SDKs and Tools Reference Guide.
Identità della forza lavoro (Utenti gestiti nel centro identità IAM)	Utilizza credenziali temporanee per firmare le richieste programmatiche a AWS CLI, AWS SDKs, o. AWS APIs	Segui le istruzioni per l'interfaccia che desideri utilizzare. • Per la AWS CLI, vedere Configurazione dell'uso AWS IAM Identity Center nella AWS CLI Guida per l'utente.AWS Command Line Interface • Per AWS SDKs gli strumenti e AWS APIs, consulta l'autenticazione di IAM Identity Center nella Guida di riferimento AWS SDKs and Tools.
IAM	Utilizza credenziali temporanee per firmare le richieste programmatiche a AWS CLI, AWS SDKs, o. AWS APIs	Seguendo le istruzioni riportate in Utilizzo delle credenziali temporanee con le

Quale utente necessita dell'accesso programmatico?	Per	Come
		AWS risorse nella Guida per l'utente IAM.
IAM	(Non consigliato) Utilizza credenziali a lungo termine per firmare richieste programmatiche a AWS CLI,, AWS SDKs o. AWS APIs	Segui le istruzioni per l'interfaccia che desideri utilizzare. • Per la AWS CLI, consulta Autenticazione tramite credenziali utente IAM nella Guida per l'utente.AWS Command Line Interface • Per gli strumenti AWS SDKs e gli strumenti, consulta Autenticazione tramite credenziali a lungo termine nella Guida di riferimento agli strumenti e agli AWS SDKs strumenti. • Per AWS APIs, consulta la sezione Gestione delle chiavi di accesso per gli utenti IAM nella Guida per l'utente IAM.

Installa e configura il AWS CLI

Se desideri utilizzare il OpenSearch Servizio APIs, devi installare la versione più recente di AWS Command Line Interface (AWS CLI). Non è necessario AWS CLI utilizzare il OpenSearch servizio dalla console e puoi iniziare senza la CLI seguendo la procedura riportata di seguito. Guida introduttiva ad Amazon OpenSearch Service

Per configurare il AWS CLI

1. Per installare la versione più recente di AWS CLI per macOS, Linux o Windows, vedi Installazione o aggiornamento della versione più recente di. AWS CLI
2. Per configurare AWS CLI la configurazione sicura dell'accesso a Servizi AWS, incluso il OpenSearch servizio, vedi Configurazione rapida con `aws configure`.
3. Per verificare la configurazione, immettete il seguente DataBrew comando al prompt dei comandi.

   ```
   aws opensearch help
   ```

   AWS CLI i comandi utilizzano l'impostazione predefinita Regione AWS della configurazione, a meno che non venga impostata con un parametro o un profilo. Per impostare Regione AWS un parametro, è possibile aggiungere il `--region` parametro a ciascun comando.

   Per impostare Regione AWS un profilo, aggiungi prima un profilo denominato nel `~/.aws/config` file o nel `%UserProfile%/.aws/config` file (per Microsoft Windows). Segui la procedura descritta in Profili denominati per AWS CLI. Successivamente, imposta le tue Regione AWS e le altre impostazioni con un comando simile a quello illustrato nell'esempio seguente.

   ```
   [profile opensearch]
   aws_access_key_id = ACCESS-KEY-ID-OF-IAM-USER
   aws_secret_access_key = SECRET-ACCESS-KEY-ID-OF-IAM-USER
   region = us-east-1
   output = text
   ```

Aprire la console

La maggior parte degli argomenti relativi alla console di questa sezione inizia dalla console di OpenSearch servizio. Se non hai già effettuato l'accesso al tuo Account AWS, accedi, quindi apri la console di OpenSearch servizio e passa alla sezione successiva per continuare a usare Service. OpenSearch

Guida introduttiva ad Amazon OpenSearch Service

Per iniziare, registrarsi per ottenere un Account AWS se non ne è già disponibile uno. Dopo aver configurato un account, completa il tutorial introduttivo per Amazon OpenSearch Service. Per ulteriori informazioni mentre apprendi come usare il servizio, consulta gli argomenti introduttivi seguenti:

- Crea un dominio.
- Dimensiona il dominio in modo appropriato per il tuo carico di lavoro.
- Controlla l'accesso al tuo dominio utilizzando una politica di accesso al dominio o un controllo granulare degli accessi.
- Indicizza i dati manualmente o da altri servizi. AWS
- Usa OpenSearch le dashboard per cercare i dati e creare visualizzazioni.
- Scopri di più sulle opzioni avanzate per la creazione di un dominio. Per ulteriori informazioni, consultare Creazione e gestione dei domini.
- Scopri come gestire gli indici nel dominio. Per ulteriori informazioni, consulta Gestione degli indici.
- Prova uno dei tutorial per lavorare con Amazon OpenSearch Service. Per ulteriori informazioni, consulta Esercitazioni.

Per informazioni sulla migrazione a OpenSearch Service da un cluster autogestito OpenSearch , consulta. the section called “Migrazione al servizio OpenSearch ”

Per informazioni più dettagliate, consulta Creazione e gestione dei domini e altri argomenti all'interno di questa guida. Per informazioni sulla migrazione a OpenSearch Service da un cluster autogestito OpenSearch , vedere. the section called “Migrazione al servizio OpenSearch ”

Puoi completare i seguenti passaggi utilizzando la console di OpenSearch servizio AWS CLI, l'o l' AWS SDK. Per informazioni sull'installazione e la configurazione di AWS CLI, consulta la Guida per l'AWS Command Line Interface utente.

Crea un dominio Amazon OpenSearch Service

> **Important**
>
> Questo è un breve tutorial per configurare un dominio Amazon OpenSearch Service di prova. Non utilizzare questo processo per creare domini di produzione. Per una versione completa della stessa procedura, consulta Creazione e gestione dei domini.

Un dominio OpenSearch di servizio è sinonimo di cluster. OpenSearch I domini sono cluster con le impostazioni, i tipi di istanza, il numero di istanze e le risorse di archiviazione specificate. È possibile creare un dominio OpenSearch di servizio utilizzando la console, il AWS CLI, o il AWS SDKs.

Per creare un dominio OpenSearch di servizio utilizzando la console

1. Vai a https://aws.amazon.come scegli Accedi alla console.
2. In Analytics, scegli Amazon OpenSearch Service.
3. Scegli Crea dominio.
4. Specificare un nome per il dominio. Gli esempi di questo tutorial utilizzano il nome movies.
5. Per il metodo di creazione del dominio, scegli Standard create.

   > **Note**
   >
   > Per configurare rapidamente un dominio di produzione con le migliori pratiche, puoi scegliere Easy create. Per lo sviluppo e il test di questo tutorial, utilizzeremo Standard create.

6. Per i modelli, scegli Dev/test.
7. Per l'opzione di distribuzione, scegli Dominio con standby.
8. Per Versione, scegliere la versione più recente.
9. Per ora, ignora le sezioni Data nodes, Warm and cold data storage, Dedicated master node, Snapshot configuration e Custom endpoint.
10. Per semplicità in questo tutorial, viene utilizzato un dominio ad accesso pubblico. In Rete, scegli Accesso pubblico.

11. Nelle impostazioni di controllo granulare degli accessi, mantieni selezionata la casella di controllo Abilita controllo granulare degli accessi. Seleziona Crea utente principale e fornisci un nome utente e una password.
12. Per adesso, ignorare le sezioni Autenticazione SAML e Autenticazione Amazon Cognito.
13. Per Policy di accesso, scegli Utilizza solo controllo granulare degli accessi. In questa esercitazione, il controllo granulare degli accessi gestisce l'autenticazione, non la policy di accesso al dominio.
14. Ignora il resto delle impostazioni per il momento e scegli Crea. I nuovi domini richiedono in genere 15-30 minuti per l'inizializzazione, ma possono richiedere più tempo a seconda della configurazione. Dopo l'inizializzazione del dominio, selezionarlo per aprire il riquadro di configurazione. Annotare l'endpoint del dominio sotto Informazioni generali (ad esempio,`https://search-my-domain.us-east-1.es.amazonaws.com`), che potrai utilizzare nel prossimo passaggio.

Avanti: carica i dati su un dominio OpenSearch di servizio per l'indicizzazione

Carica i dati su Amazon OpenSearch Service per l'indicizzazione

> **Important**
>
> Questo è un breve tutorial per caricare una piccola quantità di dati di test su Amazon Service. OpenSearch Per ulteriori informazioni sul caricamento dei dati in un dominio di produzione, consultare Indicizzazione dei dati.

Puoi caricare dati in un dominio di OpenSearch servizio utilizzando la riga di comando o la maggior parte dei linguaggi di programmazione.

Per brevità e comodità, le richieste di esempio seguenti utilizzano curl, un comune client HTTP. I client come curl non sono in grado di eseguire la firma della richiesta, necessaria se la policy d'accesso specifica utenti o ruoli IAM. Per completare correttamente questo processo, è necessario utilizzare un controllo granulare degli accessi con un nome utente e una password principali, come configurato nella Fase 1.

È possibile installare curl su Windows e utilizzarlo dal prompt dei comandi, ma consigliamo di utilizzare uno strumento come Cygwin o Windows Subsystem for Linux. I sistemi macOS e la maggior parte delle distribuzioni Linux includono già curl.

Opzione 1: Caricamento di un singolo documento

Esegui il comando seguente per aggiungere un singolo documento al dominio movies:

```
curl -XPUT -u 'master-user:master-user-password' 'domain-endpoint/movies/_doc/1' -d
 '{"director": "Burton, Tim", "genre": ["Comedy","Sci-Fi"], "year": 1996, "actor":
 ["Jack Nicholson","Pierce Brosnan","Sarah Jessica Parker"], "title": "Mars Attacks!"}'
 -H 'Content-Type: application/json'
```

Nel comando, fornisci il nome utente e la password che hai creato nel passaggio 1.

Per una spiegazione dettagliata di questo comando e su come effettuare richieste firmate al OpenSearch Servizio, vedereIndicizzazione dei dati.

Opzione 2: Caricamento di più documenti

Per caricare un file JSON contenente più documenti in un dominio di OpenSearch servizio

1. Creare un file locale denominato `bulk_movies.json`. Copiare e incollare il seguente contenuto nel file e aggiungere un carattere newline finale:

```
{ "index" : { "_index": "movies", "_id" : "2" } }
{"director": "Frankenheimer, John", "genre": ["Drama", "Mystery", "Thriller",
 "Crime"], "year": 1962, "actor": ["Lansbury, Angela", "Sinatra, Frank", "Leigh,
 Janet", "Harvey, Laurence", "Silva, Henry", "Frees, Paul", "Gregory, James",
 "Bissell, Whit", "McGiver, John", "Parrish, Leslie", "Edwards, James", "Flowers,
 Bess", "Dhiegh, Khigh", "Payne, Julie", "Kleeb, Helen", "Gray, Joe", "Nalder,
 Reggie", "Stevens, Bert", "Masters, Michael", "Lowell, Tom"], "title": "The
 Manchurian Candidate"}
{ "index" : { "_index": "movies", "_id" : "3" } }
{"director": "Baird, Stuart", "genre": ["Action", "Crime", "Thriller"], "year":
 1998, "actor": ["Downey Jr., Robert", "Jones, Tommy Lee", "Snipes, Wesley",
 "Pantoliano, Joe", "Jacob, Ir\u00e8ne", "Nelligan, Kate", "Roebuck, Daniel",
 "Malahide, Patrick", "Richardson, LaTanya", "Wood, Tom", "Kosik, Thomas",
 "Stellate, Nick", "Minkoff, Robert", "Brown, Spitfire", "Foster, Reese",
 "Spielbauer, Bruce", "Mukherji, Kevin", "Cray, Ed", "Fordham, David", "Jett,
 Charlie"], "title": "U.S. Marshals"}
{ "index" : { "_index": "movies", "_id" : "4" } }
{"director": "Ray, Nicholas", "genre": ["Drama", "Romance"], "year": 1955, "actor":
 ["Hopper, Dennis", "Wood, Natalie", "Dean, James", "Mineo, Sal", "Backus, Jim",
 "Platt, Edward", "Ray, Nicholas", "Hopper, William", "Allen, Corey", "Birch,
 Paul", "Hudson, Rochelle", "Doran, Ann", "Hicks, Chuck", "Leigh, Nelson",
```

```
 "Williams, Robert", "Wessel, Dick", "Bryar, Paul", "Sessions, Almira", "McMahon,
 David", "Peters Jr., House"], "title": "Rebel Without a Cause"}
```

2. Eseguire il comando seguente nella directory locale in cui è memorizzato il file per caricarlo nel dominio movies:

```
curl -XPOST -u 'master-user:master-user-password' 'domain-endpoint/movies/_bulk' --
data-binary @bulk_movies.json -H 'Content-Type: application/x-ndjson'
```

Per ulteriori informazioni sul formato di file bulk, consulta Indicizzazione dei dati

Successivo: Ricerca di documenti

Cerca documenti in Amazon OpenSearch Service

Per cercare documenti in un dominio Amazon OpenSearch Service, utilizza l'API OpenSearch di ricerca. In alternativa, puoi utilizzare OpenSearch le dashboard per cercare documenti nel dominio.

Come eseguire la ricerca di documenti dalla riga di comando

Esegui il comando seguente per cercare la parola mars nel dominio movies:

```
curl -XGET -u 'master-user:master-user-password' 'domain-endpoint/movies/_search?
q=mars&pretty=true'
```

Se nella pagina precedente hai utilizzato il blocco di dati, prova a cercare invece rebel.

Noterai una risposta simile alla seguente:

```
{
  "took" : 5,
  "timed_out" : false,
  "_shards" : {
    "total" : 5,
    "successful" : 5,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
```

```
    "total" : {
      "value" : 1,
      "relation" : "eq"
    },
    "max_score" : 0.2876821,
    "hits" : [
      {
        "_index" : "movies",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 0.2876821,
        "_source" : {
          "director" : "Burton, Tim",
          "genre" : [
            "Comedy",
            "Sci-Fi"
          ],
          "year" : 1996,
          "actor" : [
            "Jack Nicholson",
            "Pierce Brosnan",
            "Sarah Jessica Parker"
          ],
          "title" : "Mars Attacks!"
        }
      }
    ]
  }
}
```

Cerca documenti utilizzando le dashboard OpenSearch

OpenSearch Dashboards è un popolare strumento di visualizzazione open source progettato per funzionare con. OpenSearch Fornisce un'interfaccia utente utile per cercare e monitorare gli indici.

Per cercare documenti da un dominio di OpenSearch servizio utilizzando Dashboards

1. Vai all'URL delle OpenSearch dashboard del tuo dominio. Puoi trovare l'URL nella dashboard del dominio nella console di OpenSearch servizio. L'URL segue il seguente formato:

   ```
   domain-endpoint/_dashboards/
   ```

2. Accedi utilizzando il nome utente e la password principali.

3. Per utilizzare Dashboards, è necessario creare almeno un modello di indice. Dashboards usa questi modelli per identificare gli indici da analizzare. Aprire il riquadro di spostamento a sinistra, scegliere Gestione stack, scegliere Modelli di indice, quindi scegliere Crea modello di indice. Per questo tutorial, digita movies.
4. Scegliere Fase successiva quindi selezionare Crea modello di indice. Dopo aver creato il modello, è possibile visualizzare i vari campi del documento, ad esempio `actor` e `director`.
5. Tornare alla pagina Modelli di indice e assicurarsi che `movies` sia impostato come valore di default. In caso contrario, seleziona il modello e scegli l'icona a forma di stella per renderlo predefinito.
6. Per iniziare a cercare i dati, aprire di nuovo il riquadro di spostamento a sinistra e scegliere Individua.
7. Nella barra di ricerca, inserire mars se è stato caricato un singolo documento oppure rebel se sono stati caricati più documenti, quindi premere Invio. È possibile provare a cercare altri termini, come i nomi di attori o registi.

Successivo: Eliminazione di un dominio

Eliminare un dominio Amazon OpenSearch Service

Poiché il dominio movies utilizzato in questo tutorial è solo a scopo di test, è consigliabile eliminarlo una volta terminate le prove in modo da evitare costi aggiuntivi.

Per eliminare un dominio OpenSearch di servizio dalla console

1. Accedi alla console di Amazon OpenSearch Service.
2. In Domini, seleziona il dominio movies.
3. Scegli Elimina e conferma l'eliminazione.

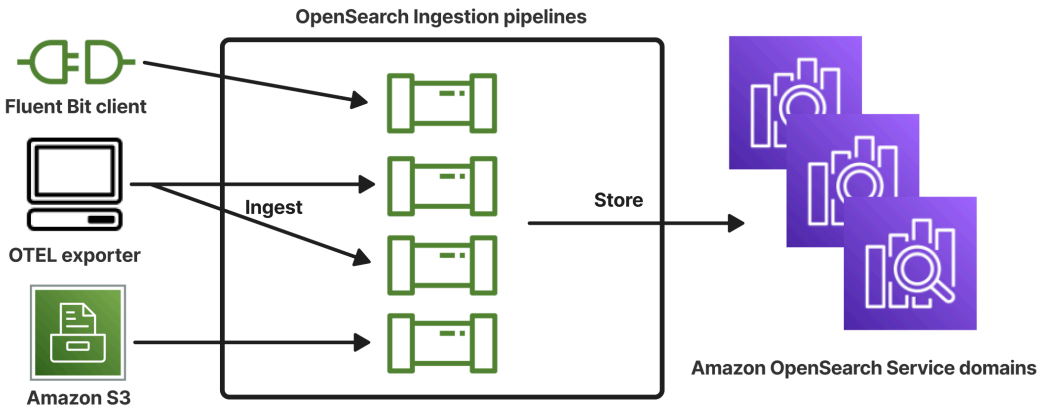
Panoramica di Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion è un raccoglitore di dati serverless completamente gestito che trasmette log, metriche e dati di tracciamento in tempo reale ai domini di Amazon Service e alle raccolte Serverless. OpenSearch OpenSearch

Con OpenSearch Ingestion, non hai più bisogno di strumenti di terze parti come Logstash o Jaeger per importare dati. Puoi configurare i tuoi produttori di dati per inviare dati a OpenSearch Ingestion, che li consegna automaticamente al dominio o alla raccolta specificati. Puoi anche trasformare i dati prima della consegna.

Poiché OpenSearch Ingestion è serverless, non è necessario gestire l'infrastruttura, applicare patch al software o scalare i cluster manualmente. È possibile effettuare il provisioning delle pipeline di ingestione direttamente in e Ingestion si occupa del resto Console di gestione AWS. OpenSearch

Come componente di Amazon OpenSearch Service, OpenSearch Ingestion si basa su Data Prepper, un raccoglitore di dati open source che filtra, arricchisce, trasforma, normalizza e aggrega i dati per l'analisi e la visualizzazione a valle.



Concetti chiave di Amazon OpenSearch Ingestion

Prima di iniziare a usare OpenSearch Ingestion, è utile comprendere questi concetti chiave.

Pipeline

Dal punto di vista OpenSearch dell'ingestione, una pipeline si riferisce a un unico raccoglitore di dati fornito che viene creato all'interno di Service. OpenSearch Puoi considerarlo come l'intero file di configurazione YAML, che include una o più pipeline secondarie. Per i passaggi per creare una pipeline di ingestione, consulta. the section called “Creazione di pipeline”

Pipeline secondaria

Le sotto-pipeline vengono definite all'interno di un file di configurazione YAML. Ogni sub-pipeline è una combinazione di una sorgente, un buffer, zero o più processori e uno o più sink. È possibile definire più sotto-pipeline in un unico file YAML, ognuna con sorgenti, processori e sink unici. Per facilitare il monitoraggio con CloudWatch e altri servizi, ti consigliamo di specificare un nome di pipeline diverso da tutte le relative pipeline secondarie.

Puoi mettere insieme più subpipeline all'interno di un singolo file YAML, in modo che l'origine di una sottopipeline sia un'altra sottopipeline e il relativo sink sia una terza sottopipeline. Per vedere un esempio, consulta the section called “OpenTelemetry Collettore”.

Origine

Il componente di input di una sottopipeline. Definisce il meccanismo attraverso il quale una pipeline consuma i record. La fonte può consumare eventi ricevendoli tramite HTTPS o leggendo da endpoint esterni come Amazon S3. Esistono due tipi di fonti: basate su push e basate su pull. Le fonti basate su push, come HTTP e OTel logs, trasmettono i record agli endpoint di ingestione. Le fonti basate su pull, come OTel trace e S3, estraggono i dati dalla fonte.

Processori

Unità di elaborazione intermedie in grado di filtrare, trasformare e arricchire i record nel formato desiderato prima di pubblicarli nel sink. Il processore è un componente opzionale di una pipeline. Se non si definisce un processore, i record vengono pubblicati nel formato definito nel codice sorgente. È possibile avere più di un processore. Una pipeline esegue i processori nell'ordine in cui vengono definiti dall'utente.

Sink

Il componente di output di una sottopipeline. Definisce una o più destinazioni in cui una sub-pipeline pubblica i record. OpenSearch Ingestion supporta OpenSearch i domini di servizio come sink. Supporta anche le condutture secondarie come sink. Ciò significa che è possibile mettere insieme più sotto-pipeline all'interno di una singola pipeline di OpenSearch ingestione (file YAML). I cluster OpenSearch autogestiti non sono supportati come sink.

Buffer

La parte del processore che funge da strato tra la sorgente e il sink. Non è possibile configurare manualmente un buffer all'interno della pipeline. OpenSearch L'ingestione utilizza una configurazione di buffer predefinita.

Route

La parte di un processore che consente agli autori della pipeline di inviare solo eventi che soddisfano determinate condizioni a diversi sink.

Una definizione di subpipeline valida deve contenere una fonte e un sink. Per ulteriori informazioni su ciascuno di questi elementi della pipeline, consultate il riferimento alla configurazione.

Vantaggi di Amazon OpenSearch Ingestion

OpenSearch L'ingestione presenta i seguenti vantaggi principali:

- Elimina la necessità di gestire manualmente una pipeline autoalimentata.
- Ridimensiona automaticamente le pipeline in base ai limiti di capacità definiti dall'utente.
- Mantiene la pipeline aggiornata con patch di sicurezza e bug.
- Offre la possibilità di collegare le pipeline al cloud privato virtuale (VPC) per un ulteriore livello di sicurezza.
- Consente di interrompere e avviare le pipeline per controllare i costi.
- Fornisce modelli di configurazione delle pipeline per i casi d'uso più diffusi per aiutarvi a iniziare a lavorare più velocemente.
- Consente di interagire a livello di codice con le pipeline tramite le varie AWS SDKs API e l'API Ingestion. OpenSearch
- Supporta il monitoraggio delle prestazioni in Amazon CloudWatch e la registrazione degli errori nei CloudWatch log.

Limitazioni di Amazon OpenSearch Ingestion

OpenSearch Ingestion presenta le seguenti limitazioni:

- Puoi importare dati solo in domini con OpenSearch versione 1.0 o successiva oppure Elasticsearch 6.8 o versione successiva. Se utilizzi la fonte di OTel traccia, ti consigliamo di utilizzare Elasticsearch 7.9 o versione successiva in modo da poter utilizzare il plug-in Dashboards. OpenSearch
- Se una pipeline sta scrivendo su un dominio di OpenSearch servizio all'interno di un VPC, la pipeline deve essere creata nello Regione AWS stesso dominio.

- È possibile configurare solo una singola origine dati all'interno di una definizione di pipeline.
- Non è possibile specificare OpenSearch cluster autogestiti come sink.
- Non è possibile specificare un endpoint personalizzato come sink. Puoi comunque scrivere su un dominio con endpoint personalizzati abilitati, ma devi specificarne l'endpoint standard.
- Non puoi specificare risorse all'interno delle Regioni opt-in come sorgenti o sink.
- Esistono alcuni vincoli sui parametri che è possibile includere in una configurazione di pipeline. Per ulteriori informazioni, consulta the section called “Requisiti e vincoli di configurazione”.

Versioni di Data Prepper supportate

OpenSearch Attualmente Ingestion supporta le seguenti versioni principali di Data Prepper:

- 2.x

Quando create una pipeline utilizzando l'editor di codice, utilizzate l'`version`opzione richiesta per specificare la versione principale di Data Prepper da utilizzare. Ad esempio,. `version: "2"` OpenSearch Ingestion recupera l'ultima versione secondaria supportata di quella versione principale e fornisce la pipeline con quella versione.

Se non utilizzate l'editor di codice per creare la pipeline, OpenSearch Ingestion esegue automaticamente il provisioning della pipeline con l'ultima versione supportata.

Attualmente, OpenSearch Ingestion fornisce le pipeline con la versione 2.7 di Data Prepper. Per informazioni, consulta le note di rilascio della versione 2.7. Non tutte le versioni secondarie di una particolare versione principale sono supportate da OpenSearch Ingestion.

Quando si aggiorna la configurazione di una pipeline, se è disponibile il supporto per una nuova versione secondaria di Data Prepper, OpenSearch Ingestion aggiorna automaticamente la pipeline all'ultima versione secondaria supportata della versione principale specificata nella configurazione della pipeline. Ad esempio, potreste avere `version: "2"` nella configurazione della pipeline e OpenSearch Ingestion inizialmente aveva fornito alla pipeline la versione 2.6.0. Quando viene aggiunto il supporto per la versione 2.7.0 e si apporta una modifica alla configurazione della pipeline, Ingestion aggiorna la pipeline alla versione 2.7.0. OpenSearch Questo processo mantiene la pipeline aggiornata con le ultime correzioni di bug e i miglioramenti delle prestazioni. OpenSearch Ingestion non può aggiornare la versione principale della pipeline a meno che non si modifichi manualmente l'`version`opzione all'interno della configurazione della pipeline. Per ulteriori informazioni, consulta the section called “Aggiornamento delle pipeline”.

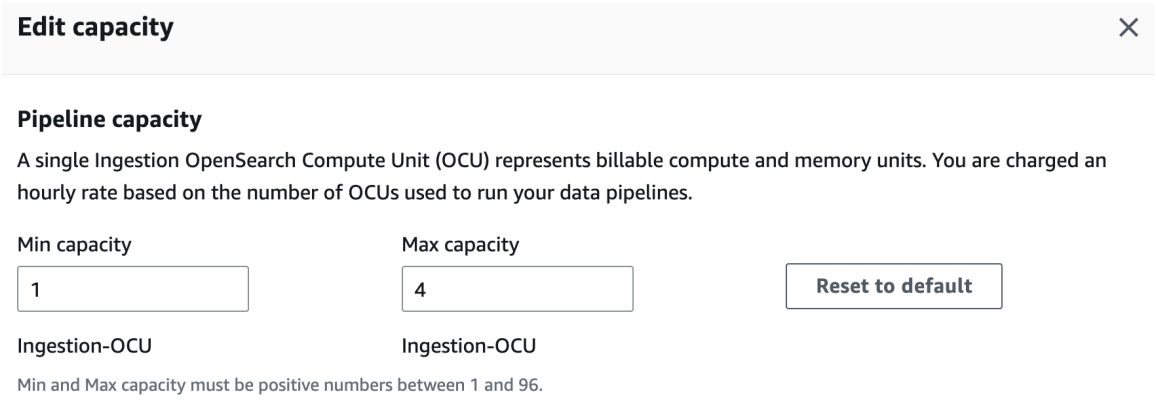
Scalabilità delle pipeline in Amazon Ingestion OpenSearch

OpenSearch Ingestion ridimensiona automaticamente la capacità della pipeline in base alle unità di calcolo di ingestione (Ingestion) minime e massime specificate. OpenSearch OCUs Ciò elimina la necessità di provisioning e gestione manuali.

Ogni OCU Ingestion è una combinazione di circa 15 GiB di memoria e 2 v. CPUs È possibile specificare i valori OCU minimi e massimi per una pipeline e OpenSearch Ingestion ridimensiona automaticamente la capacità della pipeline in base a questi limiti.

Quando create una pipeline, specificate i seguenti valori:

- Capacità minima: la pipeline può ridurre la capacità fino a questo numero di ingestioni. OCUs La capacità minima specificata è anche la capacità iniziale di una pipeline.
- Capacità massima: la pipeline può aumentare la capacità fino a questo numero di ingestioni. OCUs



Assicurati che la capacità massima di una pipeline sia sufficientemente elevata da gestire i picchi di carico di lavoro e che la capacità minima sia sufficientemente bassa da ridurre al minimo i costi quando la pipeline non è occupata. In base alle impostazioni, OpenSearch Ingestion ridimensiona automaticamente il numero di Ingestion per consentire alla pipeline di elaborare il carico di lavoro di importazione. OCUs In qualsiasi momento specifico, ti vengono addebitati solo i costi di OCUs Ingestion utilizzati attivamente dalla tua pipeline.

La capacità allocata alla pipeline di OpenSearch Ingestion aumenta e diminuisce in base ai requisiti di elaborazione della pipeline e al carico generato dall'applicazione client. Quando la capacità è limitata, OpenSearch Ingestion aumenta allocando più unità di calcolo (GiB di memoria). Quando la pipeline elabora carichi di lavoro più piccoli o non elabora affatto i dati, può essere ridimensionata fino al livello minimo di ingestione configurato. OCUs

È possibile specificare un minimo di 1 OCU di ingestione, un massimo di 96 ingestioni per le pipeline stateless e un massimo di 48 Ingestion OCUs per le pipeline stateful. OCUs Si consiglia un minimo di 2 Ingestion per sorgenti basate su push. OCUs Quando il buffering persistente è abilitato, è possibile specificare un minimo di 2 e un massimo di 384 Ingestion. OCUs

Data una pipeline di log standard con un'unica fonte, un pattern grok semplice e un sink, ogni unità di elaborazione può supportare fino a 2 MiB al secondo. Per pipeline di log più complesse con più processori, ogni unità di calcolo potrebbe supportare un carico di importazione inferiore. In base alla capacità della pipeline e all'utilizzo delle risorse, inizia il processo di scalabilità di OpenSearch Ingestion.

Per garantire un'elevata disponibilità, gli ingestion OCUs sono distribuiti tra le zone di disponibilità (). AZs Il numero di AZs dipende dalla capacità minima specificata.

Ad esempio, se si specifica un minimo di 2 unità di calcolo, le Ingestion OCUs in uso in un dato momento vengono distribuite uniformemente su 2. AZs Se si specifica un minimo di 3 o più unità di calcolo, le Ingestion OCUs vengono distribuite uniformemente su 3. AZs Ti consigliamo di effettuare il provisioning di almeno due Ingestion OCUs per garantire una disponibilità del 99,9% per le tue pipeline di importazione.

Non ti viene addebitata la fattura per Ingestion OCUs quando una pipeline si trova negli stati,, e. `Create failed` `Creating` `Deleting` `Stopped`

Per istruzioni su come configurare e recuperare le impostazioni di capacità per una pipeline, consulta. the section called “Creazione di pipeline”

OpenSearch Prezzi di ingestione

In qualsiasi momento specifico, paghi solo per il numero di ingestioni assegnate a una pipeline, indipendentemente dal fatto OCUs che i dati scorrano attraverso la pipeline. OpenSearch Ingestion soddisfa immediatamente i carichi di lavoro aumentando o diminuendo la capacità della pipeline in base all'utilizzo.

Per i dettagli completi sui prezzi, consulta i prezzi OpenSearch di Amazon Service.

Supportato Regioni AWS

OpenSearch L'ingestione è disponibile in un sottoinsieme di Regioni AWS tale OpenSearch servizio disponibile in. Per un elenco delle regioni supportate, consulta gli endpoint e le quote di Amazon OpenSearch Service nel. Riferimenti generali di AWS

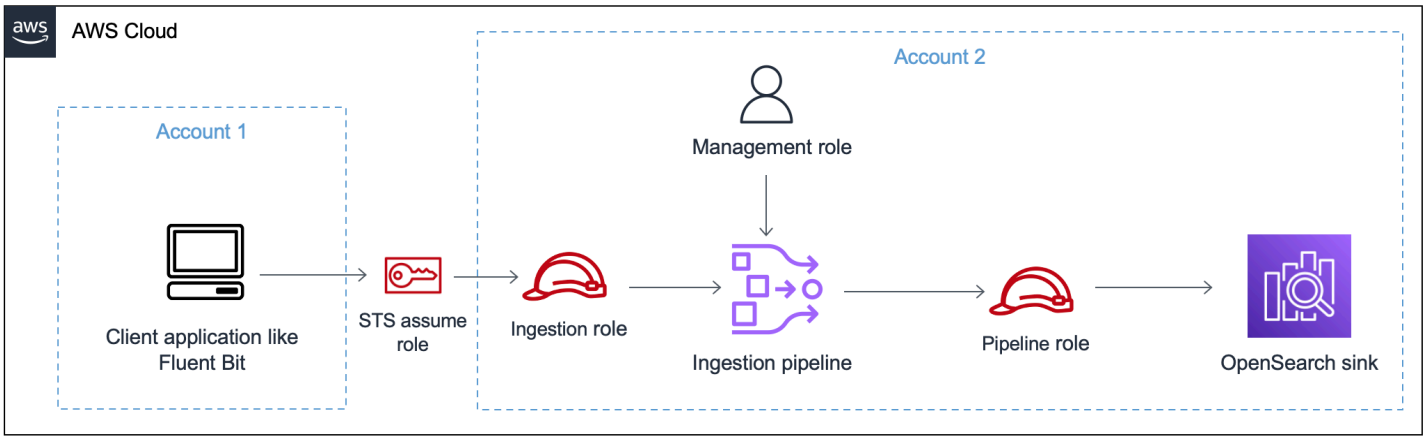
Configurazione di ruoli e utenti in Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion utilizza una varietà di modelli di autorizzazioni e ruoli IAM per consentire alle applicazioni di origine di scrivere nelle pipeline e consentire alle pipeline di scrivere nei sink. Prima di iniziare a importare dati, devi creare uno o più ruoli IAM con autorizzazioni specifiche in base al tuo caso d'uso.

Per configurare una pipeline di successo sono necessari almeno i seguenti ruoli.

Nome	Description
Ruolo della pipeline	Il ruolo pipeline fornisce le autorizzazioni necessarie affinché una pipeline possa leggere dall'origine e scrivere nel dominio o nel sink di raccolta. È possibile creare manualmente il ruolo pipeline oppure fare in modo che OpenSearch Ingestion lo crei automaticamente.
Ruolo di ingestione	Il ruolo di importazione contiene l'`osis:Ingest` autorizzazione per la risorsa della pipeline. Questa autorizzazione consente alle fonti basate su push di importare dati in una pipeline.

L'immagine seguente mostra una configurazione tipica di una pipeline, in cui un'origine dati come Amazon S3 o Fluent Bit sta scrivendo su una pipeline in un account diverso. In questo caso, il client deve assumere il ruolo di ingestione per accedere alla pipeline. Per ulteriori informazioni, consulta the section called “Inserimento tra più account”.



Per una semplice guida alla configurazione, consulta. the section called “Tutorial: inserisci dati in un dominio”

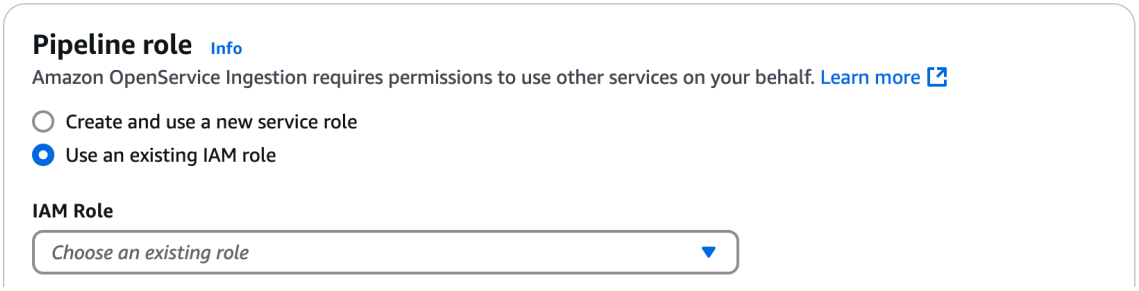
Argomenti

- the section called “Ruolo della pipeline”
- the section called “Ruolo di ingestione”
- the section called “Inserimento tra più account”

Ruolo della pipeline

Una pipeline necessita di determinate autorizzazioni per leggere dal codice sorgente e scrivere nel relativo sink. Queste autorizzazioni dipendono dall'applicazione client o da Servizio AWS quella che sta scrivendo nella pipeline e dal fatto che il sink sia un dominio di OpenSearch servizio, una raccolta OpenSearch Serverless o Amazon S3. Inoltre, una pipeline potrebbe aver bisogno delle autorizzazioni per estrarre fisicamente i dati dall'applicazione di origine (se l'origine è un plug-in basato su pull) e delle autorizzazioni per scrivere su una coda di lettere morte di S3, se abilitate.

Quando crei una pipeline, hai la possibilità di specificare un ruolo IAM esistente creato manualmente o fare in modo che OpenSearch Ingestion crei automaticamente il ruolo della pipeline in base all'origine e al sink selezionati. L'immagine seguente mostra come specificare il ruolo della pipeline in. Console di gestione AWS



Argomenti

- Automatizzazione della creazione dei ruoli della pipeline
- Creazione manuale del ruolo della pipeline

Automatizzazione della creazione dei ruoli della pipeline

Puoi scegliere di fare in modo che OpenSearch Ingestion crei il ruolo della pipeline per te. Identifica automaticamente le autorizzazioni richieste dal ruolo in base all'origine e ai sink configurati. Crea un ruolo IAM con il prefisso `OpenSearchIngestion-` e il suffisso che inserisci. Ad esempio,

se si immette `PipelineRole` come suffisso, OpenSearch Ingestion crea un ruolo denominato. `OpenSearchIngestion-PipelineRole`

La creazione automatica del ruolo pipeline semplifica il processo di configurazione e riduce la probabilità di errori di configurazione. Automatizzando la creazione dei ruoli, è possibile evitare l'assegnazione manuale delle autorizzazioni, garantendo l'applicazione delle politiche corrette senza rischiare configurazioni errate della sicurezza. Ciò consente inoltre di risparmiare tempo e migliora la conformità alla sicurezza applicando le migliori pratiche, garantendo al contempo la coerenza tra più implementazioni della pipeline.

È possibile fare in modo che OpenSearch Ingestion crei automaticamente il ruolo della pipeline solo in. Console di gestione AWS Se utilizzi l' AWS CLI API OpenSearch Ingestion o uno di questi SDKs, devi specificare un ruolo pipeline creato manualmente.

Per fare in modo che OpenSearch Ingestion crei il ruolo per te, seleziona Crea e usa un nuovo ruolo di servizio.

> **Important**
>
> È comunque necessario modificare manualmente la politica di accesso al dominio o alla raccolta per concedere l'accesso al ruolo della pipeline. Per i domini che utilizzano un controllo granulare degli accessi, è inoltre necessario mappare il ruolo della pipeline a un ruolo di backend. È possibile eseguire questi passaggi prima o dopo aver creato la pipeline. Per istruzioni, consultate i seguenti argomenti:
>
> - Configurare l'accesso ai dati per il dominio
> - Configura l'accesso ai dati e alla rete per la raccolta

Creazione manuale del ruolo della pipeline

Potresti preferire creare manualmente il ruolo pipeline se hai bisogno di un maggiore controllo sulle autorizzazioni per soddisfare specifici requisiti di sicurezza o conformità. La creazione manuale consente di personalizzare i ruoli per adattarli all'infrastruttura esistente o alle strategie di gestione degli accessi. Puoi anche scegliere la configurazione manuale per integrare il ruolo con altri Servizi AWS o assicurarti che sia in linea con le tue esigenze operative specifiche.

Per scegliere un ruolo pipeline creato manualmente, seleziona Usa un ruolo IAM esistente e scegli un ruolo esistente. Il ruolo deve disporre di tutte le autorizzazioni necessarie per ricevere dati dalla

fonte selezionata e scrivere nel sink selezionato. Nelle sezioni seguenti viene descritto come creare manualmente un ruolo di pipeline.

Argomenti

- Autorizzazioni per leggere da una fonte
- Autorizzazioni per scrivere su un sink di dominio
- Autorizzazioni per la scrittura in un sink di raccolta
- Autorizzazioni per scrivere su Amazon S3 o su una coda di lettere non scritte

Autorizzazioni per leggere da una fonte

Una pipeline di OpenSearch ingestione necessita dell'autorizzazione per leggere e ricevere dati dalla fonte specificata. Ad esempio, per un'origine Amazon DynamoDB, sono necessarie autorizzazioni come e. `dynamodb:DescribeTable` `dynamodb:DescribeStream` Per esempi di politiche di accesso ai ruoli della pipeline per fonti comuni, come Amazon S3, Fluent Bit e Collector, OpenTelemetry consulta. the section called “Integrazione delle pipeline”

Autorizzazioni per scrivere su un sink di dominio

Una pipeline di OpenSearch Ingestion necessita dell'autorizzazione per scrivere su un dominio OpenSearch di servizio configurato come sink. Queste autorizzazioni includono la possibilità di descrivere il dominio e inviargli richieste HTTP. Queste autorizzazioni sono le stesse per i domini pubblici e VPC. Per istruzioni su come creare un ruolo di pipeline e specificarlo nella politica di accesso al dominio, consulta Consentire alle pipeline di accedere ai domini.

Autorizzazioni per la scrittura in un sink di raccolta

Una pipeline di OpenSearch Ingestion necessita dell'autorizzazione per scrivere su una raccolta OpenSearch Serverless configurata come sink. Queste autorizzazioni includono la possibilità di descrivere la raccolta e inviarle richieste HTTP.

Innanzitutto, assicurati che la politica di accesso al ruolo della pipeline conceda le autorizzazioni richieste. Quindi, includi questo ruolo in una politica di accesso ai dati e concedigli le autorizzazioni per creare indici, aggiornare indici, descrivere gli indici e scrivere documenti all'interno della raccolta. Per istruzioni su come completare ciascuno di questi passaggi, consulta Consentire alle pipeline di accedere alle raccolte.

Autorizzazioni per scrivere su Amazon S3 o su una coda di lettere non scritte

Se specifichi Amazon S3 come destinazione sink per la tua pipeline o se abiliti una dead-letter queue (DLQ), il ruolo pipeline deve consentirle di accedere al bucket S3 che hai specificato come destinazione.

Associa una politica di autorizzazioni separata al ruolo della pipeline che fornisce l'accesso DLQ. Al ruolo deve essere concessa almeno l'`S3:PutObject`azione sulla risorsa bucket:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Sid": "WriteToS3DLQ",
      "Effect": "Allow",
      "Action": "s3:PutObject",
      "Resource": "arn:aws:s3:::my-dlq-bucket/*"
    }
  ]
}
```

Ruolo di ingestione

Il ruolo di ingestione è un ruolo IAM che consente ai servizi esterni di interagire e inviare dati in modo sicuro a una pipeline di Ingestion. OpenSearch Per le fonti basate su push, come Amazon Security Lake, questo ruolo deve concedere le autorizzazioni per inserire i dati nella pipeline, tra cui. `osis:Ingest` Per le fonti basate su pull, come Amazon S3, il ruolo deve OpenSearch consentire a Ingestion di assumerlo e accedere ai dati con le autorizzazioni necessarie.

Argomenti

- Ruolo di inserimento per sorgenti basate su push
- Ruolo di inserimento per sorgenti basate su pull
- Inserimento tra più account

Ruolo di inserimento per sorgenti basate su push

Per le fonti basate su push, i dati vengono inviati o inseriti nella pipeline di ingestione da un altro servizio, come Amazon Security Lake o Amazon DynamoDB. In questo scenario, il ruolo di importazione richiede almeno l'autorizzazione a interagire con la pipeline. `osis:Ingest`

La seguente policy di accesso IAM dimostra come concedere questa autorizzazione al ruolo di ingestione:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "osis:Ingest"
      ],
      "Resource": "arn:aws:osis:us-east-1:111122223333:pipeline/pipeline-name/*"
    }
  ]
}
```

Ruolo di inserimento per sorgenti basate su pull

Per le fonti basate su pull, la pipeline OpenSearch Ingestion estrae o recupera attivamente i dati da una fonte esterna, come Amazon S3. In questo caso, la pipeline deve assumere un ruolo di pipeline IAM che conceda le autorizzazioni necessarie per accedere all'origine dati. In questi scenari, il ruolo di ingestione è sinonimo di ruolo di pipeline.

Il ruolo deve includere una relazione di fiducia che consenta a OpenSearch Ingestion di assumerlo e autorizzazioni specifiche per l'origine dati. Per ulteriori informazioni, consulta the section called “Autorizzazioni per leggere da una fonte”.

Inserimento tra più account

Potrebbe essere necessario importare dati in una pipeline da un altro Account AWS, ad esempio un account di applicazione. Per configurare l'ingestione tra account, definite un ruolo di importazione

all'interno dello stesso account della pipeline e stabilite una relazione di fiducia tra il ruolo di inserimento e l'account dell'applicazione:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
     "Effect": "Allow",
     "Principal": {
       "AWS": "arn:aws:iam::444455556666:root"
      },
     "Action": "sts:AssumeRole"
  }]
}
```

Quindi, configura l'applicazione in modo che assuma il ruolo di ingestione. L'account dell'applicazione deve concedere al ruolo dell'applicazione AssumeRolele autorizzazioni per il ruolo di ingestione nell'account pipeline.

Per i passaggi dettagliati e gli esempi di policy IAM, consulta. the section called “Fornire l'accesso all'importazione su più account”

Concedere alle pipeline OpenSearch di Amazon Ingestion l'accesso ai domini

Una pipeline Amazon OpenSearch Ingestion necessita dell'autorizzazione per scrivere nel dominio OpenSearch di servizio configurato come sink. Per fornire l'accesso, configuri un ruolo AWS Identity and Access Management (IAM) con una politica di autorizzazioni restrittiva che limita l'accesso al dominio a cui una pipeline invia i dati. Ad esempio, potresti voler limitare una pipeline di ingestione solo al dominio e agli indici necessari per supportarne il caso d'uso.

> **Important**
>
> È possibile scegliere di creare manualmente il ruolo della pipeline oppure fare in modo che OpenSearch Ingestion lo crei automaticamente durante la creazione della pipeline. Se scegliete la creazione automatica dei ruoli, OpenSearch Ingestion aggiunge tutte le autorizzazioni necessarie alla politica di accesso al ruolo della pipeline in base all'origine e

al sink scelti. Crea un ruolo di pipeline in IAM con il prefisso `OpenSearchIngestion-` e il suffisso che inserisci. Per ulteriori informazioni, consulta the section called “Ruolo della pipeline”.
Se avete scelto OpenSearch Ingestion di creare automaticamente il ruolo della pipeline, dovete comunque includere il ruolo nella politica di accesso al dominio e mapparlo a un ruolo di backend (se il dominio utilizza il controllo di accesso granulare), prima o dopo la creazione della pipeline. Vedi il passaggio 2 per le istruzioni.

Argomenti

- Fase 1: Creare il ruolo della pipeline
- Fase 2: Configurare l'accesso ai dati per il dominio

Fase 1: Creare il ruolo della pipeline

Al ruolo pipeline deve essere associata una politica di autorizzazioni che gli consenta di inviare dati al domain sink. Deve inoltre avere una relazione di fiducia che consenta a OpenSearch Ingestion di assumere il ruolo. Per istruzioni su come allegare una policy a un ruolo, consulta Aggiungere i permessi di identità IAM nella Guida per l'utente IAM.

La seguente policy di esempio dimostra il privilegio minimo che puoi fornire in un ruolo di pipeline per la scrittura su un singolo dominio:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "es:DescribeDomain",
            "Resource": "arn:aws:es:*:111122223333:domain/*"
        },
        {
            "Effect": "Allow",
            "Action": "es:ESHttp*",
            "Resource": "arn:aws:es:*:111122223333:domain/domain-name/*"
        }
```

```
    ]
}
```

Se prevedi di riutilizzare il ruolo per scrivere su più domini, puoi rendere la politica più ampia sostituendo il nome di dominio con un carattere jolly (). *

Il ruolo deve avere la seguente relazione di fiducia, che consente a OpenSearch Ingestion di assumere il ruolo di pipeline:

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"osis-pipelines.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Fase 2: Configurare l'accesso ai dati per il dominio

Affinché una pipeline possa scrivere dati su un dominio, il dominio deve disporre di una politica di accesso a livello di dominio che consenta al ruolo della pipeline di accedervi.

Il seguente esempio di policy di accesso al dominio consente al ruolo pipeline denominato di `pipeline-role` scrivere dati nel dominio denominato: `ingestion-domain`

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
```

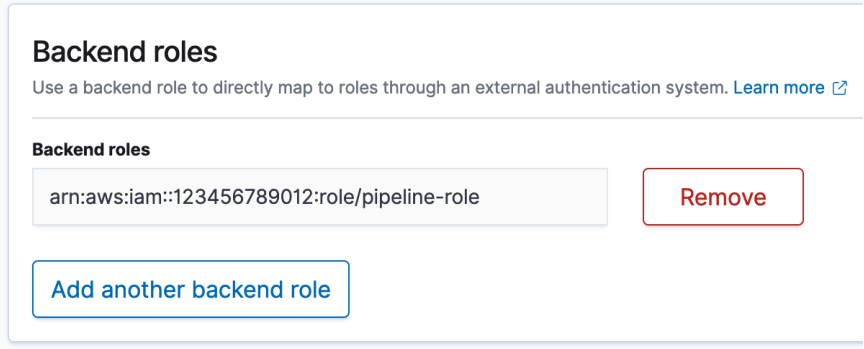
```
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::111122223333:role/pipeline-role"
            },
            "Action": [
                "es:DescribeDomain",
                "es:ESHttp*"
            ],
            "Resource": "arn:aws:es:us-east-1:111122223333:domain/domain-name/*"
        }
    ]
}
```

Mappa il ruolo della pipeline (solo per i domini che utilizzano un controllo di accesso granulare)

Se il tuo dominio utilizza un controllo granulare degli accessi per l'autenticazione, è necessario eseguire ulteriori passaggi per fornire alla pipeline l'accesso a un dominio. I passaggi variano a seconda della configurazione del dominio:

- Scenario 1: ruolo principale e ruolo della pipeline diversi: se utilizzi un IAM Amazon Resource Name (ARN) come utente principale e questo è diverso dal ruolo pipeline, devi mappare il ruolo della pipeline al ruolo di backend. OpenSearch `all_access` Questo aggiunge il ruolo della pipeline come utente principale aggiuntivo. Per ulteriori informazioni, consulta Utenti master aggiuntivi.
- Scenario 2: Utente principale nel database utenti interno: se il dominio utilizza un utente principale nel database utenti interno e l'autenticazione di base HTTP per le OpenSearch dashboard, non è possibile passare il nome utente e la password principali direttamente nella configurazione della pipeline. Invece, associa il ruolo della pipeline al ruolo di backend. OpenSearch `all_access` Questo aggiunge il ruolo della pipeline come utente principale aggiuntivo. Per ulteriori informazioni, consulta Utenti master aggiuntivi.
- Scenario 3: stesso ruolo principale e ruolo della pipeline (non comune): se utilizzi un ARN IAM come utente principale ed è lo stesso ARN che utilizzi come ruolo pipeline, non devi intraprendere ulteriori azioni. La pipeline dispone delle autorizzazioni necessarie per scrivere sul dominio. Questo scenario non è comune perché la maggior parte degli ambienti utilizza un ruolo di amministratore o un altro ruolo come ruolo principale.

L'immagine seguente mostra come mappare il ruolo della pipeline a un ruolo di backend:



Concedere alle pipeline OpenSearch di Amazon Ingestion l'accesso alle raccolte

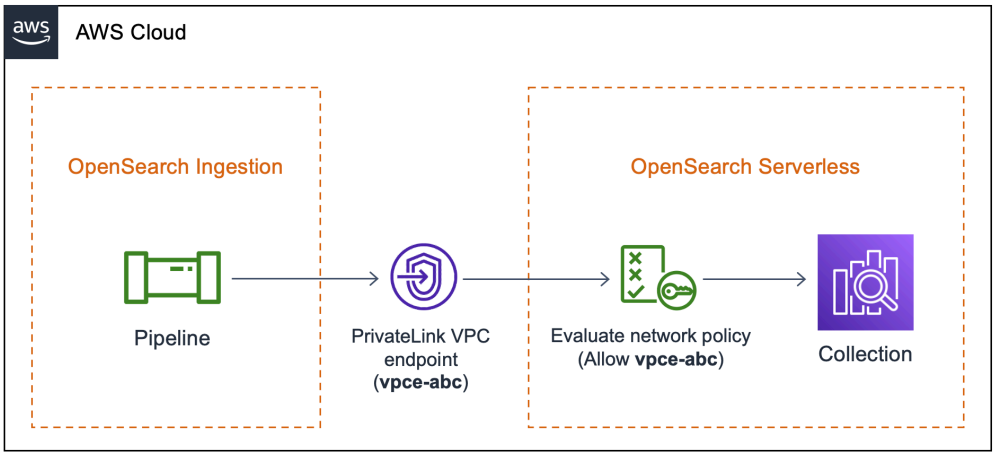
Una pipeline Amazon OpenSearch Ingestion può scrivere su una raccolta pubblica OpenSearch Serverless o su una raccolta VPC. Per fornire l'accesso alla raccolta, configuri un ruolo di pipeline AWS Identity and Access Management (IAM) con una politica di autorizzazioni che garantisca l'accesso alla raccolta. La pipeline assume questo ruolo per firmare le richieste al sink di raccolta Serverless. OpenSearch

Important

È possibile scegliere di creare manualmente il ruolo della pipeline oppure fare in modo che OpenSearch Ingestion lo crei automaticamente durante la creazione della pipeline. Se scegliete la creazione automatica dei ruoli, OpenSearch Ingestion aggiunge tutte le autorizzazioni necessarie alla politica di accesso al ruolo della pipeline in base all'origine e al sink scelti. Crea un ruolo di pipeline in IAM con il prefisso `OpenSearchIngestion-` e il suffisso che inserisci. Per ulteriori informazioni, consulta the section called “Ruolo della pipeline”.

Se avete scelto OpenSearch Ingestion di creare il ruolo pipeline per voi, dovete comunque includere il ruolo nella politica di accesso ai dati della raccolta, prima o dopo la creazione della pipeline. Vedi il passaggio 2 per le istruzioni.

Durante la creazione della pipeline, OpenSearch Ingestion crea una AWS PrivateLink connessione tra la pipeline e la raccolta Serverless. OpenSearch Tutto il traffico proveniente dalla pipeline passa attraverso questo endpoint VPC e viene indirizzato alla raccolta. Per raggiungere la raccolta, all'endpoint deve essere concesso l'accesso alla raccolta tramite una politica di accesso alla rete.



Argomenti

- Fase 1: Creare il ruolo della pipeline
- Fase 2: Configurazione dei dati e dell'accesso alla rete per la raccolta

Fase 1: Creare il ruolo della pipeline

Al ruolo pipeline deve essere associata una politica di autorizzazioni che gli consenta di inviare dati al sink di raccolta. Deve inoltre avere un rapporto di fiducia che consenta a OpenSearch Ingestion di assumere il ruolo. Per istruzioni su come allegare una policy a un ruolo, consulta Aggiungere i permessi di identità IAM nella Guida per l'utente IAM.

La seguente policy di esempio mostra il privilegio minimo che puoi fornire in una policy di accesso ai ruoli della pipeline per consentirle di scrivere nelle raccolte:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "Statement1",
            "Effect": "Allow",
            "Action": [
                "aoss:APIAccessAll",
                "aoss:BatchGetCollection",
                "aoss:CreateSecurityPolicy",
                "aoss:GetSecurityPolicy",
```

```
                "aoss:UpdateSecurityPolicy"
            ],
            "Resource": "*"
        }
    ]
}
```

Il ruolo deve avere la seguente relazione di fiducia, che consente a OpenSearch Ingestion di assumerlo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "osis-pipelines.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

Fase 2: Configurazione dei dati e dell'accesso alla rete per la raccolta

Crea una raccolta OpenSearch Serverless con le seguenti impostazioni. Per istruzioni su come creare una raccolta, consultathe section called “Creazione di raccolte”.

Politica di accesso ai dati

Crea una politica di accesso ai dati per la raccolta che conceda le autorizzazioni richieste al ruolo della pipeline. Esempio:

```
[
  {
    "Rules": [
      {
```

```
        "Resource": [
          "index/collection-name/*"
        ],
        "Permission": [
          "aoss:CreateIndex",
          "aoss:UpdateIndex",
          "aoss:DescribeIndex",
          "aoss:WriteDocument"
        ],
        "ResourceType": "index"
      }
    ],
    "Principal": [
      "arn:aws:iam::account-id:role/pipeline-role"
    ],
    "Description": "Pipeline role access"
  }
]
```

> **Note**
>
> Nell'`Principal`elemento, specifica l'Amazon Resource Name (ARN) del ruolo della pipeline.

Politica di accesso alla rete

A ogni raccolta creata in OpenSearch Serverless è associata almeno una politica di accesso alla rete. Le politiche di accesso alla rete determinano se la raccolta è accessibile su Internet da reti pubbliche o se è necessario accedervi privatamente. Per ulteriori informazioni sulle politiche di rete, vederethe section called “Accesso alla rete”.

All'interno di una politica di accesso alla rete, puoi specificare solo endpoint OpenSearch VPC gestiti senza server. Per ulteriori informazioni, consulta the section called “Endpoint VPC del piano dati”. Tuttavia, affinché la pipeline possa scrivere nella raccolta, la policy deve anche concedere l'accesso all'endpoint VPC OpenSearch che Ingestion crea automaticamente tra la pipeline e la raccolta. Pertanto, se si sceglie una raccolta OpenSearch Serverless come sink di destinazione per una pipeline, è necessario inserire il nome della politica di rete associata nel campo Nome della politica di rete.

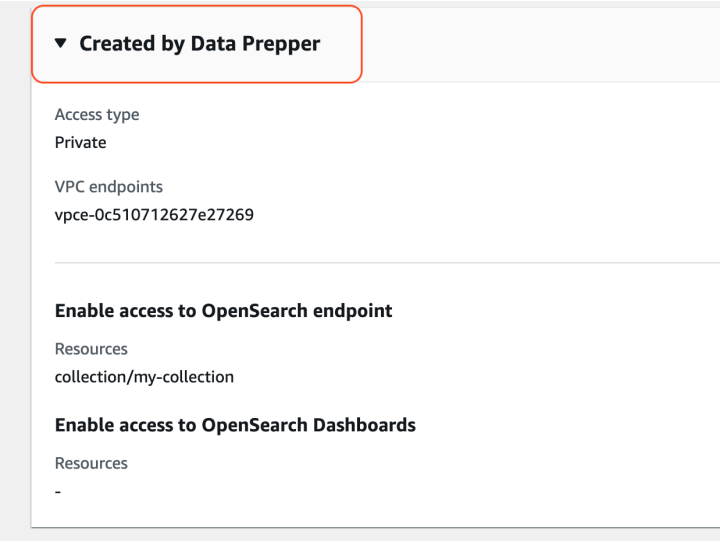
Durante la creazione della pipeline, OpenSearch Ingestion verifica l'esistenza della politica di rete specificata. Se non esiste, OpenSearch Ingestion la crea. Se esiste, OpenSearch Ingestion lo

aggiorna aggiungendovi una nuova regola. La regola concede l'accesso all'endpoint VPC che collega la pipeline e la raccolta.

Esempio:

```
{
   "Rules":[
      {
         "Resource":[
            "collection/my-collection"
         ],
         "ResourceType":"collection"
      }
   ],
   "SourceVPCEs":[
      "vpce-0c510712627e27269" # The ID of the VPC endpoint that OpenSearch Ingestion
 creates between the pipeline and collection
   ],
   "Description":"Created by Data Prepper"
}
```

Nella console, tutte le regole che OpenSearch Ingestion aggiunge alle policy di rete sono denominate Created by Data Prepper:



> **Note**
>
> In generale, una regola che specifica l'accesso pubblico per una raccolta ha la precedenza su una regola che specifica l'accesso privato. Pertanto, se la policy aveva già configurato

l'accesso pubblico, questa nuova regola aggiunta da OpenSearch Ingestion non modifica effettivamente il comportamento della politica. Per ulteriori informazioni, consulta the section called “Priorità delle policy”.

Se si arresta o si elimina la pipeline, OpenSearch Ingestion elimina l'endpoint VPC tra la pipeline e la raccolta. Modifica inoltre la politica di rete per rimuovere l'endpoint VPC dall'elenco degli endpoint consentiti. Se riavvii la pipeline, ricrea l'endpoint VPC e aggiorna nuovamente la politica di rete con l'ID dell'endpoint.

Guida introduttiva ad Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion supporta l'importazione di dati in domini di OpenSearch servizio gestiti e raccolte Serverless. OpenSearch I seguenti tutorial illustrano i passaggi di base per rendere operativa una pipeline.

Il primo tutorial mostra come usare Amazon OpenSearch Ingestion per configurare una pipeline semplice e inserire dati in un dominio Amazon Service. OpenSearch

Il secondo tutorial mostra come usare Amazon OpenSearch Ingestion per configurare una pipeline semplice e inserire dati in una raccolta Amazon Serverless. OpenSearch

Note

La creazione della pipeline avrà esito negativo se non configuri le autorizzazioni corrette. Vedi the section called “Configurazione di ruoli e utenti” per una migliore comprensione dei ruoli richiesti prima di creare una pipeline.

Argomenti

- Tutorial: importazione di dati in un dominio utilizzando Amazon Ingestion OpenSearch
- Tutorial: Inserimento di dati in una raccolta con Amazon Ingestion OpenSearch

Tutorial: importazione di dati in un dominio utilizzando Amazon Ingestion OpenSearch

Questo tutorial mostra come usare Amazon OpenSearch Ingestion per configurare una pipeline semplice e inserire dati in un dominio Amazon Service. OpenSearch Una pipeline è una risorsa che OpenSearch Ingestion fornisce e gestisce. È possibile utilizzare una pipeline per filtrare, arricchire, trasformare, normalizzare e aggregare i dati per l'analisi e la visualizzazione a valle in Service. OpenSearch

Questo tutorial illustra i passaggi di base per avviare rapidamente una pipeline. Per istruzioni più complete, consultathe section called “Creazione di pipeline”.

In questo tutorial completerai le seguenti fasi:

1. Crea un dominio.
2. Crea una pipeline.
3. Inserisci alcuni dati di esempio.

All'interno del tutorial, creerai le seguenti risorse:

- Un dominio denominato su `ingestion-domain` cui la pipeline scrive
- Una pipeline denominata `ingestion-pipeline`

Autorizzazioni richieste

Per completare questo tutorial, l'utente o il ruolo deve avere una politica basata sull'identità allegata con le seguenti autorizzazioni minime. Queste autorizzazioni consentono di creare un ruolo di pipeline e allegare una policy (`iam:Create*eiam:Attach*`), creare o modificare un dominio () e lavorare con `es:*` pipelines (). `osis:*`

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
```

```
         "Resource":"*",
         "Action":[
            "osis:*",
            "iam:Create*",
            "iam:Attach*",
            "es:*"
         ]
      },
      {
         "Resource":[
            "arn:aws:iam::111122223333:role/OpenSearchIngestion-PipelineRole"
         ],
         "Effect":"Allow",
         "Action":[
            "iam:CreateRole",
            "iam:AttachRolePolicy",
            "iam:PassRole"
         ]
      }
   ]
}
```

Fase 1: Creare il ruolo della pipeline

Innanzitutto, crea un ruolo che la pipeline assumerà per accedere al OpenSearch service domain sink. Includerai questo ruolo nella configurazione della pipeline più avanti in questo tutorial.

Per creare il ruolo della pipeline

1. Apri la AWS Identity and Access Management console all'indirizzo https://console.aws.amazon.com/iamv2/.
2. Scegli Politiche, quindi scegli Crea politica.
3. In questo tutorial, inserirai i dati in un dominio chiamato`ingestion-domain`, che creerai nel passaggio successivo. Seleziona JSON e incolla la seguente politica nell'editor. Sostituiscilo *`your-account-id`* con l'ID del tuo account e, se necessario, modifica la regione.

   JSON

```
{
    "Version":"2012-10-17",
```

```
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "es:DescribeDomain",
            "Resource": "arn:aws:es:us-east-1:111122223333:domain/ingestion-
domain"
        },
        {
            "Effect": "Allow",
            "Action": "es:ESHttp*",
            "Resource": "arn:aws:es:us-east-1:111122223333:domain/ingestion-
domain/*"
        }
    ]
}
```

Se desideri scrivere dati su un dominio esistente, sostituiscili `ingestion-domain` con il nome del tuo dominio.

> **Note**
>
> Per semplicità in questo tutorial, utilizziamo un'ampia politica di accesso. Negli ambienti di produzione, tuttavia, ti consigliamo di applicare una politica di accesso più restrittiva al tuo ruolo di pipeline. Per un esempio di policy che fornisce le autorizzazioni minime richieste, vedi. the section called “Concedere alle pipeline l'accesso ai domini”

4. Scegliete Avanti, scegliete Avanti e assegnate un nome alla vostra policy pipeline-policy.
5. Scegli Crea policy.
6. Quindi, crea un ruolo e allega la policy ad esso. Selezionare Roles (Ruoli), quindi selezionare Create role (Crea ruolo).
7. Scegli una politica di fiducia personalizzata e incolla la seguente politica nell'editor:

   JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
```

```
            "Principal":{
               "Service":"osis-pipelines.amazonaws.com"
            },
            "Action":"sts:AssumeRole"
         }
      ]
}
```

8. Scegli Next (Successivo). Quindi cerca e seleziona pipeline-policy (che hai appena creato).
9. Scegli Avanti e assegna un nome al ruolo. PipelineRole
10. Scegli Crea ruolo.

Ricorda l'Amazon Resource Name (ARN) del ruolo (ad esempio,`arn:aws:iam::`*`your-account-id`*`:role/PipelineRole`). Ne avrai bisogno quando creerai la tua pipeline.

Passaggio 2: crea un dominio

Innanzitutto, crea un dominio denominato in `ingestion-domain` cui inserire i dati.

Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa e crea un dominio che soddisfi i seguenti requisiti:

- È in esecuzione OpenSearch 1.0 o versione successiva oppure Elasticsearch 7.4 o versione successiva
- Utilizza l'accesso pubblico
- Non utilizza un controllo degli accessi a grana fine

> **Note**
>
> Questi requisiti hanno lo scopo di garantire la semplicità di questo tutorial. Negli ambienti di produzione, puoi configurare un dominio con accesso VPC utilizzando un controllo degli and/or accessi granulare. Per utilizzare un controllo granulare degli accessi, consulta Mappare il ruolo della pipeline.

Il dominio deve avere una politica di accesso che conceda l'autorizzazione al ruolo `OpenSearchIngestion-PipelineRole` IAM, che OpenSearch Service creerà per te nella fase successiva. La pipeline assumerà questo ruolo per inviare i dati al domain sink.

Assicurati che il dominio abbia la seguente politica di accesso a livello di dominio, che concede al ruolo della pipeline l'accesso al dominio. Sostituisci la regione e l'ID dell'account con i tuoi:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/OpenSearchIngestion-PipelineRole"
      },
      "Action": "es:*",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/ingestion-domain/*"
    }
  ]
}
```

Per ulteriori informazioni sulla creazione di politiche di accesso a livello di dominio, consulta. the section called “Policy basate sulle risorse”

Se hai già creato un dominio, modifica la politica di accesso esistente per fornire le autorizzazioni di cui sopra. `OpenSearchIngestion-PipelineRole`

Fase 3: Creare una pipeline

Ora che hai un dominio, puoi creare una pipeline.

Come creare una pipeline

1. Nella console di Amazon OpenSearch Service, scegli Pipelines dal riquadro di navigazione a sinistra.
2. Scegliere Create pipeline (Crea pipeline).
3. Seleziona la pipeline vuota, quindi scegli Seleziona blueprint.
4. In questo tutorial, creeremo una semplice pipeline che utilizza il plug-in di origine HTTP. Il plugin accetta i dati di registro in un formato di matrice JSON. Specificheremo un singolo dominio di OpenSearch servizio come sink e inseriremo tutti i dati nell'`application_logs`indice.

Nel menu Sorgente, scegli HTTP. Per il percorso, inserisci /logs.

5. Per semplicità, in questo tutorial, configureremo l'accesso pubblico alla pipeline. Per le opzioni di rete di origine, scegli Accesso pubblico. Per informazioni sulla configurazione dell'accesso al VPC, vedere. the section called “Configurazione dell'accesso VPC per le pipeline”
6. Scegli Next (Successivo).
7. Per Processore, inserisci Data e scegli Aggiungi.
8. Abilita A partire dall'ora di ricezione. Lascia tutte le altre impostazioni come predefinite.
9. Scegli Next (Successivo).
10. Configura i dettagli del lavandino. Per il tipo di OpenSearch risorsa, scegli Cluster gestito. Quindi scegli il dominio di OpenSearch servizio che hai creato nella sezione precedente.

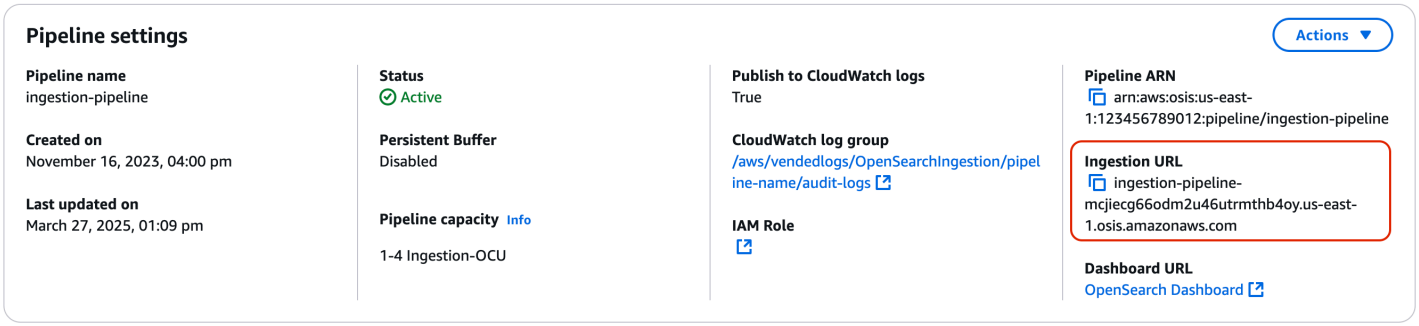
    Per il nome dell'indice, inserisci application_logs. OpenSearch Ingestion crea automaticamente questo indice nel dominio se non esiste già.
11. Scegli Next (Successivo).
12. Assegna un nome alla pipeline di ingestione-pipeline. Lasciate le impostazioni di capacità come predefinite.
13. Per il ruolo Pipeline, seleziona Crea e usa un nuovo ruolo di servizio. Il ruolo pipeline fornisce le autorizzazioni necessarie affinché una pipeline possa scrivere nel domain sink e leggere da fonti basate su pull. Selezionando questa opzione, OpenSearch consenti a Ingestion di creare il ruolo per te, anziché crearlo manualmente in IAM. Per ulteriori informazioni, consulta the section called “Configurazione di ruoli e utenti”.
14. Per il suffisso del nome del ruolo di servizio, immettere. PipelineRole In IAM, il ruolo avrà il seguente formato`arn:aws:iam::`*`your-account-id`*`:role/OpenSearchIngestion-`**`PipelineRole`**.
15. Scegli Next (Successivo). Controlla la configurazione della pipeline e scegli Crea pipeline. La pipeline impiega 5-10 minuti per diventare attiva.

Fase 4: Inserimento di alcuni dati di esempio

Quando lo stato della pipeline è impostato`Active`, puoi iniziare a importare i dati al suo interno. È necessario firmare tutte le richieste HTTP alla pipeline utilizzando Signature Version 4. Utilizza uno strumento HTTP come Postman o awscurl per inviare alcuni dati alla pipeline. Come per l'indicizzazione dei dati direttamente su un dominio, l'importazione dei dati in una pipeline richiede sempre un ruolo IAM o una chiave di accesso IAM e una chiave segreta.

> **Note**
>
> Il principale che firma la richiesta deve disporre dell'autorizzazione IAM. `osis:Ingest`

Innanzitutto, ottieni l'URL di importazione dalla pagina delle impostazioni di Pipeline:



Quindi, inserisci alcuni dati di esempio. La seguente richiesta utilizza awscurl per inviare un singolo file di registro alla pipeline:

```
awscurl --service osis --region us-east-1 \
    -X POST \
    -H "Content-Type: application/json" \
    -d
 '[{"time":"2014-08-11T11:40:13+00:00","remote_addr":"122.226.223.69","status":"404","request":
 http://www.k2proxy.com//hello.html HTTP/1.1","http_user_agent":"Mozilla/4.0
 (compatible; WOW64; SLCC2;)"}]' \
    https://pipeline-endpoint.us-east-1.osis.amazonaws.com/logs
```

Dovresti vedere una risposta. `200 OK` Se ricevi un errore di autenticazione, potrebbe essere dovuto al fatto che stai importando dati da un account diverso da quello in cui si trova la pipeline. Per informazioni, consulta the section called “Risoluzione dei problemi relativi alle autorizzazioni”.

Ora, interroga l'`application_logs`indice per assicurarti che la voce di registro sia stata inserita correttamente:

```
awscurl --service es --region us-east-1 \
     -X GET \
     https://search-ingestion-domain.us-east-1.es.amazonaws.com/application_logs/
_search | json_pp
```

Esempio di risposta:

```
{
   "took":984,
   "timed_out":false,
   "_shards":{
      "total":1,
      "successful":5,
      "skipped":0,
      "failed":0
   },
   "hits":{
      "total":{
         "value":1,
         "relation":"eq"
      },
      "max_score":1.0,
      "hits":[
         {
            "_index":"application_logs",
            "_type":"_doc",
            "_id":"z6VY_IMBRpceX-DU6V4O",
            "_score":1.0,
            "_source":{
               "time":"2014-08-11T11:40:13+00:00",
               "remote_addr":"122.226.223.69",
               "status":"404",
               "request":"GET http://www.k2proxy.com//hello.html HTTP/1.1",
               "http_user_agent":"Mozilla/4.0 (compatible; WOW64; SLCC2;)",
               "@timestamp":"2022-10-21T21:00:25.502Z"
            }
         }
      ]
   }
}
```


Risoluzione dei problemi relativi alle autorizzazioni

Se hai seguito i passaggi del tutorial e continui a riscontrare errori di autenticazione quando tenti di importare dati, è possibile che il ruolo della scrittura in una pipeline sia Account AWS diverso da quello della pipeline stessa. In questo caso, devi creare e assumere un ruolo che ti consenta specificamente di importare dati. Per istruzioni, consulta the section called “Fornire l'accesso all'importazione su più account”.

Risorse correlate

Questo tutorial ha presentato un semplice caso d'uso di ingestione di un singolo documento tramite HTTP. Negli scenari di produzione, configurerai le tue applicazioni client (come Fluent Bit, Kubernetes o OpenTelemetry Collector) per inviare dati a una o più pipeline. Le tue pipeline saranno probabilmente più complesse del semplice esempio di questo tutorial.

Per iniziare a configurare i client e ad acquisire dati, consulta le seguenti risorse:

- Creazione e gestione delle pipeline
- Configurazione dei client per l'invio di dati a Ingestion OpenSearch
- Documentazione Data Prepper

Tutorial: Inserimento di dati in una raccolta con Amazon Ingestion OpenSearch

Questo tutorial mostra come usare Amazon OpenSearch Ingestion per configurare una pipeline semplice e inserire dati in una raccolta Amazon Serverless. OpenSearch Una pipeline è una risorsa che Ingestion fornisce e gestisce. OpenSearch È possibile utilizzare una pipeline per filtrare, arricchire, trasformare, normalizzare e aggregare i dati per l'analisi e la visualizzazione a valle in Service. OpenSearch

Per un tutorial che dimostra come inserire dati in un dominio di servizio fornito, consulta. OpenSearch the section called “Tutorial: inserisci dati in un dominio”

Completerai i seguenti passaggi in questo tutorial:.

1. Creare una raccolta.
2. Crea una pipeline.
3. Inserisci alcuni dati di esempio.

All'interno del tutorial, creerai le seguenti risorse:

- Una raccolta denominata su `ingestion-collection` cui la pipeline scriverà
- Una pipeline denominata `ingestion-pipeline-serverless`

Autorizzazioni richieste

Per completare questo tutorial, l'utente o il ruolo deve avere una politica basata sull'identità allegata con le seguenti autorizzazioni minime. Queste autorizzazioni consentono di creare un ruolo di pipeline e allegare una policy (`iam:Create*eiam:Attach*`), creare o modificare una raccolta () e lavorare con `aoss:*` pipelines (). `osis:*`

Inoltre, sono necessarie diverse autorizzazioni IAM per creare automaticamente il ruolo della pipeline e passarlo a OpenSearch Ingestion in modo che possa scrivere dati nella raccolta.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Resource":"*",
         "Action":[
            "osis:*",
            "iam:Create*",
            "iam:Attach*",
            "aoss:*"
         ]
      },
      {
         "Resource":[
            "arn:aws:iam::111122223333:role/OpenSearchIngestion-PipelineRole"
         ],
         "Effect":"Allow",
         "Action":[
            "iam:CreateRole",
            "iam:AttachRolePolicy",
            "iam:PassRole"
         ]
      }
   ]
}
```

Fase 1: Creare una raccolta

Innanzitutto, crea una raccolta in cui inserire i dati. Daremo un nome alla raccolta`ingestion-collection`.

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Scegli Raccolte dalla barra di navigazione a sinistra e scegli Crea raccolta.
3. Assegna un nome alla raccolta ingestion-collection.
4. Per Sicurezza, scegliete Standard create.
5. In Impostazioni di accesso alla rete, modifica il tipo di accesso in Pubblico.
6. Mantenere tutte le altre impostazioni come valori predefiniti e scegliere Successivo.
7. Ora configura una politica di accesso ai dati per la raccolta. Deseleziona Abbina automaticamente le impostazioni dei criteri di accesso.
8. Per il metodo di definizione, scegli JSON e incolla la seguente politica nell'editor. Questa politica fa due cose:

   - Consente al ruolo della pipeline di scrivere nella raccolta.
   - Consente di leggere i contenuti della raccolta. Successivamente, dopo aver inserito alcuni dati di esempio nella pipeline, interrogherete la raccolta per assicurarvi che i dati siano stati inseriti e scritti correttamente nell'indice.

```
[
  {
    "Rules": [
      {
        "Resource": [
          "index/ingestion-collection/*"
        ],
        "Permission": [
          "aoss:CreateIndex",
          "aoss:UpdateIndex",
          "aoss:DescribeIndex",
          "aoss:ReadDocument",
          "aoss:WriteDocument"
        ],
        "ResourceType": "index"
      }
    ],
```

```
    "Principal": [
      "arn:aws:iam::your-account-id:role/OpenSearchIngestion-PipelineRole",
      "arn:aws:iam::your-account-id:role/Admin"
    ],
    "Description": "Rule 1"
  }
]
```

9. Modifica `Principal` gli elementi per includere il tuo ID. Account AWS Per il secondo principale, specificate un utente o un ruolo che potete utilizzare per interrogare la raccolta in un secondo momento.
10. Scegli Next (Successivo). Assegna un nome alla politica di accesso pipeline-collection-accesse scegli nuovamente Avanti.
11. Rivedi la configurazione della raccolta e scegli Submit (Invia).

Fase 2: Creare una pipeline

Ora che hai una collezione, puoi creare una pipeline.

Come creare una pipeline

1. Nella console di Amazon OpenSearch Service, scegli Pipelines dal riquadro di navigazione a sinistra.
2. Scegliere Create pipeline (Crea pipeline).
3. Seleziona la pipeline vuota, quindi scegli Seleziona blueprint.
4. In questo tutorial, creeremo una semplice pipeline che utilizza il plug-in di origine HTTP. Il plugin accetta i dati di registro in un formato di matrice JSON. Specificheremo una singola raccolta OpenSearch Serverless come sink e inseriremo tutti i dati nell'indice. `my_logs`

   Nel menu Sorgente, scegli HTTP. Per il percorso, inserisci /logs.
5. Per semplicità, in questo tutorial, configureremo l'accesso pubblico alla pipeline. Per le opzioni di rete di origine, scegli Accesso pubblico. Per informazioni sulla configurazione dell'accesso al VPC, vedere. the section called “Configurazione dell'accesso VPC per le pipeline”
6. Scegli Next (Successivo).
7. Per Processore, inserisci Data e scegli Aggiungi.
8. Abilita A partire dall'ora di ricezione. Lascia tutte le altre impostazioni come predefinite.
9. Scegli Next (Successivo).

10. Configura i dettagli del lavandino. Per il tipo di OpenSearch risorsa, scegli Collection (Serverless). Quindi scegli la raccolta OpenSearch di servizi che hai creato nella sezione precedente.

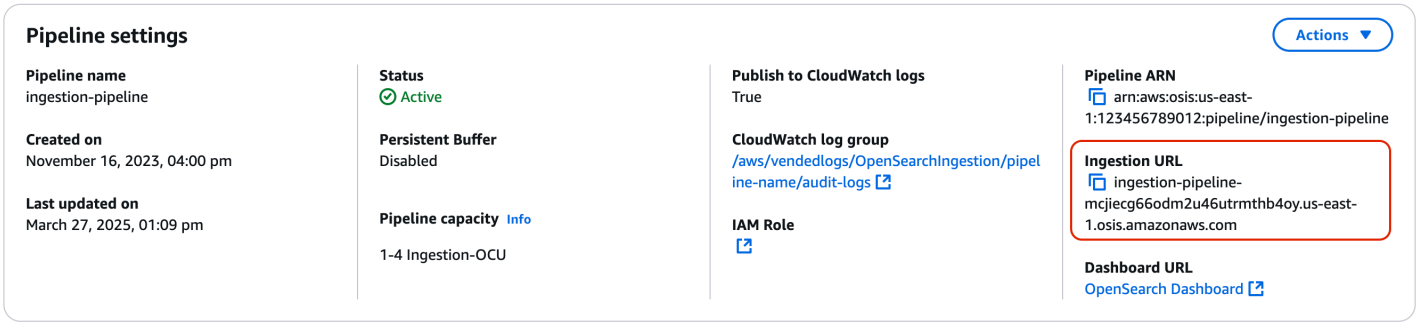
    Lasciate il nome della politica di rete come predefinito. Per il nome dell'indice, inserisci my_logs. OpenSearch Ingestion crea automaticamente questo indice nella raccolta se non esiste già.
11. Scegli Next (Successivo).
12. Assegna un nome alla pipeline. ingestion-pipeline-serverless Lasciate le impostazioni di capacità come predefinite.
13. Per il ruolo Pipeline, seleziona Crea e usa un nuovo ruolo di servizio. Il ruolo pipeline fornisce le autorizzazioni necessarie affinché una pipeline possa scrivere nel sink di raccolta e leggere da fonti basate su pull. Selezionando questa opzione, OpenSearch consenti a Ingestion di creare il ruolo per te, anziché crearlo manualmente in IAM. Per ulteriori informazioni, consulta the section called “Configurazione di ruoli e utenti”.
14. Per il suffisso del nome del ruolo di servizio, immettere. PipelineRole In IAM, il ruolo avrà il seguente formato`arn:aws:iam::`*`your-account-id`*`:role/OpenSearchIngestion-`**`PipelineRole`**.
15. Scegli Next (Successivo). Controlla la configurazione della pipeline e scegli Crea pipeline. La pipeline impiega 5-10 minuti per diventare attiva.

Fase 3: Inserimento di alcuni dati di esempio

Quando lo stato della pipeline è impostato`Active`, puoi iniziare a importare dati al suo interno. È necessario firmare tutte le richieste HTTP alla pipeline utilizzando Signature Version 4. Utilizza uno strumento HTTP come Postman o awscurl per inviare alcuni dati alla pipeline. Come per l'indicizzazione dei dati direttamente in una raccolta, l'importazione dei dati in una pipeline richiede sempre un ruolo IAM o una chiave di accesso IAM e una chiave segreta.

> **Note**
>
> Il principale che firma la richiesta deve disporre dell'autorizzazione IAM. `osis:Ingest`

Innanzitutto, ottieni l'URL di importazione dalla pagina delle impostazioni di Pipeline:



Quindi, invia alcuni dati di esempio al percorso di ingestione. La seguente richiesta di esempio utilizza awscurl per inviare un singolo file di registro alla pipeline:

```
awscurl --service osis --region us-east-1 \
    -X POST \
    -H "Content-Type: application/json" \
    -d
 '[{"time":"2014-08-11T11:40:13+00:00","remote_addr":"122.226.223.69","status":"404","request":
 http://www.k2proxy.com//hello.html HTTP/1.1","http_user_agent":"Mozilla/4.0
 (compatible; WOW64; SLCC2;)"}]' \
    https://pipeline-endpoint.us-east-1.osis.amazonaws.com/logs
```

Dovresti vedere una risposta. `200 OK`

Ora, interroga l'`my_logs`indice per assicurarti che la voce di registro sia stata inserita correttamente:

```
awscurl --service aoss --region us-east-1 \
     -X GET \
     https://collection-id.us-east-1.aoss.amazonaws.com/my_logs/_search | json_pp
```

Esempio di risposta:

```
{
   "took":348,
   "timed_out":false,
   "_shards":{
      "total":0,
      "successful":0,
      "skipped":0,
      "failed":0
   },
   "hits":{
      "total":{
```

```
         "value":1,
         "relation":"eq"
      },
      "max_score":1.0,
      "hits":[
         {
            "_index":"my_logs",
            "_id":"1%3A0%3ARJgDvIcBTy5m12xrKE-y",
            "_score":1.0,
            "_source":{
               "time":"2014-08-11T11:40:13+00:00",
               "remote_addr":"122.226.223.69",
               "status":"404",
               "request":"GET http://www.k2proxy.com//hello.html HTTP/1.1",
               "http_user_agent":"Mozilla/4.0 (compatible; WOW64; SLCC2;)",
               "@timestamp":"2023-04-26T05:22:16.204Z"
            }
         }
      ]
   }
}
```


Risorse correlate

Questo tutorial ha presentato un semplice caso d'uso di assimilazione di un singolo documento tramite HTTP. Negli scenari di produzione, configurerai le tue applicazioni client (come Fluent Bit, Kubernetes o OpenTelemetry Collector) per inviare dati a una o più pipeline. Le tue pipeline saranno probabilmente più complesse del semplice esempio di questo tutorial.

Per iniziare a configurare i client e ad acquisire dati, consulta le seguenti risorse:

- Creazione e gestione delle pipeline
- Configurazione dei client per l'invio di dati a Ingestion OpenSearch
- Documentazione Data Prepper

Panoramica delle funzionalità della pipeline in Amazon Ingestion OpenSearch

Amazon OpenSearch Ingestion fornisce pipeline costituite da una fonte, un buffer, zero o più processori e uno o più sink. Le pipeline di ingestione sono alimentate da Data Prepper come motore

di dati. Per una panoramica dei vari componenti di una pipeline, vedere. the section called “Concetti chiave”

Le seguenti sezioni forniscono una panoramica di alcune delle funzionalità più comunemente utilizzate in Amazon OpenSearch Ingestion.

> **Note**
>
> Questo non è un elenco esaustivo delle funzionalità disponibili per le pipeline. Per una documentazione completa di tutte le funzionalità della pipeline disponibili, consulta la documentazione di Data Prepper. Tieni presente che OpenSearch Ingestion impone alcuni vincoli ai plugin e alle opzioni che puoi utilizzare. Per ulteriori informazioni, consulta the section called “Plugin e opzioni supportati”.

Argomenti

- Buffering persistente
- Divisione
- Concatenamento
- Code di lettere morte
- Gestione degli indici
- End-to-end riconoscimento
- Contropressione alla fonte

Buffering persistente

Un buffer persistente archivia i dati in un buffer basato su disco su più zone di disponibilità per migliorare la durabilità dei dati. È possibile utilizzare il buffering persistente per importare dati da tutte le fonti basate su push supportate senza configurare un buffer autonomo. Queste fonti includono HTTP e OpenTelemetry per i log, le tracce e le metriche. Per abilitare il buffering persistente, scegli Abilita buffer persistente quando crei o aggiorni una pipeline. Per ulteriori informazioni, consulta the section called “Creazione di pipeline”.

OpenSearch L'ingestione determina dinamicamente il numero di file OCUs da utilizzare per il buffering persistente, tenendo conto dell'origine dei dati, delle trasformazioni di streaming e della

destinazione del sink. Poiché ne alloca una parte OCUs al buffering, potrebbe essere necessario aumentare il valore minimo e massimo per mantenere lo stesso throughput di ingestione. OCUs Le pipeline conservano i dati nel buffer per un massimo di 72 ore.

Se abiliti il buffering persistente per una pipeline, le dimensioni massime predefinite del payload delle richieste sono le seguenti:

- Fonti HTTP: 10 MB
- OpenTelemetry sorgenti — 4 MB

Per le sorgenti HTTP, è possibile aumentare la dimensione massima del payload a 20 MB. La dimensione del payload della richiesta include l'intera richiesta HTTP, che in genere contiene più eventi. Ogni evento non può superare i 3,5 MB.

Le pipeline con buffering persistente dividono le unità di pipeline configurate tra unità di calcolo e unità buffer. Se una pipeline utilizza un processore che richiede un uso intensivo della CPU come grok, key-value o split string, alloca le unità in un rapporto 1:1. buffer-to-compute Altrimenti, le alloca in un rapporto 3:1, privilegiando sempre le unità di calcolo.

Esempio:

- Pipeline con grok e 2 unità max: 1 unità di calcolo e 1 unità buffer
- Pipeline con grok e 5 unità max: 3 unità di calcolo e 2 unità buffer
- Pipeline senza processori e massimo 2 unità: 1 unità di calcolo e 1 unità buffer
- Pipeline senza processori e massimo 4 unità: 1 unità di calcolo e 3 unità buffer
- Pipeline con grok e 5 unità max: 2 unità di calcolo e 3 unità buffer

Per impostazione predefinita, le pipeline utilizzano an Chiave di proprietà di AWS per crittografare i dati del buffer. Queste pipeline non necessitano di autorizzazioni aggiuntive per il ruolo della pipeline.

In alternativa, puoi specificare una chiave gestita dal cliente e aggiungere le seguenti autorizzazioni IAM al ruolo della pipeline:

JSON

```
{
```

```
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "KeyAccess",
            "Effect": "Allow",
            "Action": [
              "kms:Decrypt",
              "kms:GenerateDataKeyWithoutPlaintext"
            ],
            "Resource": "arn:aws:kms:us-
east-1:111122223333:key/ASIAIOSFODNN7EXAMPLE"
        }
    ]
}
```

Per ulteriori informazioni, consulta Customer managed keys nella Guida per sviluppatori AWS Key Management Service .

> **Note**
>
> Se disabiliti il buffering persistente, la pipeline inizia a funzionare interamente con il buffering in memoria.

Divisione

È possibile configurare una pipeline di OpenSearch ingestione per suddividere gli eventi in entrata in una sottopipeline, in modo da eseguire diversi tipi di elaborazione sullo stesso evento in entrata.

La pipeline di esempio seguente divide gli eventi in entrata in due sotto-pipeline. Ogni sottopipeline utilizza il proprio processore per arricchire e manipolare i dati, quindi invia i dati a indici diversi. OpenSearch

```
version: "2"
log-pipeline:
  source:
    http:
    ...
  sink:
    - pipeline:
```

```
        name: "logs_enriched_one_pipeline"
    - pipeline:
        name: "logs_enriched_two_pipeline"

logs_enriched_one_pipeline:
  source:
    pipeline:
      name: "log-pipeline"
  processor:
   ...
  sink:
    - opensearch:
        # Provide a domain or collection endpoint
        # Enable the 'serverless' flag if the sink is an OpenSearch Serverless
 collection
        aws:
          ...
        index: "enriched_one_logs"

logs_enriched_two_pipeline:
  source:
    pipeline:
      name: "log-pipeline"
  processor:
   ...
  sink:
    - opensearch:
        # Provide a domain or collection endpoint
        # Enable the 'serverless' flag if the sink is an OpenSearch Serverless
 collection
        aws:
          ...
          index: "enriched_two_logs"
```

Concatenamento

È possibile concatenare più sotto-pipeline per eseguire l'elaborazione e l'arricchimento dei dati in blocchi. In altre parole, è possibile arricchire un evento in entrata con determinate funzionalità di elaborazione in una sottopipeline, quindi inviarlo a un'altra sottopipeline per un ulteriore arricchimento con un processore diverso e infine inviarlo al relativo sink. OpenSearch

Nell'esempio seguente, la `log_pipeline` sub-pipeline arricchisce un evento di registro in entrata con un set di processori, quindi invia l'evento a un indice denominato. OpenSearch `enriched_logs`

La pipeline invia lo stesso evento alla `log_advanced_pipeline` pipeline secondaria, che lo elabora e lo invia a un indice diverso denominato. OpenSearch `enriched_advanced_logs`

```
version: "2"
log-pipeline:
  source:
    http:
    ...
  processor:
    ...
  sink:
    - opensearch:
        # Provide a domain or collection endpoint
        # Enable the 'serverless' flag if the sink is an OpenSearch Serverless
 collection
        aws:
          ...
          index: "enriched_logs"
    - pipeline:
        name: "log_advanced_pipeline"

log_advanced_pipeline:
  source:
    pipeline:
      name: "log-pipeline"
  processor:
   ...
  sink:
    - opensearch:
        # Provide a domain or collection endpoint
        # Enable the 'serverless' flag if the sink is an OpenSearch Serverless
 collection
        aws:
          ...
          index: "enriched_advanced_logs"
```

Code di lettere morte

Le code di lettere morte (DLQs) sono destinazioni di eventi che una pipeline non riesce a scrivere in un sink. In OpenSearch Ingestion, è necessario specificare un bucket Amazon S3 con autorizzazioni di scrittura appropriate da utilizzare come DLQ. Puoi aggiungere una configurazione DLQ a ogni sink all'interno di una pipeline. Quando una pipeline incontra errori di scrittura, crea oggetti DLQ nel

bucket S3 configurato. Gli oggetti DLQ esistono all'interno di un file JSON come una serie di eventi falliti.

Una pipeline scrive eventi nel DLQ quando viene soddisfatta una delle seguenti condizioni:

- Il numero massimo di tentativi per il OpenSearch lavandino è stato esaurito. OpenSearch L'ingestione richiede un minimo di 16 per questa impostazione.
- Il sink rifiuta gli eventi dovuti a una condizione di errore.

Configurazione

Per configurare una coda di lettere non scritte per una pipeline secondaria, scegli Abilita S3 DLQ quando configuri la destinazione del sink. Quindi, specifica le impostazioni richieste per la coda. Per ulteriori informazioni, consulta Configurazione nella documentazione di Data Prepper DLQ.

I file scritti in questo S3 DLQ hanno il seguente schema di denominazione:

```
dlq-v${version}-${pipelineName}-${pluginId}-${timestampIso8601}-${uniqueId}
```

Per istruzioni su come configurare manualmente il ruolo della pipeline per consentire l'accesso al bucket S3 su cui il DLQ scrive, consulta. the section called “Autorizzazioni per scrivere su Amazon S3 o su una coda di lettere non scritte”

Esempio

Considerate il seguente file DLQ di esempio:

```
dlq-v2-apache-log-pipeline-opensearch-2023-04-05T15:26:19.152938Z-e7eb675a-
f558-4048-8566-dac15a4f8343
```

Ecco un esempio di dati che non sono stati scritti nel sink e che vengono inviati al bucket DLQ S3 per ulteriori analisi:

```
Record_0
pluginId            "opensearch"
pluginName          "opensearch"
pipelineName        "apache-log-pipeline"
failedData
```

```
index    "logs"
indexId   null
status    0
message  "Number of retries reached the limit of max retries (configured value 15)"
document
log      "sample log"
timestamp     "2023-04-14T10:36:01.070Z"

Record_1
pluginId            "opensearch"
pluginName          "opensearch"
pipelineName        "apache-log-pipeline"
failedData
index               "logs"
indexId   null
status    0
message  "Number of retries reached the limit of max retries (configured value 15)"
document
log                 "another sample log"
timestamp           "2023-04-14T10:36:01.071Z"
```

Gestione degli indici

Amazon OpenSearch Ingestion offre molte funzionalità di gestione degli indici, tra cui le seguenti.

Creazione di indici

È possibile specificare un nome di indice in un sink di pipeline e OpenSearch Ingestion crea l'indice quando effettua il provisioning della pipeline. Se esiste già un indice, la pipeline lo utilizza per indicizzare gli eventi in entrata. Se si arresta e si riavvia una pipeline o se si aggiorna la configurazione YAML, la pipeline tenta di creare nuovi indici se non esistono già. Una pipeline non può mai eliminare un indice.

I sinks di esempio seguenti creano due indici quando viene eseguito il provisioning della pipeline:

```
sink:
  - opensearch:
      index: apache_logs
  - opensearch:
      index: nginx_logs
```

Generazione di nomi e modelli di indici

È possibile generare nomi di indice dinamici utilizzando variabili dai campi degli eventi in arrivo. Nella configurazione sink, usa il formato `string${}` per segnalare l'interpolazione delle stringhe e usa un puntatore JSON per estrarre i campi dagli eventi. Le opzioni per sono o. `index_type` `custom` `management_disabled` Poiché l'`index_type`impostazione predefinita è per i OpenSearch domini e `custom` `management_disabled` per le raccolte OpenSearch Serverless, può essere lasciata non impostata.

Ad esempio, la seguente pipeline seleziona il `metadataType` campo dagli eventi in entrata per generare i nomi degli indici.

```
pipeline:
  ...
  sink:
    opensearch:
      index: "metadata-${metadataType}"
```

La seguente configurazione continua a generare un nuovo indice ogni giorno o ogni ora.

```
pipeline:
  ...
  sink:
    opensearch:
      index: "metadata-${metadataType}-%{yyyy.MM.dd}"

pipeline:
  ...
  sink:
    opensearch:
      index: "metadata-${metadataType}-%{yyyy.MM.dd.HH}"
```

Il nome dell'indice può anche essere una stringa semplice con un modello data-ora come suffisso, ad esempio. `my-index-%{yyyy.MM.dd}` Quando il sink invia dati a OpenSearch, sostituisce il modello data-ora con l'ora UTC e crea un nuovo indice per ogni giorno, ad esempio. `my-index-2022.01.25` Per ulteriori informazioni, consultate la classe. DateTimeFormatter

Questo nome di indice può anche essere una stringa formattata (con o senza un suffisso del modello data-ora), ad esempio. `my-${index}-name` Quando il sink invia dati a OpenSearch, sostituisce

la `"${index}"` parte con il valore dell'evento in fase di elaborazione. Se il formato è`"${index1/index2/index3}"`, sostituisce il campo `index1/index2/index3` con il relativo valore nell'evento.

Generazione di un documento IDs

Una pipeline può generare un ID di documento durante l'indicizzazione dei documenti su. OpenSearch Può dedurre questi documenti IDs dai campi all'interno degli eventi in arrivo.

Questo esempio utilizza il `uuid` campo di un evento in entrata per generare un ID del documento.

```
pipeline:
  ...
  sink:
    opensearch:
      index_type: custom
      index: "metadata-${metadataType}-%{yyyy.MM.dd}"
      "document_id": "uuid"
```

Nell'esempio seguente, il processore Add entries unisce i campi `uuid` e `other_field` l'evento in entrata per generare un ID del documento.

L'`create`azione garantisce che i documenti identici IDs non vengano sovrascritti. La pipeline elimina i documenti duplicati senza alcun nuovo tentativo o evento DLQ. Si tratta di un'aspettativa ragionevole per gli autori della pipeline che utilizzano questa azione, poiché l'obiettivo è evitare l'aggiornamento dei documenti esistenti.

```
pipeline:
  ...
  processor:
   - add_entries:
      entries:
        - key: "my_doc_id_field"
          format: "${uuid}-${other_field}"
  sink:
    - opensearch:
        ...
        action: "create"
        document_id: "my_doc_id"
```

Potresti voler impostare l'ID del documento di un evento su un campo di un oggetto secondario. Nell'esempio seguente, il plugin OpenSearch sink utilizza l'oggetto secondario `info/id` per generare un ID del documento.

```
sink:
  - opensearch:
       ...
       document_id: info/id
```

Dato il seguente evento, la pipeline genererà un documento con il `_id` campo impostato su: `json001`

```
{
   "fieldA":"arbitrary value",
   "info":{
      "id":"json001",
      "fieldA":"xyz",
      "fieldB":"def"
   }
}
```

Generazione del routing IDs

È possibile utilizzare l'`routing_field`opzione all'interno del plug-in OpenSearch sink per impostare il valore di una proprietà di routing del documento (`_routing`) su un valore proveniente da un evento in arrivo.

Il routing supporta la sintassi del puntatore JSON, quindi sono disponibili anche campi annidati, non solo campi di primo livello.

```
sink:
  - opensearch:
       ...
       routing_field: metadata/id
       document_id: id
```

Dato il seguente evento, il plugin genera un documento con il campo impostato su: `_routing` `abcd`

```
{
   "id":"123",
```

```
   "metadata":{
      "id":"abcd",
      "fieldA":"valueA"
   },
   "fieldB":"valueB"
}
```

Per istruzioni su come creare modelli di indice che le pipeline possono utilizzare durante la creazione dell'indice, vedete Modelli di indice.

End-to-end riconoscimento

OpenSearch L'ingestione garantisce la durata e l'affidabilità dei dati tracciandone la trasmissione dalla fonte ai pozzi nelle condutture stateless mediante riconoscimento. end-to-end

> **Note**
>
> Attualmente, solo il plug-in sorgente S3 supporta il riconoscimento. end-to-end

Con il end-to-end riconoscimento, il plug-in pipeline source crea un set di riconoscimenti per monitorare un batch di eventi. Riceve un riconoscimento positivo quando tali eventi vengono inviati con successo ai rispettivi sink o un riconoscimento negativo quando nessuno degli eventi non può essere inviato ai rispettivi sink.

In caso di guasto o arresto anomalo di un componente della pipeline, o se una fonte non riceve una conferma, la fonte scade e intraprende le azioni necessarie, come riprovare o registrare l'errore. Se nella pipeline sono configurati più sink o più subpipeline, i riconoscimenti a livello di evento vengono inviati solo dopo che l'evento è stato inviato a tutti i sink di tutte le pipeline secondarie. Se un sink ha un DLQ configurato, i riconoscimenti tengono traccia anche degli eventi scritti nel DLQ. end-to-end

Per abilitare il end-to-end riconoscimento, espandi Opzioni aggiuntive nella configurazione di origine di Amazon S3 e scegli end-to-end Abilita il riconoscimento dei messaggi.

Contropressione alla fonte

Una pipeline può subire una contropressione quando è impegnata nell'elaborazione dei dati o se i suoi sink sono temporaneamente inattivi o rallentano l'acquisizione dei dati. OpenSearch Ingestion ha diversi modi di gestire la contropressione a seconda del plug-in di origine utilizzato da una pipeline.

Origine HTTP

Le pipeline che utilizzano il plug-in di origine HTTP gestiscono la contropressione in modo diverso a seconda del componente della pipeline che è congestionato:

- Buffer: quando i buffer sono pieni, la pipeline inizia a restituire lo stato HTTP `REQUEST_TIMEOUT` con il codice di errore 408 all'endpoint di origine. Quando i buffer vengono liberati, la pipeline riavvia l'elaborazione degli eventi HTTP.
- Thread di origine: quando tutti i thread di origine HTTP sono impegnati nell'esecuzione di richieste e la dimensione della coda delle richieste non elaborate ha superato il numero massimo consentito di richieste, la pipeline inizia a restituire lo stato HTTP `TOO_MANY_REQUESTS` con il codice di errore 429 all'endpoint di origine. Quando la coda delle richieste scende al di sotto della dimensione massima consentita, la pipeline riavvia l'elaborazione delle richieste.

OTel fonte

Quando i buffer sono pieni per le pipeline che utilizzano OpenTelemetry sorgenti (OTel log, OTel metriche e OTel trace), la pipeline inizia a restituire lo stato HTTP `REQUEST_TIMEOUT` con il codice di errore 408 all'endpoint di origine. Quando i buffer vengono liberati, la pipeline riprende a elaborare gli eventi.

Fonte S3

Quando i buffer sono pieni per le pipeline con una sorgente S3, le pipeline interrompono l'elaborazione delle notifiche SQS. Man mano che i buffer vengono liberati, le pipeline riprendono a elaborare le notifiche.

Se un sink è inattivo o non è in grado di inserire dati e la end-to-end conferma è abilitata per la fonte, la pipeline interrompe l'elaborazione delle notifiche SQS finché non riceve un riconoscimento riuscito da tutti i sink.

Creazione di pipeline Amazon OpenSearch Ingestion

Una pipeline è il meccanismo utilizzato da Amazon OpenSearch Ingestion per spostare i dati dalla fonte (da cui provengono i dati) al relativo sink (dove vanno i dati). In OpenSearch Ingestion, il sink sarà sempre un singolo dominio Amazon OpenSearch Service, mentre la fonte dei dati potrebbe essere costituita da client come Amazon S3, Fluent Bit o Collector. OpenTelemetry

Per ulteriori informazioni, consulta Pipelines nella documentazione. OpenSearch

Argomenti

- Prerequisiti e ruolo IAM richiesto
- Autorizzazioni IAM richieste
- Specificare la versione della pipeline
- Specificare il percorso di ingestione
- Creazione di pipeline
- Monitoraggio dello stato della creazione della pipeline
- Lavorare con i progetti

Prerequisiti e ruolo IAM richiesto

Per creare una pipeline OpenSearch di ingestione, è necessario disporre delle seguenti risorse:

- Un ruolo IAM, chiamato ruolo pipeline, che OpenSearch Ingestion assume per scrivere nel sink. Puoi creare questo ruolo in anticipo oppure puoi fare in modo che OpenSearch Ingestion lo crei automaticamente durante la creazione della pipeline.
- Un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless che funga da sink. Se stai scrivendo su un dominio, deve essere in esecuzione OpenSearch 1.0 o versione successiva oppure Elasticsearch 7.4 o versione successiva. Il sink deve disporre di una politica di accesso che conceda le autorizzazioni appropriate al ruolo della pipeline IAM.

Per istruzioni su come creare queste risorse, consulta i seguenti argomenti:

- the section called “Concedere alle pipeline l'accesso ai domini”
- the section called “Concedere alle pipeline l'accesso alle raccolte”

> **Note**
>
> Se stai scrivendo su un dominio che utilizza un controllo granulare degli accessi, devi completare alcuni passaggi aggiuntivi. Per informazioni, consulta the section called “Mappa il ruolo della pipeline (solo per i domini che utilizzano un controllo di accesso granulare)”.

Autorizzazioni IAM richieste

OpenSearch Ingestion utilizza le seguenti autorizzazioni IAM per creare pipeline:

- `osis:CreatePipeline`— Creare una pipeline.
- `osis:ValidatePipeline`— Verificare se la configurazione di una tubazione è valida.
- `iam:CreateRole`e `iam:AttachPolicy` — Fai in modo che OpenSearch Ingestion crei automaticamente il ruolo della pipeline per te.
- `iam:PassRole`— Passa il ruolo della pipeline a OpenSearch Ingestion in modo che possa scrivere dati nel dominio. Questa autorizzazione deve riguardare la risorsa relativa al ruolo della pipeline o semplicemente * se si prevede di utilizzare ruoli diversi in ciascuna pipeline.

Ad esempio, la seguente politica concede l'autorizzazione a creare una pipeline:

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Resource":"*",
         "Action":[
            "osis:CreatePipeline",
            "osis:ListPipelineBlueprints",
            "osis:ValidatePipeline"
         ]
      },
      {
         "Resource":[
            "arn:aws:iam::111122223333:role/pipeline-role"
         ],
         "Effect":"Allow",
         "Action":[
            "iam:CreateRole",
            "iam:AttachRolePolicy",
            "iam:PassRole"
         ]
      }
   ]
```

```
}
```

OpenSearch Ingestion include anche un'autorizzazione chiamata`osis:Ingest`, necessaria per inviare richieste firmate alla pipeline utilizzando Signature Version 4. Per ulteriori informazioni, consulta the section called “Creazione di un ruolo di importazione”.

> **Note**
>
> Inoltre, il primo utente che crea una pipeline in un account deve disporre delle autorizzazioni per l'azione. `iam:CreateServiceLinkedRole` Per ulteriori informazioni, consulta pipeline role resource.

Per ulteriori informazioni su ciascuna autorizzazione, vedere Azioni, risorse e chiavi di condizione per l' OpenSearch ingestione nel Service Authorization Reference.

Specificare la versione della pipeline

Quando crei una pipeline utilizzando l'editor di configurazione, devi specificare la versione principale di Data Prepper che verrà eseguita dalla pipeline. Per specificare la versione, includete l'`version`opzione nella configurazione della pipeline:

```
version: "2"
log-pipeline:
  source:
    ...
```

Quando scegliete Crea, OpenSearch Ingestion determina l'ultima versione secondaria disponibile della versione principale specificata e fornisce la pipeline con quella versione. Ad esempio, se si specifica e l'ultima versione supportata di Data Prepper è la 2.1.1`version: "2"`, OpenSearch Ingestion esegue il provisioning della pipeline con la versione 2.1.1. Non mostriamo pubblicamente la versione secondaria in esecuzione nella tua pipeline.

Per aggiornare la pipeline quando è disponibile una nuova versione principale di Data Prepper, modifica la configurazione della pipeline e specifica la nuova versione. Non puoi effettuare il downgrade di una pipeline a una versione precedente.

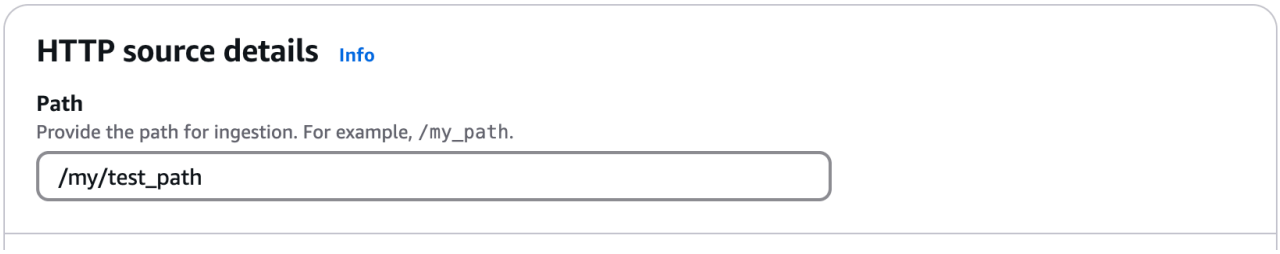
> **Note**
>
> OpenSearch Ingestion non supporta immediatamente le nuove versioni di Data Prepper non appena vengono rilasciate. Si verificherà un certo ritardo tra il momento in cui una nuova versione sarà disponibile pubblicamente e il momento in cui sarà supportata in Ingestion. OpenSearch Inoltre, OpenSearch Ingestion potrebbe esplicitamente non supportare del tutto determinate versioni principali o secondarie. Per un elenco completo, consulta the section called “Versioni di Data Prepper supportate”.

Ogni volta che apporti una modifica alla pipeline che avvia una blue/green distribuzione, OpenSearch Ingestion può aggiornarla all'ultima versione secondaria della versione principale attualmente configurata per la pipeline. Per ulteriori informazioni, vedere. the section called “Implementazioni blu/verdi per gli aggiornamenti della pipeline” OpenSearch Ingestion non può modificare la versione principale della pipeline a meno che non aggiorni esplicitamente l'`version`opzione all'interno della configurazione della pipeline.

Specificare il percorso di ingestione

Per le fonti basate su pull come OTel trace e OTel metrics, OpenSearch Ingestion richiede l'opzione aggiuntiva nella configurazione del codice sorgente. `path` Il percorso è una stringa come`/log/ingest`, che rappresenta il percorso URI per l'ingestione. Questo percorso definisce l'URI utilizzato per inviare dati alla pipeline.

Ad esempio, supponiamo di specificare il seguente percorso per una pipeline con un'origine HTTP:



Quando si inseriscono dati nella pipeline, è necessario specificare il seguente endpoint nella configurazione del client:. `https://`*`pipeline-name-abc123`*`.us-west-2.osis.amazonaws.com/`*`my/test_path`*

Il percorso deve iniziare con una barra (/) e può contenere i caratteri speciali '-', '_', ' . 'e'/', oltre al `${pipelineName}` segnaposto. Se si utilizza `${pipelineName}` (ad esempio/

`${pipelineName}/test_path`), OpenSearch Ingestion sostituisce la variabile con il nome della sottopipeline associata.

Creazione di pipeline

Questa sezione descrive come creare pipeline di OpenSearch ingestione utilizzando la console di OpenSearch servizio e il. AWS CLI

Console

Per creare una pipeline, accedi alla console di Amazon OpenSearch Service su https://console.aws.amazon.com/aos/osis/home e scegli Create pipeline.

Seleziona una pipeline vuota o scegli un blueprint di configurazione. I blueprint includono una pipeline preconfigurata per una varietà di casi d'uso comuni. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Scegli Seleziona blueprint.

Configura la fonte

1. Se parti da una pipeline vuota, seleziona una fonte dal menu a discesa. Le fonti disponibili potrebbero includere altre Servizi AWS fonti o OpenTelemetry HTTP. Per ulteriori informazioni, consulta the section called “Integrazione delle pipeline”.
2. A seconda della fonte scelta, configura impostazioni aggiuntive per la fonte. Ad esempio, per utilizzare Amazon S3 come sorgente, devi specificare l'URL della coda Amazon SQS dalla pipeline di ricezione dei messaggi. Per un elenco dei plugin di origine supportati e dei collegamenti alla relativa documentazione, consulta. the section called “Plugin e opzioni supportati”
3. Per alcune fonti, è necessario specificare le opzioni di rete di origine. Scegli tra Accesso VPC o Accesso pubblico. Se si sceglie Public access (Accesso pubblico), andare al passaggio successivo. Se scegli l'accesso VPC, configura le seguenti impostazioni:

Impostazione	Description
Gestione degli endpoint	Scegliete se volete creare voi stessi gli endpoint del cloud privato virtuale (VPC) o lasciateli creare da OpenSearch Ingestion per voi. Per impostazione predefinita, la gestione degli endpoint utilizza gli endpoint gestiti da Ingestion. OpenSearch

Impostazione	Description
VPC	Scegli l'ID del VPC che desideri utilizzare. Il VPC e la pipeline devono trovarsi nello stesso ambiente. Regione AWS
Sottoreti	Scegli una o più sottoreti. OpenSearch Il servizio inserirà un endpoint VPC e interfacce di rete elastiche nelle sottoreti.
Gruppi di sicurezza	Scegli uno o più gruppi di sicurezza VPC che consentano all'applicazione richiesta di raggiungere la pipeline di OpenSearch ingestione sulle porte (80 o 443) e sui protocolli (HTTP o HTTPs) esposti dalla pipeline.
Opzioni di collegamento VPC	Se la tua fonte è un endpoint autogestito, collega la pipeline a un VPC. Scegli una delle opzioni CIDR predefinite fornite o utilizza un CIDR personalizzato.

Per ulteriori informazioni, consulta the section called “Configurazione dell'accesso VPC per le pipeline”.

4. Scegli Next (Successivo).

Configura il processore

Aggiungi uno o più processori alla tua pipeline. I processori sono componenti all'interno di una sottopipeline che consentono di filtrare, trasformare e arricchire gli eventi prima di pubblicare i record nel dominio o nel collection sink. Per un elenco dei processori supportati e i collegamenti alla relativa documentazione, consulta. the section called “Plugin e opzioni supportati”

Puoi scegliere Azioni e aggiungere quanto segue:

- Routing condizionale: indirizza gli eventi a diversi sink in base a condizioni specifiche. Per ulteriori informazioni, consulta Routing condizionale.
- Sub-pipeline: ogni sub-pipeline è una combinazione di un'unica fonte, zero o più processori e un singolo sink. Solo una sottopipeline può avere un'origine esterna. Tutte le altre devono avere sorgenti che siano altre sotto-pipeline all'interno della configurazione generale della pipeline. Una configurazione a tubazione singola può contenere da 1 a 10 tubazioni secondarie.

Scegli Next (Successivo).

Configura sink

Seleziona la destinazione in cui la pipeline pubblica i record. Ogni tubazione secondaria deve contenere almeno un sink. È possibile aggiungere un massimo di 10 lavandini a una tubazione.

Per i OpenSearch lavandini, configurate i seguenti campi:

Impostazione	Description
Nome della politica di rete (Solo sink serverless)	Se hai selezionato una raccolta OpenSearch Serverless, inserisci un nome di policy di rete. OpenSearch Ingestion crea la policy se non esiste o la aggiorna con una regola che concede l'accesso all'endpoint VPC che collega la pipeline e la raccolta. Per ulteriori informazioni, consulta the section called “Concedere alle pipeline l'accesso alle raccolte”.
Nome dell'indice	Il nome dell'indice a cui la pipeline invia i dati. OpenSearch Ingestion crea questo indice se non esiste già.
Opzioni di mappatura dell'indice	Scegliete in che modo la pipeline archivia e indicizza i documenti e i relativi campi nel sink. OpenSearch Se selezioni Mappatura dinamica, OpenSearch aggiunge automaticamente i campi quando indicizzi un documento. Se selezioni Personalizza mappatura, inserisci un modello di mappatura dell'indice. Per ulteriori informazioni, consulta Modelli di indice.
Abilita DLQ	Configura una dead-letter queue (DLQ) di Amazon S3 per la pipeline. Per ulteriori informazioni, consulta the section called “Code di lettere morte”.
Impostazioni aggiuntive	Configura le opzioni avanzate per il OpenSearch lavandino. Per ulteriori informazioni, consulta Opzioni di configurazione nella documentazione di Data Prepper.

Per aggiungere un sink Amazon S3, scegli Aggiungi sink e Amazon S3. Per ulteriori informazioni, consulta the section called “Amazon S3 come destinazione”.

Scegli Next (Successivo).

Configura la pipeline

Configurate le seguenti impostazioni aggiuntive della pipeline:

Impostazione	Description
Nome della pipeline	Un nome univoco per la pipeline.
Buffer persistente	Un buffer persistente archivia i dati in un buffer basato su disco su più zone di disponibilità. Per ulteriori informazioni, consulta the section called “Bufferin g persistente”. Se abiliti il buffering persistente, seleziona la AWS Key Management Service chiave per crittografare i dati del buffer.
Capacità della pipeline	La capacità minima e massima della pipeline, in Ingestion OpenSearch Compute Units (). OCUs Per ulteriori informazioni, consulta the section called “Scalabilità delle pipeline”.
Ruolo della pipeline	Il ruolo IAM che fornisce le autorizzazioni necessarie alla pipeline per scrivere nel sink e leggere da fonti basate su pull. Puoi creare il ruolo tu stesso o fare in modo che OpenSearch Ingestion lo crei per te in base al caso d'uso selezionato. Per ulteriori informazioni, consulta the section called “Configurazione di ruoli e utenti”.
Tag	Aggiungi uno o più tag alla tua pipeline. Per ulteriori informazioni, consulta the section called “Etichettatura delle pipeline”.
Opzioni di pubblicazione dei log	Abilita la pubblicazione dei log della pipeline su Amazon CloudWatch Logs. Ti consigliamo di abilitare la pubblicazione dei log in modo da poter risolvere più facilmente i problemi relativi alla pipeline. Per ulteriori informazioni, consulta the section called “Monitoraggio dei log della pipeline”.

Scegli Avanti., quindi rivedi la configurazione della pipeline e scegli Crea pipeline.

OpenSearch Ingestion esegue un processo asincrono per creare la pipeline. Una volta raggiunto lo stato della pipeline, puoi iniziare a importare i dati`Active`.

AWS CLI

Il comando create-pipeline accetta la configurazione della pipeline come stringa o all'interno di un file .yaml o .json. Se fornite la configurazione come stringa, ogni nuova riga deve essere scappata con. `\n` Ad esempio, `"log-pipeline:\n source:\n http:\n processor:\n - grok:\n ..."`

Il seguente comando di esempio crea una pipeline con la seguente configurazione:

- Minimo 4 ingestioni OCUs, massimo 10 ingestioni OCUs
- Fornito all'interno di un cloud privato virtuale (VPC)
- Pubblicazione dei log abilitata

```
aws osis create-pipeline \
  --pipeline-name my-pipeline \
  --min-units 4 \
  --max-units 10 \
  --log-publishing-options
 IsLoggingEnabled=true,CloudWatchLogDestination={LogGroup="MyLogGroup"} \
  --vpc-options
 SecurityGroupIds={sg-12345678,sg-9012345},SubnetIds=subnet-1212234567834asdf \
  --pipeline-configuration-body "file://pipeline-config.yaml" \
  --pipeline-role-arn  arn:aws:iam::1234456789012:role/pipeline-role
```

OpenSearch Ingestion esegue un processo asincrono per creare la pipeline. Una volta raggiunto lo stato della pipeline, puoi iniziare a importare i dati`Active`. Per controllare lo stato della pipeline, utilizzate il comando. GetPipeline

OpenSearch API di ingestione

Per creare una pipeline OpenSearch di ingestione utilizzando l'API Ingestion, chiama l' OpenSearch operazione. CreatePipeline

Dopo aver creato correttamente la pipeline, puoi configurare il client e iniziare a importare dati nel tuo dominio di servizio. OpenSearch Per ulteriori informazioni, consulta the section called “Integrazione delle pipeline”.

Monitoraggio dello stato della creazione della pipeline

È possibile tenere traccia dello stato di una pipeline man mano che OpenSearch Ingestion la rifornisce e la prepara per l'importazione dei dati.

Console

Dopo aver creato inizialmente una pipeline, questa passa attraverso più fasi man mano OpenSearch che Ingestion la prepara per l'inserimento dei dati. Per visualizzare le varie fasi della creazione della pipeline, scegliete il nome della pipeline per visualizzarne la pagina delle impostazioni della pipeline. In Stato, scegliete Visualizza dettagli.

Una pipeline passa attraverso le seguenti fasi prima di essere disponibile per l'acquisizione dei dati:

- Convalida: convalida della configurazione della pipeline. Una volta completata questa fase, tutte le convalide hanno avuto esito positivo.
- Crea ambiente: preparazione e approvvigionamento delle risorse. Al termine di questa fase, è stato creato il nuovo ambiente di pipeline.
- Distribuisci pipeline: distribuzione della pipeline. Una volta completata questa fase, la pipeline è stata implementata con successo.
- Verifica dello stato della pipeline: verifica dello stato della pipeline. Una volta completata questa fase, tutti i controlli sanitari sono stati superati.
- Abilita il traffico: consente alla pipeline di importare dati. Una volta completata questa fase, puoi iniziare a importare i dati nella pipeline.

CLI

Utilizzate il get-pipeline-change-progresscomando per controllare lo stato di una pipeline. La seguente AWS CLI richiesta verifica lo stato di una pipeline denominata: `my-pipeline`

```
aws osis get-pipeline-change-progress \
    --pipeline-name my-pipeline
```

Risposta:

```
{
   "ChangeProgressStatuses": {
```

```
      "ChangeProgressStages": [
         {
            "Description": "Validating pipeline configuration",
            "LastUpdated": 1.671055851E9,
            "Name": "VALIDATION",
            "Status": "PENDING"
         }
      ],
      "StartTime": 1.671055851E9,
      "Status": "PROCESSING",
      "TotalNumberOfStages": 5
   }
}
```

OpenSearch API di ingestione

Per tenere traccia dello stato della creazione della pipeline utilizzando l'API OpenSearch Ingestion, chiama l'operazione. GetPipelineChangeProgress

Lavorare con i progetti

Invece di creare una definizione di pipeline partendo da zero, puoi utilizzare i blueprint di configurazione, che sono modelli preconfigurati per scenari di ingestione comuni come Trace Analytics o i log di Apache. I blueprint di configurazione consentono di effettuare facilmente il provisioning delle pipeline senza dover creare una configurazione da zero.

Console

Per utilizzare un blueprint di pipeline

1. Accedi alla console OpenSearch Ingestion da osis/home. https://console.aws.amazon.com/aos/ Ti troverai nella pagina Pipelines.
2. Scegliere Create pipeline (Crea pipeline).
3. Seleziona un blueprint dall'elenco dei casi d'uso, quindi scegli Seleziona blueprint. La configurazione della pipeline viene compilata con una pipeline secondaria per il caso d'uso selezionato.

   Il blueprint della pipeline non è valido così com'è. È necessario specificare impostazioni aggiuntive a seconda della fonte selezionata.

CLI

Per ottenere un elenco di tutti i progetti disponibili utilizzando il AWS CLI, invia una list-pipeline-blueprintsrichiesta.

```
aws osis list-pipeline-blueprints
```

La richiesta restituisce un elenco di tutti i blueprint disponibili.

Per ottenere informazioni più dettagliate su un progetto specifico, usa il get-pipeline-blueprintcomando:

```
aws osis get-pipeline-blueprint --blueprint-name AWS-ApacheLogPipeline
```

Questa richiesta restituisce il contenuto del blueprint della pipeline di log di Apache:

```
{
   "Blueprint":{
      "PipelineConfigurationBody":"###\n  # Limitations: https://docs.aws.amazon.com/
opensearch-service/latest/ingestion/ingestion.html#ingestion-limitations\n###\n###\n
 # apache-log-pipeline:\n    # This pipeline receives logs via http (e.g. FluentBit),
 extracts important values from the logs by matching\n    # the value in the 'log' key
 against the grok common Apache log pattern. The grokked logs are then sent\n    # to
 OpenSearch to an index named 'logs'\n###\n\nversion: \"2\"\napache-log-pipeline:\n
 source:\n    http:\n      # Provide the path for ingestion. ${pipelineName} will be
 replaced with pipeline name configured for this pipeline.\n      # In this case it
 would be \"/apache-log-pipeline/logs\". This will be the FluentBit output URI value.
\n      path: \"/${pipelineName}/logs\"\n  processor:\n    - grok:\n        match:\n
         log: [ \"%{COMMONAPACHELOG_DATATYPED}\" ]\n  sink:\n    - opensearch:\n
   # Provide an AWS OpenSearch Service domain endpoint\n        # hosts: [ \"https://
search-mydomain-1a2a3a4a5a6a7a8a9a0a9a8a7a.us-east-1.es.amazonaws.com\" ]\n        aws:
\n          # Provide the region of the domain.\n          # region: \"us-east-1\"\n
        # Enable the 'serverless' flag if the sink is an Amazon OpenSearch Serverless
 collection\n          # serverless: true\n        index: \"logs\"\n        # Enable
 the S3 DLQ to capture any failed requests in an S3 bucket\n        # dlq:\n
  # s3:\n            # Provide an S3 bucket\n            # bucket: \"your-dlq-bucket-
name\"\n            # Provide a key path prefix for the failed requests\n
  # key_path_prefix: \"${pipelineName}/logs/dlq\"\n            # Provide the region
 of the bucket.\n            # region: \"us-east-1\"\n            # Provide a Role
 ARN with access to the bucket. This role should have a trust relationship with osis-
pipelines.amazonaws.com\n"
```

```
      "BlueprintName":"AWS-ApacheLogPipeline"
   }
}
```

OpenSearch API di ingestione

Per ottenere informazioni sui progetti di pipeline utilizzando l'API OpenSearch Ingestion, utilizza le operazioni and. ListPipelineBlueprintsGetPipelineBlueprint

Visualizzazione delle pipeline di Amazon OpenSearch Ingestion

Puoi visualizzare i dettagli su una pipeline di Amazon OpenSearch Ingestion utilizzando l' Console di gestione AWS, la o l' OpenSearch API AWS CLI Ingestion.

Console

Per visualizzare una pipeline

1. Accedi alla console di Amazon OpenSearch Service su https://console.aws.amazon.com/aos/osis/home. Ti troverai nella pagina Pipelines.
2. (Facoltativo) Per visualizzare le pipeline con uno stato particolare, scegli Qualsiasi stato e seleziona uno stato in base al quale filtrare.

   Una tubazione può avere i seguenti stati:

   - `Active`— La pipeline è attiva e pronta per l'inserimento dei dati.
   - `Creating`— La pipeline è in fase di creazione.
   - `Updating`— La pipeline è in fase di aggiornamento.
   - `Deleting`— La pipeline viene eliminata.
   - `Create failed`— La pipeline non può essere creata.
   - `Update failed`— La pipeline non può essere aggiornata.
   - `Stop failed`— La pipeline non può essere interrotta.
   - `Start failed`— La pipeline non può essere avviata.
   - `Stopping`— La pipeline viene interrotta.
   - `Stopped`— La pipeline è interrotta e può essere riavviata in qualsiasi momento.
   - `Starting`— La pipeline si sta avviando.

Non ti viene addebitata la fattura per l'ingestione OCUs quando una pipeline si trova negli statiCreate failed,, Creating e. Deleting Stopped

CLI

Per visualizzare le pipeline utilizzando, invia una richiesta list-pipelines AWS CLI:

```
aws osis list-pipelines
```

La richiesta restituisce un elenco di tutte le pipeline esistenti:

```
{
    "NextToken": null,
    "Pipelines": [
        {,
            "CreatedAt": 1.671055851E9,
            "LastUpdatedAt": 1.671055851E9,
            "MaxUnits": 4,
            "MinUnits": 2,
            "PipelineArn": "arn:aws:osis:us-west-2:123456789012:pipeline/log-pipeline",
            "PipelineName": "log-pipeline",
            "Status": "ACTIVE",
            "StatusReason": {
                "Description": "The pipeline is ready to ingest data."
            }
        },
            "CreatedAt": 1.671055851E9,
            "LastUpdatedAt": 1.671055851E9,
            "MaxUnits": 2,
            "MinUnits": 8,
            "PipelineArn": "arn:aws:osis:us-west-2:123456789012:pipeline/another-
pipeline",
            "PipelineName": "another-pipeline",
            "Status": "CREATING",
            "StatusReason": {
                "Description": "The pipeline is being created. It is not able to ingest
 data."
            }
        }
    ]
}
```

Per ottenere informazioni su una singola pipeline, usa il comando get-pipeline:

```
aws osis get-pipeline --pipeline-name "my-pipeline"
```

La richiesta restituisce informazioni di configurazione per la pipeline specificata:

```
{
    "Pipeline": {
        "PipelineName": "my-pipeline",
        "PipelineArn": "arn:aws:osis:us-east-1:123456789012:pipeline/my-pipeline",
        "MinUnits": 9,
        "MaxUnits": 10,
        "Status": "ACTIVE",
        "StatusReason": {
            "Description": "The pipeline is ready to ingest data."
        },
        "PipelineConfigurationBody": "log-pipeline:\n source:\n http:\n processor:\n
 - grok:\n match:\nlog: [ '%{COMMONAPACHELOG}' ]\n - date:\n from_time_received: true
\n destination: \"@timestamp\"\n  sink:\n - opensearch:\n hosts: [ \"https://search-
mdp-performance-test-duxkb4qnycd63rpy6svmvyvfpi.us-east-1.es.amazonaws.com\" ]\n index:
 \"apache_logs\"\n aws_sts_role_arn: \"arn:aws:iam::123456789012:role/my-domain-role
\"\n  aws_region: \"us-east-1\"\n  aws_sigv4: true",,
        "CreatedAt": "2022-10-01T15:28:05+00:00",
        "LastUpdatedAt": "2022-10-21T21:41:08+00:00",
        "IngestEndpointUrls": [
            "my-pipeline-123456789012.us-east-1.osis.amazonaws.com"
        ]
    }
}
```

OpenSearch API di ingestione

Per visualizzare le pipeline OpenSearch di Ingestion utilizzando l'API OpenSearch Ingestion, chiamate le operazioni and. ListPipelinesGetPipeline

Aggiornamento delle pipeline di Amazon OpenSearch Ingestion

Puoi aggiornare le pipeline OpenSearch di Amazon Ingestion utilizzando l' Console di gestione AWS, la o l' OpenSearch API AWS CLI Ingestion. OpenSearch Ingestion avvia una distribuzione quando aggiorni una blue/green configurazione di pipeline. Per ulteriori informazioni, consulta the section called “Implementazioni blu/verdi per gli aggiornamenti della pipeline”.

Argomenti

- Considerazioni
- Autorizzazioni richieste
- Aggiornamento delle pipeline
- Implementazioni blu/verdi per gli aggiornamenti della pipeline

Considerazioni

Quando aggiorni una pipeline, tenete presente quanto segue:

- Non puoi aggiornare il nome o le impostazioni di rete di una pipeline.
- Se la pipeline scrive su un sink di dominio VPC, non puoi tornare indietro e modificare il sink in un dominio VPC diverso dopo la creazione della pipeline. È necessario eliminare e ricreare la pipeline con il nuovo sink. Puoi comunque passare il sink da un dominio VPC a un dominio pubblico, da un dominio pubblico a un dominio VPC o da un dominio pubblico a un altro dominio pubblico.
- È possibile passare il sink della pipeline in qualsiasi momento da un dominio di OpenSearch servizio pubblico a una raccolta Serverless. OpenSearch
- Quando si aggiorna la configurazione di origine, processore o sink per una pipeline, OpenSearch Ingestion avvia una distribuzione. blue/green Per ulteriori informazioni, consulta the section called “Implementazioni blu/verdi per gli aggiornamenti della pipeline”.
- Quando si aggiorna la configurazione di origine, processore o sink per una pipeline, OpenSearch Ingestion aggiorna automaticamente la pipeline all'ultima versione secondaria supportata della versione principale di Data Prepper utilizzata nella pipeline. Questo processo mantiene la pipeline aggiornata con le ultime correzioni di bug e i miglioramenti delle prestazioni.
- Puoi comunque aggiornare la tua pipeline quando viene interrotta.

Autorizzazioni richieste

OpenSearch Ingestion utilizza le seguenti autorizzazioni IAM per l'aggiornamento delle pipeline:

- `osis:UpdatePipeline`— Aggiornare una pipeline.
- `osis:ValidatePipeline`— Verificare se la configurazione di una tubazione è valida.
- `iam:PassRole`— Passa il ruolo della pipeline a OpenSearch Ingestion in modo che possa scrivere dati nel dominio. Questa autorizzazione è richiesta solo se stai aggiornando la configurazione della pipeline, non se stai modificando altre impostazioni come la pubblicazione dei log o i limiti di capacità.

Ad esempio, la seguente politica concede l'autorizzazione ad aggiornare una pipeline:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Resource": "*",
            "Action": [
                "osis:UpdatePipeline",
                "osis:ValidatePipeline"
            ]
        },
        {
            "Resource": [
                "arn:aws:iam::111122223333:role/pipeline-role"
            ],
            "Effect": "Allow",
            "Action": [
                "iam:PassRole"
            ]
        }
    ]
}
```

Aggiornamento delle pipeline

Puoi aggiornare le pipeline OpenSearch di Amazon Ingestion utilizzando l' Console di gestione AWS, la o l' OpenSearch API AWS CLI Ingestion.

Console

Come aggiornare una pipeline

1. Accedi alla console di Amazon OpenSearch Service su https://console.aws.amazon.com/aos/osis/home. Ti troverai nella pagina Pipelines.
2. Scegli una pipeline per aprirne le impostazioni. Quindi, scegliete una delle opzioni di modifica.
3. Una volta completate le modifiche, scegli Save (Salva).

CLI

Per aggiornare una pipeline utilizzando AWS CLI, invia una richiesta di update-pipeline. La seguente richiesta di esempio carica un nuovo file di configurazione e aggiorna i valori di capacità minima e massima:

```
aws osis update-pipeline \
  --pipeline-name "my-pipeline" \
  --pipline-configuration-body "file://new-pipeline-config.yaml" \
  --min-units 11 \
  --max-units 18
```

OpenSearch API di ingestione

Per aggiornare una pipeline OpenSearch di ingestione utilizzando l'API Ingestion, chiamate l' OpenSearch operazione. UpdatePipeline

Implementazioni blu/verdi per gli aggiornamenti della pipeline

OpenSearch L'ingestione avvia un processo di distribuzione blu/verde quando si aggiorna una configurazione della pipeline.

Blue/green refers to the practice of creating a new environment for pipeline updates and routing traffic to the new environment after those updates are complete. The practice minimizes downtime and maintains the original environment in the event that deployment to the new environment is unsuccessful. Blue/greenle distribuzioni di per sé non hanno alcun impatto sulle prestazioni, ma le prestazioni potrebbero cambiare se la configurazione della pipeline cambia in modo tale da alterare le prestazioni.

OpenSearch L'ingestione blocca l'auto-scaling durante le implementazioni. blue/green Continuerai a ricevere addebiti solo per il traffico verso la vecchia pipeline fino a quando non viene reindirizzato alla nuova pipeline. Una volta che il traffico è stato reindirizzato, ti verrà addebitato solo il costo della nuova pipeline. Non ti vengono mai addebitati costi per due pipeline contemporaneamente.

Quando aggiorni la configurazione di origine, processore o sink di una pipeline, OpenSearch Ingestion può aggiornare automaticamente la pipeline all'ultima versione secondaria supportata della versione principale in esecuzione sulla pipeline. Ad esempio, potreste avere `version: "2"` nella configurazione della pipeline e OpenSearch Ingestion inizialmente aveva fornito alla pipeline la versione 2.1.0. Quando viene aggiunto il supporto per la versione 2.1.1 e si apporta una modifica alla configurazione della pipeline, Ingestion aggiorna la pipeline alla versione 2.1.1. OpenSearch

Questo processo mantiene la pipeline aggiornata con le ultime correzioni di bug e i miglioramenti delle prestazioni. OpenSearch Ingestion non può aggiornare la versione principale della pipeline a meno che non si modifichi manualmente l'`version`opzione all'interno della configurazione della pipeline.

Gestione dei costi della OpenSearch pipeline di Amazon Ingestion

Puoi avviare e interrompere le pipeline di importazione in Amazon OpenSearch Ingestion per controllare il flusso di dati in base alle tue esigenze. L'arresto di una pipeline interrompe l'elaborazione dei dati preservando le configurazioni, quindi puoi riavviarla senza riconfigurarla. Questo può aiutare a ottimizzare i costi, gestire l'utilizzo delle risorse o risolvere i problemi. Quando si interrompe una pipeline, OpenSearch Ingestion non elabora i dati in entrata, ma i dati precedentemente importati rimangono disponibili in. OpenSearch

L'avvio e l'arresto semplificano i processi di configurazione e smontaggio delle pipeline utilizzate per lo sviluppo, il test o attività simili che non richiedono una disponibilità continua. Mentre la pipeline è interrotta, non ti viene addebitato alcun costo per le ore di Ingestion OCU. Puoi comunque aggiornare le pipeline interrotte e queste riceveranno aggiornamenti automatici delle versioni secondarie e patch di sicurezza.

L'arresto e l'avvio di una pipeline comporteranno la rielaborazione di tutti i dati dall'inizio per le pipeline basate su pull (DDB, S3, DocDB, ecc.). Quando si arresta una pipeline, tutti gli endpoint VPC gestiti dal servizio creati dalla pipeline vengono rimossi. Per le pipeline con endpoint VPC autogestiti, devi ricreare l'endpoint VPC nel tuo account quando riavvii la pipeline. Per ulteriori informazioni, consulta the section called “Endpoint VPC autogestiti”.

> **Note**
>
> Se la tua pipeline ha una capacità in eccesso ma deve rimanere operativa, valuta la possibilità di modificare i limiti di capacità massima anziché interromperla e riavviarla. Questo può aiutare a gestire i costi garantendo al contempo che la pipeline continui a elaborare i dati in modo efficiente. Per ulteriori dettagli, consultare the section called “Scalabilità delle pipeline”.

I seguenti argomenti spiegano come avviare e arrestare le pipeline utilizzando l'API Console di gestione AWS e AWS CLI OpenSearch Ingestion.

Argomenti

- Interruzione di una pipeline di Amazon OpenSearch Ingestion
- Avvio di una pipeline Amazon OpenSearch Ingestion

Interruzione di una pipeline di Amazon OpenSearch Ingestion

Per utilizzare una pipeline di OpenSearch ingestione o eseguire l'amministrazione, devi sempre iniziare con una pipeline attiva, quindi arrestare la pipeline e riavviarla. Mentre la pipeline è interrotta, non ti vengono addebitati i costi per le ore di Ingestion OCU.

Console

Per fermare una pipeline

1. Accedi alla console di Amazon OpenSearch Service su https://console.aws.amazon.com/aos/osis/home. Ti troverai nella pagina Pipelines.
2. Scegli una pipeline. È possibile eseguire l'operazione di arresto da questa pagina o passare alla pagina dei dettagli della tubazione che si desidera interrompere.
3. Per Azioni, scegli Stop pipeline.

   Se una pipeline non può essere interrotta e avviata, l'azione Stop pipeline non è disponibile.

AWS CLI

Per arrestare una pipeline utilizzando il AWS CLI, chiamate il comando stop-pipeline con i seguenti parametri:

- `--pipeline-name`— il nome della pipeline.

Example

```
aws osis stop-pipeline --pipeline-name my-pipeline
```

OpenSearch API di ingestione

Per interrompere una pipeline utilizzando l'API OpenSearch Ingestion, richiamate l'StopPipelineoperazione con il seguente parametro:

- `PipelineName`— il nome della pipeline.

Avvio di una pipeline Amazon OpenSearch Ingestion

Avviate sempre una pipeline OpenSearch di ingestione a partire da una pipeline che si trova già nello stato di interruzione. La pipeline mantiene le impostazioni di configurazione, come i limiti di capacità, le impostazioni di rete e le opzioni di pubblicazione dei log.

Il riavvio di una pipeline richiede in genere diversi minuti.

Console

Per avviare una pipeline

1. Accedi alla console di Amazon OpenSearch Service su https://console.aws.amazon.com/aos/osis/home. Ti troverai nella pagina Pipelines.
2. Scegli una pipeline. È possibile eseguire l'operazione di avvio da questa pagina o accedere alla pagina dei dettagli della pipeline che si desidera avviare.
3. Per Azioni, scegli Avvia pipeline.

AWS CLI

Per avviare una pipeline utilizzando il AWS CLI, chiamate il comando start-pipeline con i seguenti parametri:

- `--pipeline-name`— il nome della pipeline.

Example

```
aws osis start-pipeline --pipeline-name my-pipeline
```

OpenSearch API di ingestione

Per avviare una pipeline OpenSearch di ingestione utilizzando l'API OpenSearch Ingestion, richiamate l'operazione con il seguente parametro: StartPipeline

- `PipelineName`— il nome della pipeline.

Eliminazione delle pipeline di Amazon OpenSearch Ingestion

Puoi eliminare una pipeline di Amazon OpenSearch Ingestion utilizzando l' Console di gestione AWS, la o l' OpenSearch API AWS CLI Ingestion. Non puoi eliminare una pipeline se lo stato è o. `Creating` `Updating`

Console

Come eliminare una pipeline

1. Accedi alla console di Amazon OpenSearch Service su https://console.aws.amazon.com/aos/osis/home. Ti troverai nella pagina Pipelines.
2. Seleziona la pipeline che desideri eliminare e scegli Azioni, Elimina.
3. Per confermare l'eliminazione, scegliere Delete (Elimina).

CLI

Per eliminare una pipeline utilizzando AWS CLI, invia una richiesta di eliminazione della pipeline:

```
aws osis delete-pipeline --pipeline-name "my-pipeline"
```

OpenSearch API di ingestione

Per eliminare una pipeline OpenSearch di ingestione utilizzando l'API OpenSearch Ingestion, richiamate l'operazione con il seguente parametro: DeletePipeline

- `PipelineName`— il nome della pipeline.

Plugin e opzioni supportati per le pipeline di Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion supporta un sottoinsieme di sorgenti, processori e sink all'interno di Data Prepper open source. OpenSearch Inoltre, OpenSearch Ingestion impone alcuni vincoli alle opzioni disponibili per ogni plug-in supportato. Le sezioni seguenti descrivono i plugin e le opzioni associate supportati da Ingestion. OpenSearch

> **Note**
>
> OpenSearch Ingestion non supporta alcun plugin buffer perché configura automaticamente un buffer predefinito. Si riceve un errore di convalida se si include un buffer nella configurazione della pipeline.

Argomenti

- Plugin supportati
- Processori stateless e processori stateful
- Requisiti e vincoli di configurazione

Plugin supportati

OpenSearch Ingestion supporta i seguenti plugin Data Prepper:

Fonti:

- DocumentDB
- DynamoDB

- HTTP
- Kafka
- Kinesis
- OpenSearch
- OTel logs
- OTel metriche
- OTel traccia
- S3

Processori:

- Aggiungere voci
- Aggregazione

- Rilevatore di anomalie
- AWS Lambda
- Convertire il tipo di ingresso
- Copia i valori
- CSV
- Data
- Ritardo
- Decomprimere
- Eliminare le voci
- Sezionare
- Elimina gli eventi
- Appiattire
- IP geografico
- Grok
- Valore chiave
- Elenco da mappare
- Stringa minuscola
- Mappa da elencare
- Evento di mutazione (serie di processori)
- Stringa mutata (serie di processori)
- Offuscare
- OTel metriche
- OTel gruppo di tracciamento
- OTel traccia
- Analizza Ion
- Analizza JSON
- Analizza XML
- Rinomina le chiavi
- Seleziona le voci
- Mappa servizio

- Evento diviso
- Stringa divisa
- convertitore di stringhe
- Stringa sostitutiva
- Trace peer forwarder
- Translate (Trasla)
- Taglia la corda
- Troncare
- Stringa maiuscola
- Agente utente
- Scrivi JSON

Lavelli:

- OpenSearch(supporta OpenSearch Service, OpenSearch Serverless ed Elasticsearch 6.8 o versioni successive)
- S3

Codec Sink:

- Avro
- NDJSON
- JSON
- Parquet

Processori stateless e processori stateful

I processori stateless eseguono operazioni come trasformazioni e filtri, mentre i processori stateful eseguono operazioni come le aggregazioni, che ricordano il risultato dell'esecuzione precedente. OpenSearch Ingestion supporta i processori stateful Aggregate e Service-map. Tutti gli altri processori supportati sono stateless.

Per le pipeline che contengono solo processori stateless, il limite di capacità massima è 96 Ingestion. OCUs Se una pipeline contiene processori con stato, il limite di capacità massima è 48 Ingestion.

OCUs Tuttavia, se una pipeline ha il buffering persistente abilitato, può avere un massimo di 384 Ingestion OCUs con solo processori stateless o 192 Ingestion se contiene processori con stato. OCUs Per ulteriori informazioni, consulta the section called “Scalabilità delle pipeline”.

End-to-end il riconoscimento è supportato solo per i processori stateless. Per ulteriori informazioni, consulta the section called “End-to-end riconoscimento”.

Requisiti e vincoli di configurazione

Se non diversamente specificato di seguito, tutte le opzioni descritte nel riferimento alla configurazione di Data Prepper per i plugin supportati sopra elencati sono consentite nelle pipeline di Ingestion. OpenSearch Le sezioni seguenti spiegano i vincoli che Ingestion impone a determinate opzioni del plug-in. OpenSearch

> **Note**
>
> OpenSearch Ingestion non supporta alcun plugin buffer perché configura automaticamente un buffer predefinito. Si riceve un errore di convalida se si include un buffer nella configurazione della pipeline.

Molte opzioni sono configurate e gestite internamente da OpenSearch Ingestion, come e. `authentication` `acm_certificate_arn` Altre opzioni, come `thread_count` `erequest_timeout`, hanno un impatto sulle prestazioni se modificate manualmente. Pertanto, questi valori sono impostati internamente per garantire prestazioni ottimali delle pipeline.

Infine, alcune opzioni non possono essere passate a OpenSearch Ingestion, ad esempio `ism_policy_file` and`sink_template`, perché sono file locali se eseguite in Data Prepper open source. Questi valori non sono supportati.

Argomenti

- Opzioni generali della pipeline
- Processore Grok
- Origine HTTP
- OpenSearch lavandino
- OTel metrics source, OTel trace source e logs source OTel
- OTel processore trace group
- OTel processore di traccia

- Processore Service-map
- Fonte S3

Opzioni generali della pipeline

Le seguenti opzioni generali di pipeline sono impostate da OpenSearch Ingestion e non sono supportate nelle configurazioni della pipeline:

- `workers`
- `delay`

Processore Grok

Le seguenti opzioni del processore Grok non sono supportate:

- `patterns_directories`
- `patterns_files_glob`

Origine HTTP

Il plugin sorgente HTTP ha i seguenti requisiti e vincoli:

- L'*`path`*opzione è obbligatoria. Il percorso è una stringa come`/log/ingest`, che rappresenta il percorso URI per l'inserimento dei log. Questo percorso definisce l'URI utilizzato per inviare dati alla pipeline. Ad esempio, `https://log-pipeline.us-west-2.osis.amazonaws.com`*`/log/ingest`*. Il percorso deve iniziare con una barra (/) e può contenere i caratteri speciali '-', '_', '. 'e'/', oltre al `${pipelineName}` segnaposto.
- Le seguenti opzioni di origine HTTP sono impostate da OpenSearch Ingestion e non sono supportate nelle configurazioni della pipeline:
  - `port`
  - `ssl`
  - `ssl_key_file`
  - `ssl_certificate_file`
  - `aws_region`
  - `authentication`

- `unauthenticated_health_check`
- `use_acm_certificate_for_ssl`
- `thread_count`
- `request_timeout`
- `max_connection_count`
- `max_pending_requests`
- `health_check_service`
- `acm_private_key_password`
- `acm_certificate_timeout_millis`
- `acm_certificate_arn`

OpenSearch lavandino

Il plugin OpenSearchsink presenta i seguenti requisiti e limitazioni.

- L'`aws`opzione è obbligatoria e deve contenere le seguenti opzioni:
  - `sts_role_arn`
  - `region`
  - `hosts`
  - `serverless`(se il sink è una raccolta OpenSearch Serverless)
- L'`sts_role_arn`opzione deve puntare allo stesso ruolo per ogni sink all'interno di un file di definizione YAML.
- L'`hosts`opzione deve specificare un endpoint del dominio OpenSearch di servizio o un endpoint di raccolta OpenSearch Serverless. Non è possibile specificare un endpoint personalizzato per un dominio; deve essere l'endpoint standard.
- Se l'`hosts`opzione è un endpoint di raccolta senza server, è necessario impostarla su. `serverless` `true` Inoltre, se il file di definizione YAML contiene l'`index_type`opzione, questa deve essere impostata su, altrimenti la convalida `management_disabled` fallisce.
- Le seguenti opzioni non sono supportate:
  - `username`
  - `password`
  - `cert`

- `proxy`
- `dlq_file`- Se desideri scaricare gli eventi non riusciti su una coda di lettere morte (DLQ), devi utilizzare l'`dlq`opzione e specificare un bucket S3.
- `ism_policy_file`
- `socket_timeout`
- `template_file`
- `insecure`

OTel metrics source, OTel trace source e logs source OTel

I plugin OTel metrics source, OTel trace source e OTel logs source presentano i seguenti requisiti e limitazioni:

- L'*`path`*opzione è obbligatoria. Il percorso è una stringa come`/log/ingest`, che rappresenta il percorso URI per l'inserimento dei log. Questo percorso definisce l'URI utilizzato per inviare dati alla pipeline. Ad esempio, `https://log-pipeline.us-west-2.osis.amazonaws.com/log/ingest`. Il percorso deve iniziare con una barra (/) e può contenere i caratteri speciali '-', '_', ' . 'e'/', oltre al `${pipelineName}` segnaposto.
- Le seguenti opzioni sono impostate da OpenSearch Ingestion e non sono supportate nelle configurazioni della pipeline:
  - `port`
  - `ssl`
  - `sslKeyFile`
  - `sslKeyCertChainFile`
  - `authentication`
  - `unauthenticated_health_check`
  - `useAcmCertForSSL`
  - `unframed_requests`
  - `proto_reflection_service`
  - `thread_count`
  - `request_timeout`
  - `max_connection_count`
  - `acmPrivateKeyPassword`

- `acmCertIssueTimeOutMillis`
- `health_check_service`
- `acmCertificateArn`
- `awsRegion`

OTel processore trace group

Il processore del gruppo di OTel tracce presenta i seguenti requisiti e limitazioni:

- L'`aws`opzione è obbligatoria e deve contenere le seguenti opzioni:
  - `sts_role_arn`
  - `region`
  - `hosts`
- L'`sts_role_arn`opzione specifica lo stesso ruolo del ruolo della pipeline specificato nella configurazione del OpenSearch sink.
- Le `insecure` opzioni `username`,`password`,`cert`, e non sono supportate.
- L'`aws_sigv4`opzione è obbligatoria e deve essere impostata su true.
- L'`serverless`opzione all'interno del plugin OpenSearch sink non è supportata. Il processore Otel trace group attualmente non funziona con le raccolte OpenSearch Serverless.
- Il numero di `otel_trace_group` processori all'interno del corpo di configurazione della pipeline non può superare 8.

OTel processore di traccia

Il processore di OTel traccia presenta i seguenti requisiti e limitazioni:

- Il valore dell'`trace_flush_interval`opzione non può superare i 300 secondi.

Processore Service-map

Il processore Service-map presenta i seguenti requisiti e limitazioni:

- Il valore dell'`window_duration`opzione non può superare i 300 secondi.

Fonte S3

Il plugin sorgente S3 presenta i seguenti requisiti e limitazioni:

- L'`aws`opzione è obbligatoria e deve contenere `region` `sts_role_arn` opzioni.
- Il valore dell'`records_to_accumulate`opzione non può superare 200.
- Il valore dell'`maximum_messages`opzione non può superare 10.
- Se specificata, l'`disable_bucket_ownership_validation`opzione deve essere impostata su false.
- Se specificata, l'`input_serialization`opzione deve essere impostata su`parquet`.

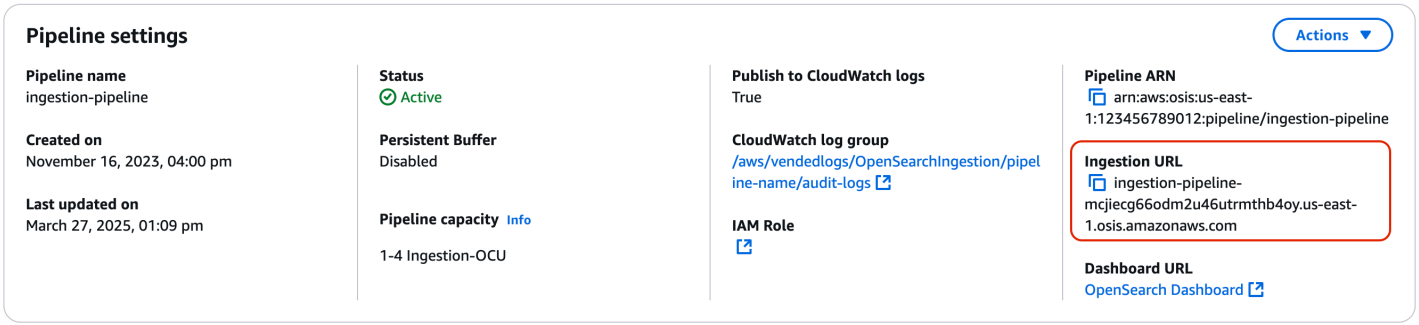
Integrazione delle pipeline OpenSearch di Amazon Ingestion con altri servizi e applicazioni

Per inserire correttamente i dati in una pipeline di Amazon OpenSearch Ingestion, devi configurare l'applicazione client (la fonte) per inviare dati all'endpoint della pipeline. La tua fonte potrebbe essere client come Fluent Bit logs, Collector o un semplice bucket S3. OpenTelemetry La configurazione esatta è diversa per ogni client.

Le differenze importanti durante la configurazione dell'origine (rispetto all'invio di dati direttamente a un dominio di OpenSearch servizio o a una raccolta OpenSearch Serverless) sono il nome del AWS servizio (`osis`) e l'endpoint host, che deve essere l'endpoint della pipeline.

Costruzione dell'endpoint di ingestione

Per importare dati in una pipeline, inviateli all'endpoint di ingestione. Per individuare l'URL di ingestione, accedete alla pagina delle impostazioni della pipeline e copiate l'URL di ingestione.



Per creare l'endpoint di ingestione completo per fonti basate su pull come OTel trace e OTel metrics, aggiungi il percorso di ingestione dalla configurazione della pipeline all'URL di ingestione.

Ad esempio, supponiamo che la configurazione della pipeline abbia il seguente percorso di inserimento:

L'endpoint di ingestione completo, specificato nella configurazione del client, assumerà il seguente formato:. `https://ingestion-pipeline-abcdefg.us-east-1.osis.amazonaws.com/my/test_path`

Creazione di un ruolo di importazione

Tutte le richieste di OpenSearch Ingestion devono essere firmate con Signature Version 4. Al ruolo che firma la richiesta deve essere almeno concessa l'autorizzazione per l'`osis:Ingest`azione, che gli consente di inviare dati a una pipeline di OpenSearch Ingestion.

Ad esempio, la seguente policy AWS Identity and Access Management (IAM) consente al ruolo corrispondente di inviare dati a una singola pipeline:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "osis:Ingest",
      "Resource": "arn:aws:osis:us-east-1:111122223333:pipeline/pipeline-name"
    }
  ]
}
```

> **Note**
>
> Per utilizzare il ruolo per tutte le pipeline, sostituite l'ARN `Resource` nell'elemento con un carattere jolly (*).

Fornire l'accesso all'importazione su più account

> **Note**
>
> Puoi fornire l'accesso all'ingestione tra account solo per le pipeline pubbliche, non per le pipeline VPC.

Potrebbe essere necessario importare dati in una pipeline da un altro Account AWS, ad esempio un account che ospita l'applicazione di origine. Se il principale che sta scrivendo su una pipeline si trova in un account diverso da quello della pipeline stessa, devi configurarlo in modo che si fidi di un altro ruolo IAM per l'inserimento dei dati nella pipeline.

Per configurare le autorizzazioni di importazione tra più account

1. Crea il ruolo di ingestione con `osis:Ingest` autorizzazione (descritto nella sezione precedente) all'interno della stessa pipeline. Account AWS Per istruzioni, consulta Creazione di ruoli IAM.
2. Allega una politica di fiducia al ruolo di inserimento che consenta a un responsabile di un altro account di assumerlo:

   JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
     "Effect": "Allow",
     "Principal": {
       "AWS": "arn:aws:iam::111122223333:root"
      },
     "Action": "sts:AssumeRole"
  }]
}
```

3. Nell'altro account, configura l'applicazione client (ad esempio, Fluent Bit) per assumere il ruolo di importazione. Affinché ciò funzioni, l'account dell'applicazione deve concedere le autorizzazioni all'utente o al ruolo dell'applicazione per assumere il ruolo di ingestione.

   L'esempio seguente di politica basata sull'identità consente al principale allegato di assumere dall'account della pipeline: `ingestion-role`

   JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "sts:AssumeRole",
      "Resource": "arn:aws:iam::111122223333:role/ingestion-role"
    }
  ]
}
```

L'applicazione client può quindi utilizzare l'AssumeRoleoperazione per assumere `ingestion-role` e importare dati nella pipeline associata.

Utilizzo di una pipeline OpenSearch di ingestione con Atlassian Services

Puoi utilizzare i plugin di origine di Atlassian Jira e Confluence per importare dati dai servizi Atlassian nella tua pipeline di Ingestion. OpenSearch Queste integrazioni consentono di creare una knowledge base unificata e ricercabile sincronizzando progetti Jira completi e spazi Confluence, mantenendo al contempo la pertinenza in tempo reale attraverso il monitoraggio continuo e la sincronizzazione automatica degli aggiornamenti.

Integrating with Jira

Trasforma la tua esperienza con Jira con potenti funzionalità di ricerca contestuale integrando i tuoi contenuti Jira in. OpenSearch Il plug-in sorgente Data Prepper Atlassian Jira ti consente di creare una knowledge base unificata e ricercabile sincronizzando progetti Jira completi, mantenendo al contempo la pertinenza in tempo reale attraverso il monitoraggio continuo e la sincronizzazione automatica degli aggiornamenti. Questa integrazione consente la

sincronizzazione dei dati con opzioni di filtro flessibili per progetti, tipi di problemi e stato specifici, garantendo che vengano importate solo le informazioni necessarie.

Per garantire una connettività sicura e affidabile, il plug-in supporta diversi metodi di autenticazione, tra cui l'autenticazione e OAuth2 l'autenticazione con chiave API di base, con la maggiore sicurezza della gestione delle credenziali utilizzando un segreto archiviato in. Gestione dei segreti AWS Dispone inoltre del rinnovo automatico dei token per un accesso ininterrotto, garantendo un funzionamento continuo. Basata sulla versione 2 dell'API di Atlassian, questa integrazione consente ai team di ottenere informazioni preziose dai dati Jira attraverso le funzionalità di ricerca avanzate di Jira. OpenSearch

Integrating with Confluence

Migliora le capacità di collaborazione e gestione delle conoscenze del tuo team integrando i contenuti di Atlassian Confluence nel plug-in sorgente Confluence di Data Prepper. OpenSearch Questa integrazione consente di creare un archivio centralizzato e ricercabile di conoscenze collettive, migliorando la scoperta delle informazioni e la produttività del team. Sincronizzando i contenuti di Confluence e monitorando continuamente gli aggiornamenti, il plugin assicura che l'indice rimanga e completo. OpenSearch up-to-date

L'integrazione offre opzioni di filtro flessibili, che consentono di importare selettivamente i contenuti da spazi o tipi di pagina specifici, adattando i contenuti sincronizzati alle esigenze dell'organizzazione. Il plugin supporta sia la chiave API di base che i metodi di OAuth2 autenticazione, con la possibilità di gestire in modo sicuro le credenziali tramite. Gestione dei segreti AWS La funzione di rinnovo automatico dei token del plug-in garantisce un accesso ininterrotto e un funzionamento senza interruzioni. Basata sull'API Confluence di Atlassian, questa integrazione consente ai team di sfruttare OpenSearch le funzionalità di ricerca avanzate di tutti i contenuti Confluence, migliorando l'accessibilità e l'utilizzo delle informazioni all'interno dell'organizzazione.

Argomenti

- Prerequisiti
- Configura un ruolo della pipeline
- Configurazione della pipeline del connettore Jira
- Configurazione della pipeline del connettore Confluence
- Coerenza dei dati
- Limitazioni

- Metriche introdotte per i connettori CloudWatch Atlassian
- Connessione di una pipeline Amazon OpenSearch Ingestion ad Atlassian Jira o Confluence tramite 2.0 OAuth

Prerequisiti

Prima di creare la pipeline di Ingestion, completa i seguenti passaggi: OpenSearch

1. Prepara le credenziali per il tuo sito Jira scegliendo una delle seguenti opzioni. OpenSearch L'ingestione richiede solo l'`ReadOnly`autorizzazione al contenuto.
   a. Opzione 1: chiave API: accedi al tuo account Atlassian e utilizza le informazioni nel seguente argomento per generare la tua chiave API:
      - Gestisci i token API per il tuo account Atlassian
   b. Opzione 2: OAuth2 — Accedi al tuo account Atlassian e usa le informazioni in esso contenute. the section called “Connessione di una pipeline Amazon OpenSearch Ingestion ad Atlassian Jira o Confluence tramite 2.0 OAuth”
2. Crea un account segreto Gestione dei segreti AWS per archiviare le credenziali create nel passaggio precedente. Effettua le seguenti scelte mentre segui la procedura:
   - Per Secret type (Tipo di segreto), scegli Other type of secret (Altro tipo di segreto).
   - Per le coppie chiave/valore, create le seguenti coppie, a seconda del tipo di autorizzazione selezionato:

   API key

   ```
   {
      "username": user-name-usualy-email-id,
      "password": api-key
   }
   ```

   OAuth 2.0

   ```
   {
      "clientId": client-id
      "clientSecret": client-secret
      "accessKey": access-key
      "refreshKey": refresh-key
   }
   ```

Dopo aver creato il segreto, copia l'Amazon Resource Name (ARN) del segreto. Lo includerai nella politica di autorizzazione dei ruoli della pipeline.

Configura un ruolo della pipeline

Il ruolo passato nella pipeline deve avere la seguente politica di lettura e scrittura allegata al segreto creato nella sezione dei prerequisiti.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "SecretReadWrite",
            "Effect": "Allow",
            "Action": [
                "secretsmanager:GetResourcePolicy",
                "secretsmanager:GetSecretValue",
                "secretsmanager:DescribeSecret",
                "secretsmanager:PutSecretValue",
                "secretsmanager:ListSecretVersionIds"
            ],
            "Resource": "arn:aws:secretsmanager:us-
east-1:111122223333:secret:secret-name-random-6-characters"
        }
    ]
}
```

Il ruolo dovrebbe inoltre avere una politica allegata per l'accesso e la scrittura nel sink prescelto. Ad esempio, se scegli OpenSearch come sink, la policy è simile alla seguente:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
```

```
            "Sid": "OpenSearchWritePolicy",
            "Effect": "Allow",
            "Action": "aoss:*",
            "Resource": "arn:aws:aoss:us-
east-1:111122223333:collection/collection-id"
        }
    ]
}
```

Configurazione della pipeline del connettore Jira

Puoi utilizzare un blueprint Atlassian Jira preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Sostituisci *placeholder values* con le informazioni appropriate.

```
version: "2"
extension:
  aws:
    secrets:
      jira-account-credentials:
        secret_id: "secret-arn"
        region: "secret-region"
        sts_role_arn: "arn:aws:iam::123456789012:role/Example-Role"
atlassian-jira-pipeline:
  source:
    jira:
      # We only support one host url for now
      hosts: ["jira-host-url"]
      acknowledgments: true
      authentication:
        # Provide one of the authentication method to use. Supported methods are
 'basic' and 'oauth2'.
        # For basic authentication, password is the API key that you generate using
 your jira account
        basic:
          username: ${{aws_secrets:jira-account-credentials:username}}
          password: ${{aws_secrets:jira-account-credentials:password}}
        # For OAuth2 based authentication, we require the following 4 key values stored
 in the secret
        # Follow atlassian instructions at the below link to generate these keys.
        # https://developer.atlassian.com/cloud/confluence/oauth-2-3lo-apps/
```

```
        # If you are using OAuth2 authentication, we also require, write permission to
 your AWS secret to
        # be able to write the renewed tokens back into the secret.
        # oauth2:
          # client_id: ${{aws_secrets:jira-account-credentials:clientId}}
          # client_secret: ${{aws_secrets:jira-account-credentials:clientSecret}}
          # access_token: ${{aws_secrets:jira-account-credentials:accessToken}}
          # refresh_token: ${{aws_secrets:jira-account-credentials:refreshToken}}
      filter:
        project:
          key:
            include:
              # This is not project name.
              # It is an alphanumeric project key that you can find under project
 details in Jira.
              - "project-key"
              - "project-key"
            # exclude:
              # - "project-key"
              # - "project-key"
        issue_type:
          include:
            - "issue-type"
            # - "Story"
            # - "Bug"
            # - "Task"
         # exclude:
             # - "Epic"
        status:
          include:
            - "ticket-status"
            # - "To Do"
            # - "In Progress"
            # - "Done"
         # exclude:
           # - "Backlog"

  sink:
    - opensearch:
        # Provide an Amazon OpenSearch Service domain endpoint
        hosts: [ "https://search-mydomain-1a2a3a4a5a6a7a8a9a0a9a8a7a.us-
east-1.es.amazonaws.com" ]
        index: "index_${getMetadata(\"project\")}"
        # Ensure adding unique document id which is the unique ticket id in this case
```

```
        document_id: '${/id}'
        aws:
          # Provide a Role ARN with access to the domain. This role should have a trust
relationship with osis-pipelines.amazonaws.com
          sts_role_arn: "arn:aws:iam::123456789012:role/Example-Role"
          # Provide the region of the domain.
          region: "us-east-1"
          # Enable the 'serverless' flag if the sink is an Amazon OpenSearch Serverless
collection
          serverless: false
          # serverless_options:
            # Specify a name here to create or update network policy for the serverless
collection
            # network_policy_name: "network-policy-name"
        # Enable the 'distribution_version' setting if the Amazon OpenSearch Service
domain is of version Elasticsearch 6.x
        # distribution_version: "es6"
        # Enable and switch the 'enable_request_compression' flag if the default
compression setting is changed in the domain.
        # See the section called “Compressione delle richieste HTTP”
        # enable_request_compression: true/false
        # Optional: Enable the S3 DLQ to capture any failed requests in an S3 bucket.
Delete this entire block if you don't want a DLQ.
        dlq:
          s3:
            # Provide an S3 bucket
            bucket: "your-dlq-bucket-name"
            # Provide a key path prefix for the failed requests
            # key_path_prefix: "kinesis-pipeline/logs/dlq"
            # Provide the region of the bucket.
            region: "us-east-1"
            # Provide a Role ARN with access to the bucket. This role should have a
trust relationship with osis-pipelines.amazonaws.com
            sts_role_arn: "arn:aws:iam::123456789012:role/Example-Role"
```

Chiave degli attributi nel codice sorgente di Jira:

1. hosts: il tuo URL Jira cloud o locale. In genere, sembra. `https://your-domain-name.atlassian.net/`
2. riconoscimenti: per garantire la consegna dei dati fino al lavandino.

3. autenticazione: descrive come desideri che la pipeline acceda alla tua istanza Jira. Scegli `Basic` o specifica `0Auth2` gli attributi chiave corrispondenti che fanno riferimento alle chiavi del tuo segreto. AWS
4. filtro: questa sezione ti aiuta a selezionare quale parte dei tuoi dati Jira estrarre e sincronizzare.
   a. progetto: Elenca le chiavi del progetto che desideri sincronizzare nella `include` sezione. Altrimenti, elenca i progetti che desideri escludere nella `exclude` sezione. Fornisci solo una delle opzioni di inclusione o esclusione alla volta.
   b. issue_type: tipi di problemi specifici che desideri sincronizzare. Segui lo `exclude` schema simile `include` o più adatto alle tue esigenze. Tieni presente che gli allegati verranno visualizzati come collegamenti di ancoraggio all'allegato originale, ma il contenuto dell'allegato non verrà estratto.
   c. status: filtro di stato specifico che desideri applicare per la query di estrazione dei dati. Se lo specifichi`include`, verranno sincronizzati solo i ticket con questi stati. Se lo specifichi`exclude`, verranno sincronizzati tutti i ticket tranne quelli con gli stati esclusi elencati.

Configurazione della pipeline del connettore Confluence

Puoi utilizzare un blueprint Atlassian Confluence preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

```
version: "2"
extension:
  aws:
    secrets:
      confluence-account-credentials:
        secret_id: "secret-arn"
        region: "secret-region"
        sts_role_arn: "arn:aws:iam::123456789012:role/Example-Role"
atlassian-confluence-pipeline:
  source:
    confluence:
      # We currently support only one host URL.
      hosts: ["confluence-host-url"]
      acknowledgments: true
      authentication:
        # Provide one of the authentication method to use. Supported methods are
 'basic' and 'oauth2'.
        # For basic authentication, password is the API key that you generate using
 your Confluence account
```

```
        basic:
          username: ${{aws_secrets:confluence-account-credentials:confluenceId}}
          password: ${{aws_secrets:confluence-account-
credentials:confluenceCredential}}
        # For OAuth2 based authentication, we require the following 4 key values stored
 in the secret
        # Follow atlassian instructions at the following link to generate these keys:
        # https://developer.atlassian.com/cloud/confluence/oauth-2-3lo-apps/
        # If you are using OAuth2 authentication, we also require write permission to
 your AWS secret to
        # be able to write the renewed tokens back into the secret.
        # oauth2:
          # client_id: ${{aws_secrets:confluence-account-credentials:clientId}}
          # client_secret: ${{aws_secrets:confluence-account-credentials:clientSecret}}
          # access_token: ${{aws_secrets:confluence-account-credentials:accessToken}}
          # refresh_token: ${{aws_secrets:confluence-account-credentials:refreshToken}}
      filter:
        space:
          key:
            include:
              # This is not space name.
              # It is a space key that you can find under space details in Confluence.
              - "space key"
              - "space key"
           # exclude:
             #  - "space key"
             #  - "space key"
        page_type:
          include:
            - "content type"
            # - "page"
            # - "blogpost"
            # - "comment"
         # exclude:
            # - "attachment"

  sink:
    - opensearch:
        # Provide an Amazon OpenSearch Service domain endpoint
        hosts: [ "https://search-mydomain-1a2a3a4a5a6a7a8a9a0a9a8a7a.us-
east-1.es.amazonaws.com" ]
         index: "index_${getMetadata(\"space\")}"
        # Ensure adding unique document id which is the unique ticket ID in this case.
        document_id: '${/id}'
```

```
        aws:
          # Provide the Amazon Resource Name (ARN) for a role with access to the
domain. This role should have a trust relationship with osis-pipelines.amazonaws.com.
          sts_role_arn: "arn:aws:iam::123456789012:role/Example-Role"
          # Provide the Region of the domain.
          region: "us-east-1"
          # Enable the 'serverless' flag if the sink is an Amazon OpenSearch Serverless
collection
          serverless: false
          # serverless_options:
            # Specify a name here to create or update network policy for the serverless
collection.
            # network_policy_name: "network-policy-name"
        # Enable the 'distribution_version' setting if the Amazon OpenSearch Service
domain is of version Elasticsearch 6.x
        # distribution_version: "es6"
        # Enable and switch the 'enable_request_compression' flag if the default
compression setting is changed in the domain.
        # For more information, see the section called “Compressione delle richieste
HTTP”.
        # enable_request_compression: true/false
        # Optional: Enable the S3 DLQ to capture any failed requests in an S3 bucket.
Delete this entire block if you don't want a DLQ.
        dlq:
          s3:
            # Provide an S3 bucket
            bucket: "your-dlq-bucket-name"
            # Provide a key path prefix for the failed requests
            # key_path_prefix: "kinesis-pipeline/logs/dlq"
            # Provide the Rregion of the bucket.
            region: "us-east-1"
            # Provide the Amazon Resource Name (ARN) for a role with access to the
bucket. This role should have a trust relationship with osis-pipelines.amazonaws.com
            sts_role_arn: "arn:aws:iam::123456789012:role/Example-Role"
```

Attributi chiave nel sorgente Confluence:

1. host: il tuo URL nel cloud o locale di Confluence. In genere, sembra `https://your-domain-name.atlassian.net/`
2. riconoscimenti: per garantire la consegna dei dati fino al lavandino.

3. autenticazione: descrive come desideri che la pipeline acceda alla tua istanza Confluence. Scegli `Basic` o `0Auth2` e specifica gli attributi chiave corrispondenti che fanno riferimento alle chiavi del tuo segreto. AWS
4. filtro: questa sezione ti aiuta a selezionare quale parte dei dati di Confluence estrarre e sincronizzare.
   a. spazio: elenca i tasti spaziatori che desideri sincronizzare nella sezione. `include` Altrimenti, elenca gli spazi che desideri escludere nella `exclude` sezione. Fornisci solo una delle opzioni di inclusione o esclusione alla volta.
   b. page_type: tipi di pagina specifici (come pagina, post del blog o allegati) che desideri sincronizzare. Segui lo `exclude` schema simile `include` o più adatto alle tue esigenze. Tieni presente che gli allegati verranno visualizzati come collegamenti di ancoraggio all'allegato originale, ma il contenuto dell'allegato non verrà estratto.

Coerenza dei dati

In base ai filtri specificati nella pipeline YAML, i progetti (o gli spazi) selezionati verranno estratti una volta e completamente sincronizzati con il sink di destinazione. Quindi il monitoraggio continuo delle modifiche acquisirà le modifiche man mano che si verificano e aggiornerà i dati nel sink. Un'eccezione è che il monitoraggio delle modifiche sincronizza solo `update` le azioni `create` e non le `delete` azioni.

Limitazioni

- Le azioni di eliminazione degli utenti non verranno sincronizzate. I dati una volta registrati nel lavandino rimarranno nel lavandino. Gli aggiornamenti sovrascriveranno il contenuto esistente con nuove modifiche se la mappatura degli ID è specificata nelle impostazioni del sink.
- Le istanze locali che utilizzano versioni precedenti del software Atlassian che non supportano quanto segue non APIs sono compatibili con questa fonte:
  - API Jira Search versione 3
    - `rest/api/3/search`
    - `rest/api/3/issue`
  - Confluenza
    - `wiki/rest/api/content/search`
    - `wiki/rest/api/content`
    - `wiki/rest/api/settings/systemInfo`

Metriche introdotte per i connettori CloudWatch Atlassian

Tipo: metriche dei connettori Jira

Origine	Metrica	Tipo di metrica
acknowledgementSetSuccesses.count	Contatore	Se i riconoscimenti sono abilitati, questa metrica fornisce il numero di ticket sincronizzati correttamente.
acknowledgementSetFailures.contare	Contatore	Se i riconoscimenti sono abilitati, questa metrica fornisce il numero di ticket che non sono stati sincronizzati.
CrawlingTime.avg	Timer	Il tempo impiegato per visualizzare tutte le nuove modifiche.
ticketFetchLatency.avg	Timer	La media di latenza dell'API Ticket Fetch.
ticketFetchLatency.max	Timer	La latenza massima dell'API di recupero dei biglietti.
Tickets richiesti.count	Contatore	Numero di richieste di recupero dei biglietti effettuate.
ticketRequestedFailed.contare	Contatore	Numero di richieste di recupero dei ticket non riuscite.
ticketRequestedSuc	Contatore	Numero di richieste di recupero dei ticket riuscite.

Origine	Metrica	Tipo di metrica
cess.coun t		
searchCal lLatency. avg	Timer	Latenza media delle chiamate dell'API di ricerca.
searchCal lLatency. max	Timer	Latenza massima delle chiamate dell'API di ricerca.
searchRes ultsFound .count	Contatore	Numero di elementi trovati in una determinata chiamata di ricerca.
searchReq uestFaile d.count	Contatore	Contano gli errori delle chiamate all'API di ricerca.
Authfailu res.count	Contatore	Numero di errori di autenticazione.

Tipo: metriche del connettore Confluence

Origine	Metrica	Tipo di metrica
acknowled gementSet Successes .count	Contatore	Se i riconoscimenti sono abilitati, questa metrica fornisce il numero di pagine sincronizzate correttamente.
acknowled gementSet Failures. contare	Contatore	Se i riconoscimenti sono abilitati, questa metrica fornisce il numero di pagine che non sono state sincronizzate.

Origine	Metrica	Tipo di metrica
CrawlingTime.avg	Timer	Il tempo impiegato per visualizzare tutte le nuove modifiche.
pageFetchLatency.avg	Timer	Latenza dell'API di recupero dei contenuti (media).
pageFetchLatency.max	Timer	Latenza dell'API di recupero dei contenuti (massima).
Pagine richieste.Count	Contatore	Numero di chiamate dell'API di recupero dei contenuti.
pageRequestFailed.count	Contatore	Numero di richieste fallite di API di recupero dei contenuti.
pageRequestedSuccess.count	Contatore	Numero di richieste riuscite di API per il recupero dei contenuti.
searchCallLatency.avg	Timer	Latenza media delle chiamate dell'API di ricerca.
searchCallLatency.max	Timer	Latenza massima delle chiamate API di ricerca
searchResultsFound.count	Contatore	Numero di elementi trovati in una determinata chiamata di ricerca.

Origine	Metrica	Tipo di metrica
searchRequestsFailed.count	Contatore	Contano gli errori delle chiamate all'API di ricerca.
Authfailures.count	Contatore	Numero di errori di autenticazione.

Connessione di una pipeline Amazon OpenSearch Ingestion ad Atlassian Jira o Confluence tramite 2.0 OAuth

Usa le informazioni in questo argomento per aiutarti a configurare e connettere una pipeline Amazon OpenSearch Ingestion a un account Jira o Confluence utilizzando l'autenticazione 2.0. OAuth Esegui questa attività quando stai completando l'the section called “Prerequisiti”utilizzo di una pipeline di OpenSearch ingestion con Atlassian Services, ma scegli di non utilizzare le credenziali della chiave API.

Argomenti

- Crea un'app di integrazione 2.0 OAuth
- Generazione e aggiornamento di un token di accesso Atlassian Developer

Crea un'app di integrazione 2.0 OAuth

Utilizza la seguente procedura per aiutarti a creare un'app di integrazione OAuth 2.0 sul sito Web di Atlassian Developer.

Per creare un'app di integrazione OAuth 2.0

1. Accedi al tuo account Atlassian Developer su https://developer.atlassian.com/console/ myapps/.
2. Scegli Create, integrazione 2.0. OAuth
3. In Nome, inserisci un nome per identificare lo scopo dell'app.
4. Seleziona la casella di controllo Accetto di essere vincolato dai termini per sviluppatori di Atlassian, quindi scegli Crea.
5. Nella barra di navigazione a sinistra, scegli Autorizzazione, quindi scegli Aggiungi.

6. Per URL di richiamata, inserisci un URL qualsiasi, ad esempio **`https://www.amazon.com`** o**`https://www.example.com`**, quindi scegli Salva modifiche.
7. Nella barra di navigazione a sinistra, scegli la pagina Autorizzazioni, quindi nella riga per Jira API, scegli Aggiungi, quindi scegli Configura. Quindi seleziona tutte le autorizzazioni di lettura di Classic Scopes (elenco riportato di seguito), quindi seleziona Salva
8. Scegli la scheda Ambiti granulari, quindi scegli Modifica ambiti per aprire la finestra di dialogo Modifica API Jira.
9. Seleziona le autorizzazioni per il plug-in sorgente che stai utilizzando:

   Jira

   ```
   read:audit-log:jira
   read:issue:jira
   read:issue-meta:jira
   read:attachment:jira
   read:comment:jira
   read:comment.property:jira
   read:field:jira
   read:field.default-value:jira
   read:field.option:jira
   read:field-configuration-scheme:jira
   read:field-configuration:jira
   read:issue-link:jira
   read:issue-link-type:jira
   read:issue-link-type:jira
   read:issue.remote-link:jira
   read:issue.property:jira
   read:resolution:jira
   read:issue-details:jira
   read:issue-type:jira
   read:issue-worklog:jira
   read:issue-field-values:jira
   read:issue.changelog:jira
   read:issue.transition:jira
   read:issue.vote:jira
   read:jira-expressions:jira
   ```

   Confluence

   ```
   read:content:confluence
   read:content-details:confluence
   ```

```
read:space-details:confluence
read:audit-log:confluence
read:page:confluence
read:blogpost:confluence
read:custom-content:confluence
read:comment:confluence
read:space:confluence
read:space.property:confluence
read:space.setting:confluence
read:content.property:confluence
read:content.metadata:confluence
read:task:confluence
read:whiteboard:confluence
read:app-data:confluence
manage:confluence-configuration
```

10. Scegli Save (Salva).

Per informazioni correlate, consulta Implementazione OAuth 2.0 (3LO) e Determinazione degli ambiti richiesti per un'operazione sul sito Web di Atlassian Developer.

Generazione e aggiornamento di un token di accesso Atlassian Developer

Utilizza la seguente procedura per aiutarti a generare e aggiornare un token di accesso Atlassian Developer sul sito web di Atlassian Developer.

Per generare e aggiornare un token di accesso Jira

1. Accedi al tuo account Atlassian Developer su myapps/. https://developer.atlassian.com/console/
2. Scegli l'app in cui hai creato. the section called “Crea un'app di integrazione 2.0 OAuth”
3. Nella barra di navigazione a sinistra, scegli Autorizzazione.
4. Copia il valore granulare dell'URL di autorizzazione dell'API Atlassian dalla parte inferiore della pagina e incollalo nell'editor di testo che preferisci.

   Il formato dell'URL è il seguente:

```
https://auth.atlassian.com/authorize?
audience=api.atlassian.com
&client_id=YOUR_CLIENT_ID
&scope=REQUESTED_SCOPE%20REQUESTED_SCOPE_TWO
&redirect_uri=https://YOUR_APP_CALLBACK_URL
```

```
&state=YOUR_USER_BOUND_VALUE
&response_type=code
&prompt=consent
```

5. Per`state=YOUR_USER_BOUND_VALUE`, modifica il valore del parametro con qualsiasi cosa tu scelga, ad esempio state=»». **`sample_text`**

   Per ulteriori informazioni, consulta A cosa serve il parametro state? sul sito Web di Atlassian Developer.

6. Nota che la `scope` sezione elenca gli ambiti granulari selezionati in un'attività precedente. Ad esempio: `scope=read%3Ajira-work%20read%3Ajira-user%20offline_access`

   `offline_access`indica che si desidera generare un. `refresh_token`

7. Apri una finestra del browser Web e inserisci l'URL di autorizzazione che hai copiato nella barra degli indirizzi della finestra del browser.

8. Quando si apre la pagina di destinazione, verifica che le informazioni siano corrette, quindi scegli Accetta per essere reindirizzato alla home page di Jira o Confluence.

9. Dopo aver caricato la home page, copia l'URL di questa pagina. Contiene il codice di autorizzazione per la tua applicazione. Questo codice viene utilizzato per generare il token di accesso. L'intera sezione successiva `code=` è il codice di autorizzazione.

10. Usa il seguente comando cURL per generare il token di accesso. Sostituisci *`placeholder values`* con le informazioni appropriate.

    > **Tip**
    >
    > Puoi anche utilizzare un servizio di terze parti come Postman.

```
curl --request POST --url 'https://auth.atlassian.com/oauth/token' \
--header 'Content-Type: application/json' \
--data '{"grant_type": "authorization_code",
"client_id": "YOUR_CLIENT_ID",
"client_secret": "YOUR_CLIENT_SECRET",
"code": "AUTHORIZATION_CODE",
"redirect_uri": "YOUR_CALLBACK_URL"}'
```

La risposta a questo comando include i valori per `access_code` e`refresh_token`.

Utilizzo di una pipeline OpenSearch di ingestione con Amazon Aurora

Puoi utilizzare una pipeline OpenSearch di importazione con Amazon Aurora per esportare dati esistenti e trasmettere le modifiche (ad esempio creazione, aggiornamento ed eliminazione) verso domini e raccolte di Amazon OpenSearch Service. La pipeline OpenSearch Ingestion incorpora l'infrastruttura Change Data Capture (CDC) per fornire un modo su larga scala e a bassa latenza per lo streaming continuo di dati da Amazon Aurora. Aurora MySQL e Aurora PostgreSQL sono supportati.

Esistono due modi per utilizzare Amazon Aurora come origine per elaborare i dati: con o senza uno snapshot iniziale completo. Uno snapshot iniziale completo è uno snapshot di tabelle specificate e questo snapshot viene esportato in Amazon S3. Da lì, una pipeline di OpenSearch ingestione lo invia a un indice in un dominio o lo partiziona in più indici di un dominio. Per mantenere i dati in Amazon Aurora e OpenSearch coerenti, la pipeline sincronizza tutti gli eventi di creazione, aggiornamento ed eliminazione nelle tabelle dei cluster Amazon Aurora con i documenti salvati nell'indice o negli indici. OpenSearch

Quando usi uno snapshot iniziale completo, la tua pipeline di OpenSearch Ingestion prima lo inserisce e poi inizia a leggere i dati dai flussi di modifiche di Amazon Aurora. Alla fine recupera e mantiene la coerenza dei dati quasi in tempo reale tra Amazon Aurora OpenSearch e.

Puoi anche utilizzare l'integrazione di OpenSearch Ingestion con Amazon Aurora per tenere traccia dei dati di modifica, acquisizione dei dati e importare tutti gli aggiornamenti in Aurora to. OpenSearch Scegli questa opzione se disponi già di uno snapshot completo di un altro meccanismo o se desideri semplicemente acquisire tutte le modifiche ai dati nel cluster Amazon Aurora.

Quando si sceglie questa opzione, è necessario configurare la registrazione binaria per Aurora MySQL o configurare la replica logica per Aurora PostgreSQL sul cluster.

Argomenti

- Aurora MySQL
- Aurora PostgreSQL

Aurora MySQL

Completa i seguenti passaggi per configurare una pipeline di OpenSearch ingestione con Amazon Aurora per Aurora MySQL.

Argomenti

- Prerequisiti MySQL di Aurora
- Fase 1: Configurare il ruolo della pipeline
- Passaggio 2: creare la pipeline
- Coerenza dei dati
- Mappatura dei tipi di dati
- Limitazioni
- CloudWatch Allarmi consigliati

Prerequisiti MySQL di Aurora

Prima di creare la pipeline di OpenSearch Ingestion, effettuate le seguenti operazioni:

1. Crea un gruppo di parametri del cluster Aurora DB personalizzato in Amazon Aurora per configurare la registrazione binaria.

   ```
   aurora_enhanced_binlog=1
   binlog_backup=0
   binlog_format=ROW
   binlog_replication_globaldb=0
   binlog_row_image=full
   binlog_row_metadata=full
   ```

   Inoltre, assicurati che il `binlog_transaction_compression` parametro non sia impostato su e che il `binlog_row_value_options` parametro non sia impostato su. `ON` `PARTIAL_JSON`
2. Seleziona o crea un cluster Aurora MySQL DB e associa il gruppo di parametri creato nel passaggio precedente al cluster DB.
3. Configura la conservazione dei log binari su 24 ore o più.
4. Configura l'autenticazione di nome utente e password sul tuo cluster Amazon Aurora utilizzando la gestione delle password con Aurora e. Gestione dei segreti AWS Puoi anche creare una username/password combinazione creando un segreto di Secrets Manager.
5. Se utilizzi la funzionalità completa di snapshot iniziale, crea un ruolo AWS KMS key e un ruolo IAM per l'esportazione dei dati da Amazon Aurora ad Amazon S3.

   Il ruolo IAM deve avere la seguente politica di autorizzazione:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "ExportPolicy",
            "Effect": "Allow",
            "Action": [
                "s3:PutObject*",
                "s3:ListBucket",
                "s3:GetObject*",
                "s3:DeleteObject*",
                "s3:GetBucketLocation"
            ],
            "Resource": [
                "arn:aws:s3:::s3-bucket-used-in-pipeline",
                "arn:aws:s3:::s3-bucket-used-in-pipeline/*"
            ]
        }
    ]
}
```

Il ruolo dovrebbe inoltre avere le seguenti relazioni di fiducia:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "export.rds.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

6. Seleziona o crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless. Per ulteriori informazioni, consulta Creazione di domini OpenSearch di servizio e Creazione di raccolte.

7. Allega una politica basata sulle risorse al tuo dominio o una politica di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal tuo cluster Amazon Aurora DB al tuo dominio o alla tua raccolta.

Fase 1: Configurare il ruolo della pipeline

Dopo aver configurato i prerequisiti della pipeline di Amazon Aurora, configura il ruolo della pipeline da utilizzare nella configurazione della pipeline. Aggiungi anche le seguenti autorizzazioni per il codice sorgente Amazon Aurora al ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
    {
    "Sid": "allowReadingFromS3Buckets",
    "Effect": "Allow",
    "Action": [
    "s3:GetObject",
    "s3:DeleteObject",
    "s3:GetBucketLocation",
    "s3:ListBucket",
    "s3:PutObject"
    ],
    "Resource": [
    "arn:aws:s3:::s3_bucket",
    "arn:aws:s3:::s3_bucket/*"
    ]
    },
    {
    "Sid": "allowNetworkInterfacesActions",
    "Effect": "Allow",
    "Action": [
    "ec2:AttachNetworkInterface",
    "ec2:CreateNetworkInterface",
    "ec2:CreateNetworkInterfacePermission",
    "ec2:DeleteNetworkInterface",
    "ec2:DeleteNetworkInterfacePermission",
    "ec2:DetachNetworkInterface",
    "ec2:DescribeNetworkInterfaces"
```

```
    ],
    "Resource": [
    "arn:aws:ec2:*:111122223333:network-interface/*",
    "arn:aws:ec2:*:111122223333:subnet/*",
    "arn:aws:ec2:*:111122223333:security-group/*"
    ]
    },
    {
    "Sid": "allowDescribeEC2",
    "Effect": "Allow",
    "Action": [
    "ec2:Describe*"
    ],
    "Resource": "*"
    },
    {
    "Sid": "allowTagCreation",
    "Effect": "Allow",
    "Action": [
    "ec2:CreateTags"
    ],
    "Resource": "arn:aws:ec2:*:111122223333:network-interface/*",
    "Condition": {
    "StringEquals": {
    "aws:RequestTag/OSISManaged": "true"
    }
    }
    },
    {
    "Sid": "AllowDescribeInstances",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBInstances"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:db:*"
    ]
    },
    {
    "Sid": "AllowDescribeClusters",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBClusters"
    ],
```

```
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:cluster:DB-id"
    ]
    },
    {
    "Sid": "AllowSnapshots",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBClusterSnapshots",
    "rds:CreateDBClusterSnapshot",
    "rds:AddTagsToResource"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:cluster:DB-id",
    "arn:aws:rds:us-east-2:111122223333:cluster-snapshot:DB-id*"
    ]
    },
    {
    "Sid": "AllowExport",
    "Effect": "Allow",
    "Action": [
    "rds:StartExportTask"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:cluster:DB-id",
    "arn:aws:rds:us-east-2:111122223333:cluster-snapshot:DB-id*"
    ]
    },
    {
    "Sid": "AllowDescribeExports",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeExportTasks"
    ],
    "Resource": "*",
    "Condition": {
    "StringEquals": {
    "aws:RequestedRegion": "us-east-2",
    "aws:ResourceAccount": "111122223333"
    }
    }
    },
    {
    "Sid": "AllowAccessToKmsForExport",
```

```
    "Effect": "Allow",
    "Action": [
    "kms:Decrypt",
    "kms:Encrypt",
    "kms:DescribeKey",
    "kms:RetireGrant",
    "kms:CreateGrant",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*"
    ],
    "Resource": [
    "arn:aws:kms:us-east-2:111122223333:key/export-key-id"
    ]
    },
    {
    "Sid": "AllowPassingExportRole",
    "Effect": "Allow",
    "Action": "iam:PassRole",
    "Resource": [
    "arn:aws:iam::111122223333:role/export-role"
    ]
    },
    {
    "Sid": "SecretsManagerReadAccess",
    "Effect": "Allow",
    "Action": [
    "secretsmanager:GetSecretValue"
    ],
    "Resource": [
    "arn:aws:secretsmanager:*:111122223333:secret:*"
    ]
    }
    ]
    }
```

Passaggio 2: creare la pipeline

Configurate una pipeline OpenSearch di ingestione simile alla seguente. La pipeline di esempio specifica un cluster Amazon Aurora come origine.

```
version: "2"
aurora-mysql-pipeline:
```

```
  source:
    rds:
      db_identifier: "cluster-id"
      engine: aurora-mysql
      database: "database-name"
      tables:
        include:
          - "table1"
          - "table2"
      s3_bucket: "bucket-name"
      s3_region: "bucket-region"
      s3_prefix: "prefix-name"
      export:
        kms_key_id: "kms-key-id"
        iam_role_arn: "export-role-arn"
      stream: true
      aws:
        sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        region: "us-east-1"
      authentication:
        username: ${{aws_secrets:secret:username}}
        password: ${{aws_secrets:secret:password}}
  sink:
    - opensearch:
        hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
        index: "${getMetadata(\"table_name\")}"
        index_type: custom
        document_id: "${getMetadata(\"primary_key\")}"
        action: "${getMetadata(\"opensearch_action\")}"
        document_version: "${getMetadata(\"document_version\")}"
        document_version_type: "external"
        aws:
          sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
          region: "us-east-1"
extension:
  aws:
    secrets:
      secret:
        secret_id: "rds-secret-id"
        region: "us-east-1"
        sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        refresh_interval: PT1H
```

Puoi usare un blueprint Amazon Aurora preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta Lavorare con i progetti.

Per utilizzare Amazon Aurora come sorgente, devi configurare l'accesso VPC per la pipeline. Il VPC scelto deve essere lo stesso VPC utilizzato dalla sorgente Amazon Aurora. Quindi scegli una o più sottoreti e uno o più gruppi di sicurezza VPC. Tieni presente che la pipeline richiede l'accesso di rete a un database Aurora MySQL, quindi dovresti anche verificare che il cluster Aurora sia configurato con un gruppo di sicurezza VPC che consenta il traffico in entrata dal gruppo di sicurezza VPC della pipeline alla porta del database. Per ulteriori informazioni, consulta Controllo dell'accesso con gruppi di sicurezza.

Se utilizzi il Console di gestione AWS per creare la tua pipeline, devi anche collegare la pipeline al tuo VPC per utilizzare Amazon Aurora come sorgente. Per farlo, trova la sezione Configurazione di rete, seleziona la casella di controllo Collega a VPC e scegli il tuo CIDR da una delle opzioni predefinite fornite oppure selezionane una personalizzata. È possibile utilizzare qualsiasi CIDR da uno spazio di indirizzi privato come definito nella RFC 1918 Best Current Practice.

Per fornire un CIDR personalizzato, seleziona Altro dal menu a discesa. Per evitare una collisione negli indirizzi IP tra OpenSearch Ingestion e Amazon Aurora, assicurati che il CIDR VPC di Amazon Aurora sia diverso dal CIDR per Ingestion. OpenSearch

Per ulteriori informazioni, consulta Configurazione dell'accesso VPC per una pipeline.

Coerenza dei dati

La pipeline garantisce la coerenza dei dati effettuando continuamente polling o ricevendo modifiche dal cluster Amazon Aurora e aggiornando i documenti corrispondenti nell'indice. OpenSearch

OpenSearch Ingestion supporta il end-to-end riconoscimento per garantire la durabilità dei dati. Quando una pipeline legge istantanee o flussi, crea dinamicamente partizioni per l'elaborazione parallela. La pipeline contrassegna una partizione come completa quando riceve un riconoscimento dopo aver acquisito tutti i record nel dominio o nella raccolta. OpenSearch Se desideri importare in una raccolta di ricerca OpenSearch Serverless, puoi generare un ID di documento nella pipeline. Se desideri importare in una raccolta di serie temporali OpenSearch Serverless, tieni presente che la pipeline non genera un ID di documento, quindi devi `document_id: "${getMetadata(\"primary_key\")}"` ometterlo nella configurazione del pipeline sink.

Una pipeline OpenSearch di Ingestion mappa anche le azioni degli eventi in entrata nelle corrispondenti azioni di indicizzazione di massa per facilitare l'acquisizione dei documenti. Ciò

mantiene i dati coerenti, in modo che ogni modifica dei dati in Amazon Aurora venga riconciliata con le corrispondenti modifiche del documento. OpenSearch

Mappatura dei tipi di dati

OpenSearch La pipeline di ingestione mappa i tipi di dati MySQL su rappresentazioni adatte all'utilizzo da parte di domini o raccolte di servizi. OpenSearch Se non è definito alcun modello di mappatura OpenSearch, determina OpenSearch automaticamente i tipi di campo con mappatura dinamica in base al primo documento inviato. È inoltre possibile definire in modo esplicito i tipi di campo più adatti alle proprie esigenze OpenSearch tramite un modello di mappatura.

La tabella seguente elenca i tipi di dati MySQL e OpenSearch i tipi di campo corrispondenti. La colonna Tipo di OpenSearch campo predefinito mostra il tipo di campo corrispondente OpenSearch se non è definita alcuna mappatura esplicita. In questo caso, determina OpenSearch automaticamente i tipi di campo con mappatura dinamica. La colonna Tipo di OpenSearch campo consigliato è il tipo di campo corrispondente che si consiglia di specificare esplicitamente in un modello di mappatura. Questi tipi di campo sono più allineati ai tipi di dati in MySQL e di solito possono abilitare migliori funzionalità di ricerca disponibili in. OpenSearch

Tipi di dati MySQL	Tipo di campo predefinito OpenSearch	Tipo di OpenSearch campo consigliato
BIGINT	Long	Long
BIGINT UNSIGNED	Long	lungo senza segno
BIT	Long	byte, corto, intero o lungo a seconda del numero di bit
DECIMAL	testo	doppio o parola chiave
DOUBLE	virgola mobile	virgola mobile a doppia precisione
FLOAT	virgola mobile	virgola mobile
INT	Long	intero
INT UNSIGNED	Long	Long

Tipi di dati MySQL	Tipo di campo predefinito OpenSearch	Tipo di OpenSearch campo consigliato
MEDIUMIN	Long	intero
MEDIUMIN UNSIGNED	Long	intero
NUMERIC	testo	doppio o parola chiave
SMALLINT	Long	short
SMALLINT UNSIGNED	Long	intero
TINYINT	Long	byte
TINYINT UNSIGNED	Long	short
BINARY	testo	binary
BLOB	testo	binary
CHAR	testo	testo
ENUM	testo	Parola chiave
LONGBLOE	testo	binary
LONGTEXT	testo	testo
MEDIUMBL B	testo	binary
MEDIUMTE T	testo	testo
SET	testo	Parola chiave
TEXT	testo	testo

Tipi di dati MySQL	Tipo di campo predefinito OpenSearch	Tipo di OpenSearch campo consigliato
TINYBLOB	testo	binary
TINYTEXT	testo	testo
VARBINAR	testo	binary
VARCHAR	testo	testo
DATE	lungo (in millisecondi di epoca)	data
DATETIME	long (in millisecondi epocali)	data
TIME	long (in millisecondi epocali)	data
TIMESTAM	long (in millisecondi epocali)	data
ANNO	long (in millisecondi epocali)	data
GEOMETR	testo (in formato WKT)	geo_shape
GEOMETR OLLECTIO	testo (in formato WKT)	geo_shape
LINESTRIN G	testo (in formato WKT)	geo_shape
MULTILINE STRING	testo (in formato WKT)	geo_shape
MULTIPOI T	testo (in formato WKT)	geo_shape
MULTIPOL GON	testo (in formato WKT)	geo_shape
POINT	testo (in formato WKT)	geo_point o geo_shape
POLYGON	testo (in formato WKT)	geo_shape

Tipi di dati MySQL	Tipo di campo predefinito OpenSearch	Tipo di OpenSearch campo consigliato
JSON	testo	oggetto

Ti consigliamo di configurare la coda di lettere morte (DLQ) nella pipeline di Ingestion. OpenSearch Se hai configurato la coda, OpenSearch Service invia tutti i documenti non riusciti che non possono essere importati a causa di errori di mappatura dinamica sulla coda.

Se le mappature automatiche falliscono, puoi utilizzare `template_type` e `template_content` nella configurazione della pipeline per definire regole di mappatura esplicite. In alternativa, puoi creare modelli di mappatura direttamente nel tuo dominio di ricerca o nella tua raccolta prima di avviare la pipeline.

Limitazioni

Considerate le seguenti limitazioni quando impostate una pipeline di OpenSearch ingestione per Aurora MySQL:

- L'integrazione supporta solo un database MySQL per pipeline.
- L'integrazione attualmente non supporta l'ingestione di dati tra regioni; il cluster e il OpenSearch dominio Amazon Aurora devono trovarsi nello stesso posto. Regione AWS
- L'integrazione attualmente non supporta l'ingestione di dati tra account; il cluster Amazon Aurora e la pipeline di importazione devono trovarsi OpenSearch nello stesso ambiente. Account AWS
- Assicurati che il cluster Amazon Aurora abbia l'autenticazione abilitata utilizzando Secrets Manager, che è l'unico meccanismo di autenticazione supportato.
- La configurazione della pipeline esistente non può essere aggiornata per inserire dati da un database and/or diverso in una tabella diversa. Per aggiornare il nome del database e/o della tabella di una pipeline, è necessario arrestare la pipeline e riavviarla con una configurazione aggiornata oppure creare una nuova pipeline.
- Le istruzioni DDL (Data Definition Language) in genere non sono supportate. La coerenza dei dati non verrà mantenuta se:
  - Le chiavi primarie vengono modificate (add/delete/rename).
  - Le tabelle vengono eliminate/troncate.
  - I nomi delle colonne o i tipi di dati vengono modificati.

- Se le tabelle MySQL da sincronizzare non hanno chiavi primarie definite, la coerenza dei dati non è garantita. Dovrai definire correttamente `document_id` l'opzione personalizzata nella configurazione del OpenSearch sink per poterla updates/deletes sincronizzare con. OpenSearch
- I riferimenti a chiavi esterne con azioni di eliminazione a cascata non sono supportati e possono causare incongruenze dei dati tra Aurora MySQL e. OpenSearch
- Versioni supportate: Aurora MySQL versione 3.05.2 e successive.

CloudWatch Allarmi consigliati

Le seguenti CloudWatch metriche sono consigliate per monitorare le prestazioni della pipeline di ingestione. Queste metriche possono aiutarti a identificare la quantità di dati elaborati dalle esportazioni, il numero di eventi elaborati dagli stream, gli errori nell'elaborazione delle esportazioni e degli eventi di flusso e il numero di documenti scritti nella destinazione. È possibile impostare CloudWatch allarmi per eseguire un'azione quando una di queste metriche supera un valore specificato per un determinato periodo di tempo.

Metrica	Description
pipeline-name .rds.Credentials Modificate	Questa metrica indica la frequenza con cui i segreti vengono ruotati. AWS
pipeline-name .rds.executorRefreshErrors	Questa metrica indica gli errori nell'aggiornamento AWS dei segreti.
pipeline-name .rds.exportRecordsTotal	Questa metrica indica il numero di record esportati da Amazon Aurora.
pipeline-name .rds.exportRecordsProcessed	Questa metrica indica il numero di record elaborati dalla pipeline di OpenSearch ingestione.

Metrica	Description
pipeline-name .rds. exportRec ordProcessingErrori	Questa metrica indica il numero di errori di elaborazione in una pipeline di OpenSearch ingestione durante la lettura dei dati da un cluster Amazon Aurora.
pipeline-name .rds. exportRec ordsSuccessTotale	Questa metrica indica il numero totale di record di esportazione elaborati correttamente.
pipeline-name .rds. exportRecordsFaile dTotale	Questa metrica indica il numero totale di record di esportazione che non sono stati elaborati.
pipeline-name .RDS.byte ricevuti	Questa metrica indica il numero totale di byte ricevuti da una pipeline di ingestione. OpenSearch
pipeline-name .rds.bytes elaborati	Questa metrica indica il numero totale di byte elaborati da una pipeline di ingestione. OpenSearch
pipeline-name .rds. streamRec ordsSuccessTotale	Questa metrica indica il numero di record elaborati con successo dallo stream.
pipeline-name .rds. streamRecordsFaile dTotale	Questa metrica indica il numero totale di record non elaborati dallo stream.

Aurora PostgreSQL

Completa i seguenti passaggi per configurare una pipeline di OpenSearch ingestione con Amazon Aurora per Aurora PostgreSQL.

Argomenti

- Prerequisiti Aurora PostgreSQL
- Fase 1: Configurare il ruolo della pipeline
- Passaggio 2: creare la pipeline
- Coerenza dei dati
- Mappatura dei tipi di dati
- Limitazioni
- CloudWatch Allarmi consigliati

Prerequisiti Aurora PostgreSQL

Prima di creare la pipeline di OpenSearch Ingestion, effettuate le seguenti operazioni:

1. Crea un gruppo di parametri del cluster DB personalizzato in Amazon Aurora per configurare la replica logica.

   ```
   rds.logical_replication=1
       aurora.enhanced_logical_replication=1
       aurora.logical_replication_backup=0
       aurora.logical_replication_globaldb=0
   ```

2. Seleziona o crea un cluster Aurora PostgreSQL DB e associa il gruppo di parametri creato nel passaggio 1 al cluster DB.
3. Configura l'autenticazione di nome utente e password sul tuo cluster Amazon Aurora utilizzando la gestione delle password con Aurora e. Gestione dei segreti AWS Puoi anche creare una username/password combinazione creando un segreto di Secrets Manager.
4. Se utilizzi la funzionalità completa di snapshot iniziale, crea un ruolo AWS KMS key e un ruolo IAM per l'esportazione dei dati da Amazon Aurora ad Amazon S3.

   Il ruolo IAM deve avere la seguente politica di autorizzazione:

JSON

```
{
        "Version":"2012-10-17",
        "Statement": [
            {
                "Sid": "ExportPolicy",
                "Effect": "Allow",
                "Action": [
                    "s3:PutObject*",
                    "s3:ListBucket",
                    "s3:GetObject*",
                    "s3:DeleteObject*",
                    "s3:GetBucketLocation"
                ],
                "Resource": [
                    "arn:aws:s3:::s3-bucket-used-in-pipeline",
                    "arn:aws:s3:::s3-bucket-used-in-pipeline/*"
                ]
            }
        ]
    }
```

Il ruolo dovrebbe inoltre avere le seguenti relazioni di fiducia:

JSON

```
{
        "Version":"2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Principal": {
                    "Service": "export.rds.amazonaws.com"
                },
                "Action": "sts:AssumeRole"
            }
        ]
    }
```

5. Seleziona o crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless. Per ulteriori informazioni, consulta Creazione di domini OpenSearch di servizio e Creazione di raccolte.

6. Allega una politica basata sulle risorse al tuo dominio o una politica di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal tuo cluster Amazon Aurora DB al tuo dominio o alla tua raccolta.

Fase 1: Configurare il ruolo della pipeline

Dopo aver configurato i prerequisiti della pipeline di Amazon Aurora, configura il ruolo della pipeline da utilizzare nella configurazione della pipeline. Aggiungi anche le seguenti autorizzazioni per il codice sorgente Amazon Aurora al ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
    {
    "Sid": "allowReadingFromS3Buckets",
    "Effect": "Allow",
    "Action": [
    "s3:GetObject",
    "s3:DeleteObject",
    "s3:GetBucketLocation",
    "s3:ListBucket",
    "s3:PutObject"
    ],
    "Resource": [
    "arn:aws:s3:::s3_bucket",
    "arn:aws:s3:::s3_bucket/*"
    ]
    },
    {
    "Sid": "allowNetworkInterfacesActions",
    "Effect": "Allow",
    "Action": [
    "ec2:AttachNetworkInterface",
    "ec2:CreateNetworkInterface",
    "ec2:CreateNetworkInterfacePermission",
    "ec2:DeleteNetworkInterface",
    "ec2:DeleteNetworkInterfacePermission",
    "ec2:DetachNetworkInterface",
    "ec2:DescribeNetworkInterfaces"
```

```
    ],
    "Resource": [
    "arn:aws:ec2:*:111122223333:network-interface/*",
    "arn:aws:ec2:*:111122223333:subnet/*",
    "arn:aws:ec2:*:111122223333:security-group/*"
    ]
    },
    {
    "Sid": "allowDescribeEC2",
    "Effect": "Allow",
    "Action": [
    "ec2:Describe*"
    ],
    "Resource": "*"
    },
    {
    "Sid": "allowTagCreation",
    "Effect": "Allow",
    "Action": [
    "ec2:CreateTags"
    ],
    "Resource": "arn:aws:ec2:*:111122223333:network-interface/*",
    "Condition": {
    "StringEquals": {
    "aws:RequestTag/OSISManaged": "true"
    }
    }
    },
    {
    "Sid": "AllowDescribeInstances",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBInstances"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:db:*"
    ]
    },
    {
    "Sid": "AllowDescribeClusters",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBClusters"
    ],
```

```
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:cluster:DB-id"
    ]
    },
    {
    "Sid": "AllowSnapshots",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBClusterSnapshots",
    "rds:CreateDBClusterSnapshot",
    "rds:AddTagsToResource"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:cluster:DB-id",
    "arn:aws:rds:us-east-2:111122223333:cluster-snapshot:DB-id*"
    ]
    },
    {
    "Sid": "AllowExport",
    "Effect": "Allow",
    "Action": [
    "rds:StartExportTask"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:cluster:DB-id",
    "arn:aws:rds:us-east-2:111122223333:cluster-snapshot:DB-id*"
    ]
    },
    {
    "Sid": "AllowDescribeExports",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeExportTasks"
    ],
    "Resource": "*",
    "Condition": {
    "StringEquals": {
    "aws:RequestedRegion": "us-east-2",
    "aws:ResourceAccount": "111122223333"
    }
    }
    },
    {
    "Sid": "AllowAccessToKmsForExport",
```

```
    "Effect": "Allow",
    "Action": [
    "kms:Decrypt",
    "kms:Encrypt",
    "kms:DescribeKey",
    "kms:RetireGrant",
    "kms:CreateGrant",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*"
    ],
    "Resource": [
    "arn:aws:kms:us-east-2:111122223333:key/export-key-id"
    ]
    },
    {
    "Sid": "AllowPassingExportRole",
    "Effect": "Allow",
    "Action": "iam:PassRole",
    "Resource": [
    "arn:aws:iam::111122223333:role/export-role"
    ]
    },
    {
    "Sid": "SecretsManagerReadAccess",
    "Effect": "Allow",
    "Action": [
    "secretsmanager:GetSecretValue"
    ],
    "Resource": [
    "arn:aws:secretsmanager:*:111122223333:secret:*"
    ]
    }
    ]
    }
```

Passaggio 2: creare la pipeline

Configura una pipeline di OpenSearch ingestione come la seguente, che specifica il cluster Aurora PostgreSQL come origine.

```
version: "2"
aurora-postgres-pipeline:
```

```
  source:
    rds:
      db_identifier: "cluster-id"
      engine: aurora-postgresql
      database: "database-name"
      tables:
        include:
          - "schema1.table1"
          - "schema2.table2"
      s3_bucket: "bucket-name"
      s3_region: "bucket-region"
      s3_prefix: "prefix-name"
      export:
        kms_key_id: "kms-key-id"
        iam_role_arn: "export-role-arn"
      stream: true
      aws:
        sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        region: "us-east-1"
      authentication:
        username: ${{aws_secrets:secret:username}}
        password: ${{aws_secrets:secret:password}}
  sink:
    - opensearch:
        hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
        index: "${getMetadata(\"table_name\")}"
        index_type: custom
        document_id: "${getMetadata(\"primary_key\")}"
        action: "${getMetadata(\"opensearch_action\")}"
        document_version: "${getMetadata(\"document_version\")}"
        document_version_type: "external"
        aws:
          sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
          region: "us-east-1"
extension:
  aws:
    secrets:
      secret:
        secret_id: "rds-secret-id"
        region: "us-east-1"
        sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        refresh_interval: PT1H
```

> **Note**
>
> Puoi usare un blueprint Amazon Aurora preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta Lavorare con i progetti.

Per utilizzare Amazon Aurora come sorgente, devi configurare l'accesso VPC per la pipeline. Il VPC scelto deve essere lo stesso VPC utilizzato dalla sorgente Amazon Aurora. Quindi scegli una o più sottoreti e uno o più gruppi di sicurezza VPC. Tieni presente che la pipeline richiede l'accesso di rete a un database Aurora MySQL, quindi dovresti anche verificare che il cluster Aurora sia configurato con un gruppo di sicurezza VPC che consenta il traffico in entrata dal gruppo di sicurezza VPC della pipeline alla porta del database. Per ulteriori informazioni, consulta Controllo dell'accesso con gruppi di sicurezza.

Se utilizzi il Console di gestione AWS per creare la tua pipeline, devi anche collegare la pipeline al tuo VPC per utilizzare Amazon Aurora come sorgente. Per fare ciò, trova la sezione Configurazione di rete, scegli Collega a VPC e scegli il tuo CIDR da una delle opzioni predefinite fornite, oppure seleziona la tua. È possibile utilizzare qualsiasi CIDR da uno spazio di indirizzi privato come definito nella Best Current Practice della RFC 1918.

Per fornire un CIDR personalizzato, seleziona Altro dal menu a discesa. Per evitare una collisione negli indirizzi IP tra OpenSearch Ingestion e Amazon Aurora, assicurati che il CIDR VPC di Amazon Aurora sia diverso dal CIDR per Ingestion. OpenSearch

Per ulteriori informazioni, consulta Configurazione dell'accesso VPC per una pipeline.

Coerenza dei dati

La pipeline garantisce la coerenza dei dati effettuando continuamente polling o ricevendo modifiche dal cluster Amazon Aurora e aggiornando i documenti corrispondenti nell'indice. OpenSearch

OpenSearch Ingestion supporta il end-to-end riconoscimento per garantire la durabilità dei dati. Quando una pipeline legge istantanee o flussi, crea dinamicamente partizioni per l'elaborazione parallela. La pipeline contrassegna una partizione come completa quando riceve un riconoscimento dopo aver acquisito tutti i record nel dominio o nella raccolta. OpenSearch Se desideri importare in una raccolta di ricerca OpenSearch Serverless, puoi generare un ID di documento nella pipeline. Se desideri importare in una raccolta di serie temporali OpenSearch Serverless, tieni presente che la pipeline non genera un ID di documento, quindi devi `document_id: "${getMetadata(\"primary_key\")}"` ometterlo nella configurazione del pipeline sink.

Una pipeline OpenSearch di Ingestion mappa anche le azioni degli eventi in entrata nelle corrispondenti azioni di indicizzazione di massa per facilitare l'acquisizione dei documenti. Ciò mantiene i dati coerenti, in modo che ogni modifica dei dati in Amazon Aurora venga riconciliata con le corrispondenti modifiche del documento. OpenSearch

Mappatura dei tipi di dati

OpenSearch La pipeline di ingestione mappa i tipi di dati Aurora PostgreSQL su rappresentazioni adatte all'uso da parte di domini o raccolte di servizi. OpenSearch Se non è definito alcun modello di mappatura OpenSearch, determina OpenSearch automaticamente i tipi di campo con una mappatura dinamica basata sul primo documento inviato. È inoltre possibile definire in modo esplicito i tipi di campo più adatti alle proprie esigenze OpenSearch tramite un modello di mappatura.

La tabella seguente elenca i tipi di dati Aurora PostgreSQL e i tipi di campo corrispondenti. OpenSearch La colonna Tipo di OpenSearch campo predefinito mostra il tipo di campo corrispondente OpenSearch se non è definita alcuna mappatura esplicita. In questo caso, determina OpenSearch automaticamente i tipi di campo con mappatura dinamica. La colonna Tipo di OpenSearch campo consigliato è il tipo di campo consigliato corrispondente da specificare esplicitamente in un modello di mappatura. Questi tipi di campo sono più allineati ai tipi di dati di Aurora PostgreSQL e in genere possono abilitare migliori funzionalità di ricerca disponibili in. OpenSearch

Tipo di dati Aurora PostgreSQL	OpenSearch Tipo di campo predefinito	Tipo di OpenSearch campo consigliato
smallint	Long	short
intero	Long	intero
bigint	Long	Long
decimal	testo	doppio o parola chiave
numerico [(p, s)]	testo	doppio o parola chiave

Tipo di dati Aurora PostgreSQL	OpenSearch Tipo di campo predefinito	Tipo di OpenSearch campo consigliato
real	virgola mobile	virgola mobile
double precision	virgola mobile	virgola mobile a doppia precisione
smallserial	Long	short
seriale	Long	intero
bigserial	Long	Long
money	oggetto	oggetto
carattere variabile (n)	testo	testo
varchar(n)	testo	testo
carattere (n)	testo	testo
char(n)	testo	testo
bpchar (n)	testo	testo
bpchar	testo	testo
testo	testo	testo
enum	testo	testo
bytea	testo	binary

Tipo di dati Aurora PostgreSQL	OpenSearch Tipo di campo predefinito	Tipo di OpenSearch campo consigliato
timestamp [(p)] [senza fuso orario]	long (in millisecondi di epoca)	data
timestamp [(p)] con fuso orario	long (in millisecondi di epoca)	data
data	long (in millisecondi epocali)	data
ora [(p)] [senza fuso orario]	long (in millisecondi di epoca)	data
ora [(p)] con fuso orario	lungo (in millisecondi di epoca)	data
intervallo [campi] [(p)]	testo (ISO8601 formato)	testo
booleano	Booleano	Booleano
point	testo (in formato WKT)	geo_shape
linea	testo (in formato WKT)	geo_shape
gamba	testo (in formato WKT)	geo_shape

Tipo di dati Aurora PostgreSQL	OpenSearch Tipo di campo predefinito	Tipo di OpenSearch campo consigliato
scatola	testo (in formato WKT)	geo_shape
path	testo (in formato WKT)	geo_shape
poligono	testo (in formato WKT)	geo_shape
cerchio	oggetto	oggetto
cidr	testo	testo
inet	testo	testo
macaddr	testo	testo
macaddr8	testo	testo
bit(n)	Long	byte, corto, intero o lungo (a seconda del numero di bit)
bit varying(n)	Long	byte, corto, intero o lungo (a seconda del numero di bit)
json	oggetto	oggetto
jsonb	oggetto	oggetto
jsonpath	testo	testo

Ti consigliamo di configurare la dead-letter queue (DLQ) nella pipeline di Ingestion. OpenSearch Se hai configurato la coda, OpenSearch Service invia tutti i documenti non riusciti che non possono essere importati a causa di errori di mappatura dinamica sulla coda.

Nel caso in cui le mappature automatiche falliscano, puoi utilizzare `template_type` e `template_content` nella configurazione della pipeline per definire regole di mappatura esplicite.

In alternativa, puoi creare modelli di mappatura direttamente nel tuo dominio di ricerca o nella tua raccolta prima di avviare la pipeline.

Limitazioni

Considerate le seguenti limitazioni quando impostate una pipeline di OpenSearch ingestione per Aurora PostgreSQL:

- L'integrazione supporta solo un database Aurora PostgreSQL per pipeline.
- L'integrazione attualmente non supporta l'ingestione di dati tra regioni; il cluster e il OpenSearch dominio Amazon Aurora devono trovarsi nello stesso posto. Regione AWS
- L'integrazione attualmente non supporta l'ingestione di dati tra account; il cluster Amazon Aurora e la pipeline di importazione devono trovarsi OpenSearch nello stesso ambiente. Account AWS
- Assicurati che nel cluster Amazon Aurora sia abilitata l'autenticazione Gestione dei segreti AWS, che è l'unico meccanismo di autenticazione supportato.
- La configurazione della pipeline esistente non può essere aggiornata per inserire dati da un database and/or diverso in una tabella diversa. Per aggiornare il nome del database e/o della tabella di una pipeline, è necessario arrestare la pipeline e riavviarla con una configurazione aggiornata oppure creare una nuova pipeline.
- Le istruzioni DDL (Data Definition Language) in genere non sono supportate. La coerenza dei dati non verrà mantenuta se:
  - Le chiavi primarie vengono modificate (add/delete/rename).
  - Le tabelle vengono eliminate/troncate.
  - I nomi delle colonne o i tipi di dati vengono modificati.
- Se le tabelle Aurora PostgreSQL da sincronizzare non hanno chiavi primarie definite, la coerenza dei dati non è garantita. Dovrai definire, personalizzare l'`document_id`opzione OpenSearch e configurare correttamente il sink per poterla sincronizzare con. updates/deletes OpenSearch
- Versioni supportate: Aurora PostgreSQL versione 16.4 e successive.

CloudWatch Allarmi consigliati

Le seguenti CloudWatch metriche sono consigliate per monitorare le prestazioni della pipeline di ingestione. Queste metriche possono aiutarti a identificare la quantità di dati elaborati dalle esportazioni, il numero di eventi elaborati dagli stream, gli errori nell'elaborazione delle esportazioni e degli eventi di flusso e il numero di documenti scritti nella destinazione. È possibile impostare

CloudWatch allarmi per eseguire un'azione quando una di queste metriche supera un valore specificato per un determinato periodo di tempo.

Metrica	Description
pipeline-name .rds.Credentials Modificate	Questa metrica indica la frequenza con cui i segreti vengono ruotati. AWS
pipeline-name .rds.executorRefreshErrors	Questa metrica indica gli errori nell'aggiornamento AWS dei segreti.
pipeline-name .rds.exportRecordsTotal	Questa metrica indica il numero di record esportati da Amazon Aurora.
pipeline-name .rds.exportRecordsProcessed	Questa metrica indica il numero di record elaborati dalla pipeline di OpenSearch ingestione.
pipeline-name .rds.exportRecordProcessingErrori	Questa metrica indica il numero di errori di elaborazione in una pipeline di OpenSearch ingestione durante la lettura dei dati da un cluster Amazon Aurora.
pipeline-name .rds.exportRecordsSuccessTotale	Questa metrica indica il numero totale di record di esportazione elaborati correttamente.
pipeline-name .rds.exportRecordsFailedTotale	Questa metrica indica il numero totale di record di esportazione che non sono stati elaborati.

Metrica	Description
pipeline-name .RDS.byte ricevuti	Questa metrica indica il numero totale di byte ricevuti da una pipeline di ingestione. OpenSearch
pipeline-name .rds.bytes elaborati	Questa metrica indica il numero totale di byte elaborati da una pipeline di ingestione. OpenSearch
pipeline-name .rds. streamRec ordsSuccessTotale	Questa metrica indica il numero di record elaborati con successo dallo stream.
pipeline-name .rds. streamRecordsFaile dTotale	Questa metrica indica il numero totale di record non elaborati dallo stream.

Utilizzo di una pipeline OpenSearch di ingestione con Amazon DynamoDB

Puoi utilizzare il plug-in DynamoDB per trasmettere eventi di tabella, come creazioni, aggiornamenti ed eliminazioni, ai domini OpenSearch Amazon Service e alle raccolte Amazon Serverless. OpenSearch La pipeline utilizza Change Data Capture (CDC) per lo streaming su larga scala e a bassa latenza.

È possibile elaborare i dati DynamoDB con o senza un'istantanea iniziale completa.

- Con un'istantanea completa: DynamoDB point-in-time utilizza il ripristino (PITR) per creare un backup e caricarlo su Amazon S3. OpenSearch L'ingestione indicizza quindi lo snapshot in uno o più indici. OpenSearch Per mantenere la coerenza, la pipeline sincronizza tutte le modifiche a DynamoDB con. OpenSearch Questa opzione richiede l'attivazione sia di PITR che di DynamoDB Streams.
- Senza un'istantanea: OpenSearch Ingestion trasmette solo i nuovi eventi DynamoDB. Scegli questa opzione se disponi già di un'istantanea o hai bisogno di uno streaming in tempo reale senza dati storici. Questa opzione richiede l'attivazione solo di DynamoDB Streams.

Per ulteriori informazioni, consulta l'integrazione Zero-ETL di DynamoDB con OpenSearch Amazon Service nella Developer Guide.Amazon DynamoDB

Argomenti

- Prerequisiti
- Fase 1: Configurare il ruolo della pipeline
- Fase 2: Creare la pipeline
- Coerenza dei dati
- Mappatura dei tipi di dati
- Limitazioni
- CloudWatch Allarmi consigliati per DynamoDB

Prerequisiti

Per configurare la pipeline, è necessario disporre di una tabella DynamoDB con DynamoDB Streams abilitato. Il tuo stream deve utilizzare il tipo di visualizzazione dello stream. `NEW_IMAGE` Tuttavia, le pipeline di OpenSearch Ingestion possono anche trasmettere eventi `NEW_AND_OLD_IMAGES` se questo tipo di visualizzazione dello stream si adatta al tuo caso d'uso.

Se utilizzi le istantanee, devi anche abilitare il point-in-time ripristino sulla tua tabella. Per ulteriori informazioni, consulta Creating a table, Enabling point-in-time recovery e Enabling a stream nella Amazon DynamoDB Developer Guide.

Fase 1: Configurare il ruolo della pipeline

Dopo aver impostato la tabella DynamoDB, imposta il ruolo pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi le seguenti autorizzazioni DynamoDB nel ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "allowRunExportJob",
            "Effect": "Allow",
            "Action": [
                "dynamodb:DescribeTable",
```

```
                "dynamodb:DescribeContinuousBackups",
                "dynamodb:ExportTableToPointInTime"
            ],
            "Resource": [
                "arn:aws:dynamodb:us-east-1:111122223333:table/my-table"
            ]
        },
        {
            "Sid": "allowCheckExportjob",
            "Effect": "Allow",
            "Action": [
                "dynamodb:DescribeExport"
            ],
            "Resource": [
                "arn:aws:dynamodb:us-east-1:111122223333:table/my-table/export/*"
            ]
        },
        {
            "Sid": "allowReadFromStream",
            "Effect": "Allow",
            "Action": [
                "dynamodb:DescribeStream",
                "dynamodb:GetRecords",
                "dynamodb:GetShardIterator"
            ],
            "Resource": [
                "arn:aws:dynamodb:us-east-1:111122223333:table/my-table/stream/*"
            ]
        },
        {
            "Sid": "allowReadAndWriteToS3ForExport",
            "Effect": "Allow",
            "Action": [
                "s3:GetObject",
                "s3:AbortMultipartUpload",
                "s3:PutObject",
                "s3:PutObjectAcl"
            ],
            "Resource": [
                "arn:aws:s3:::amzn-s3-demo-bucket/export-folder/*"
            ]
        }
    ]
```

```
}
```

Puoi anche utilizzare una chiave gestita dal AWS KMS cliente per crittografare i file di dati di esportazione. Per decrittografare gli oggetti esportati, specificate `s3_sse_kms_key_id` l'ID della chiave nella configurazione di esportazione della pipeline con il seguente formato:. `arn:aws:kms:region:account-id:key/my-key-id` La seguente politica include le autorizzazioni richieste per l'utilizzo di una chiave gestita dal cliente:

```
{
    "Sid": "allowUseOfCustomManagedKey",
    "Effect": "Allow",
    "Action": [
        "kms:GenerateDataKey",
        "kms:Decrypt"
    ],
    "Resource": arn:aws:kms:region:account-id:key/my-key-id
}
```

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestione come la seguente, che specifica DynamoDB come origine. Questa pipeline di esempio inserisce i dati `table-a` con lo snapshot PITR, seguiti dagli eventi di DynamoDB Streams. Una posizione iniziale di `LATEST` indica che la pipeline deve leggere i dati più recenti da DynamoDB Streams.

```
version: "2"
cdc-pipeline:
  source:
    dynamodb:
      tables:
      - table_arn: "arn:aws:dynamodb:region:account-id:table/table-a"
        export:
          s3_bucket: "my-bucket"
          s3_prefix: "export/"
        stream:
          start_position: "LATEST"
      aws:
        region: "us-east-1"
  sink:
  - opensearch:
```

```
      hosts: ["https://search-mydomain.region.es.amazonaws.com"]
      index: "${getMetadata(\"table-name\")}"
      index_type: custom
      normalize_index: true
      document_id: "${getMetadata(\"primary_key\")}"
      action: "${getMetadata(\"opensearch_action\")}"
      document_version: "${getMetadata(\"document_version\")}"
      document_version_type: "external"
```

È possibile utilizzare un blueprint DynamoDB preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Coerenza dei dati

OpenSearch Ingestion supporta il riconoscimento per garantire la durabilità dei dati. end-to-end Quando una pipeline legge istantanee o flussi, crea dinamicamente partizioni per l'elaborazione parallela. La pipeline contrassegna una partizione come completa quando riceve un riconoscimento dopo aver acquisito tutti i record nel dominio o nella raccolta. OpenSearch

Se desideri importare in una raccolta di ricerca OpenSearch Serverless, puoi generare un ID di documento nella pipeline. Se desideri importare in una raccolta di serie temporali OpenSearch Serverless, tieni presente che la pipeline non genera un ID di documento.

Una pipeline OpenSearch di Ingestion mappa anche le azioni degli eventi in entrata nelle corrispondenti azioni di indicizzazione di massa per facilitare l'acquisizione dei documenti. Ciò mantiene i dati coerenti, in modo che ogni modifica dei dati in DynamoDB venga riconciliata con le corrispondenti modifiche del documento. OpenSearch

Mappatura dei tipi di dati

OpenSearch Il servizio mappa dinamicamente i tipi di dati in ogni documento in entrata al tipo di dati corrispondente in DynamoDB. La tabella seguente mostra come OpenSearch Service mappa automaticamente vari tipi di dati.

Tipo di dati	OpenSearch	DynamoDB
Numero	OpenSearch mappa automaticamente i dati numerici. Se il numero è un numero intero, lo OpenSearch mappa come	DynamoDB supporta i numeri.

Tipo di dati	OpenSearch	DynamoDB
	valore lungo. Se il numero è frazionario, lo OpenSearch mappa come valore float. OpenSearch mappa dinamicamente vari attributi in base al primo documento inviato. Se si dispone di una combinazi one di tipi di dati per lo stesso attributo in DynamoDB, ad esempio un numero intero e un numero frazionario, la mappatura potrebbe non riuscire. Ad esempio, se il primo documento ha un attributo che è un numero intero e un documento successivo ha lo stesso attributo come numero frazionario, OpenSearch non riesce a importare il secondo documento. In questi casi, è necessario fornire un modello di mappatura esplicito, come il seguente: `{ "template": { "mappings": { "properties": { "MixedNumberAttribute": { "type": "float" } } } } }` Se hai bisogno di una doppia precision e, usa la mappatura dei campi di tipo stringa. Non esiste un tipo numerico	

Tipo di dati	OpenSearch	DynamoDB
	equivalente che supporti 38 cifre di precisione in. OpenSearch	
Set di numeri	OpenSearch mappa automaticamente un set di numeri in una matrice di valori lunghi o a virgola mobile. Come per i numeri scalari, ciò dipende dal fatto che il primo numero ingerito sia un numero intero o un numero frazionario. È possibile fornire mappature per i set di numeri nello stesso modo in cui si mappano le stringhe scalari.	DynamoDB supporta tipi che rappresentano insiemi di numeri.
Stringa	OpenSearch mappa automaticamente i valori delle stringhe come testo. In alcune situazioni, come i valori enumerati, è possibile eseguire il mapping al tipo di parola chiave. L'esempio seguente mostra come mappare un attributo DynamoDB `PartType` denominato a una parola chiave. OpenSearch `{ "template": { "mappings": { "properties": { "PartType": { "type": "keyword" } } } } }`	DynamoDB supporta le stringhe.

Tipo di dati	OpenSearch	DynamoDB
Set di stringhe	OpenSearch mappa automaticamente un set di stringhe in un array di stringhe. È possibile fornire mappature per i set di stringhe nello stesso modo in cui si mappano le stringhe scalari.	DynamoDB supporta tipi che rappresentano set di stringhe.
Binario	OpenSearch mappa automaticamente i dati binari come testo. È possibile fornire una mappatura in cui scriverli come campi binari. OpenSearch L'esempio seguente mostra come mappare un attributo DynamoDB `ImageData` denominato su un campo binario. OpenSearch `{ "template": { "mappings": { "properties": { "ImageData": { "type": "binary" } } } } }`	DynamoDB supporta gli attributi di tipo binario.
Set binario	OpenSearch mappa automaticamente un set binario in una matrice di dati binari come testo. È possibile fornire mappature per i set di numeri nello stesso modo in cui si esegue la mappatura di un sistema binario scalare.	DynamoDB supporta tipi che rappresentano insiemi di valori binari.

Tipo di dati	OpenSearch	DynamoDB
Booleano	OpenSearch mappa un tipo booleano DynamoDB in un tipo booleano. OpenSearch	DynamoDB supporta gli attributi di tipo booleano.
Null	OpenSearch può importare documenti con il tipo null DynamoDB. Salva il valore come valore nullo nel documento. Non esiste alcuna mappatura per questo tipo e questo campo non è indicizzato o ricercabile. Se lo stesso nome di attributo viene utilizzato per un tipo nullo e successiv amente passa a un tipo diverso, ad esempio string, OpenSearch crea una mappatura dinamica per il primo valore non nullo. I valori successivi possono ancora essere valori nulli di DynamoDB.	DynamoDB supporta gli attributi di tipo null.

Tipo di dati	OpenSearch	DynamoDB
Eseguire la mappatura	OpenSearch mappa gli attributi della mappa di DynamoDB ai campi annidati. Le stesse mappature si applicano all'interno di un campo nidificato. L'esempio seguente mappa una stringa in un campo nidificato a un tipo di parola chiave in: OpenSearch `{ "template": { "mappings": { "properties": { "AdditionalDescriptions": { "properties": { "PartType": { "type": "keyword" } } } } } } }`	DynamoDB supporta gli attributi dei tipi di mappa.

Tipo di dati	OpenSearch	DynamoDB
List	OpenSearch fornisce risultati diversi per gli elenchi DynamoDB, a seconda del contenuto dell'elenco. Quando un elenco contiene tutti i tipi scalari dello stesso tipo (ad esempio, un elenco di tutte le stringhe), OpenSearch inserisce l'elenco come un array di quel tipo. Funziona per i tipi stringa, numerica, booleana e null. Le restrizioni per ciascuno di questi tipi sono le stesse delle restrizioni per uno scalare di quel tipo. È inoltre possibile fornire mappature per elenchi di mappe utilizzando la stessa mappatura utilizzata per una mappa. Non puoi fornire un elenco di tipi misti.	DynamoDB supporta gli attributi del tipo di elenco.

Tipo di dati	OpenSearch	DynamoDB
Imposta	OpenSearch fornisce risultati diversi per i set DynamoDB a seconda del contenuto del set. Quando un set contiene tutti i tipi scalari dello stesso tipo (ad esempio, un insieme di tutte le stringhe), OpenSearch inserisce il set come un array di quel tipo. Funziona per i tipi stringa, numerica, booleana e null. Le restrizioni per ciascuno di questi tipi sono le stesse delle restrizioni per uno scalare di quel tipo. È inoltre possibile fornire mappature per set di mappe utilizzando la stessa mappatura utilizzata per una mappa. Non puoi fornire un set di tipi misti.	DynamoDB supporta i tipi che rappresentano i set.

Ti consigliamo di configurare la dead-letter queue (DLQ) nella pipeline di Ingestion. OpenSearch Se hai configurato la coda, OpenSearch Service invia tutti i documenti non riusciti che non possono essere importati a causa di errori di mappatura dinamica sulla coda.

Nel caso in cui le mappature automatiche falliscano, puoi utilizzare `template_type` e `template_content` nella configurazione della pipeline per definire regole di mappatura esplicite. In alternativa, puoi creare modelli di mappatura direttamente nel tuo dominio di ricerca o nella tua raccolta prima di avviare la pipeline.

Limitazioni

Considerate le seguenti limitazioni quando impostate una pipeline di OpenSearch ingestione per DynamoDB:

- L'integrazione OpenSearch di Ingestion con DynamoDB attualmente non supporta l'ingestione tra regioni. La tabella DynamoDB OpenSearch e la pipeline di ingestione devono trovarsi nella stessa posizione. Regione AWS
- La tabella DynamoDB OpenSearch e la pipeline di ingestione devono trovarsi nella stessa posizione. Account AWS
- Una pipeline OpenSearch di ingestione supporta solo una tabella DynamoDB come origine.
- DynamoDB Streams archivia i dati in un registro solo per un massimo di 24 ore. Se l'inserimento da un'istantanea iniziale di una tabella di grandi dimensioni richiede 24 ore o più, si verificherà una perdita iniziale di dati. Per mitigare questa perdita di dati, stimate la dimensione della tabella e configurate le unità di calcolo appropriate delle pipeline di Ingestion. OpenSearch

CloudWatch Allarmi consigliati per DynamoDB

Le seguenti CloudWatch metriche sono consigliate per monitorare le prestazioni della pipeline di ingestione. Queste metriche possono aiutarti a identificare la quantità di dati elaborati dalle esportazioni, la quantità di eventi elaborati dagli stream, gli errori nell'elaborazione delle esportazioni e degli eventi di streaming e il numero di documenti scritti nella destinazione. È possibile impostare CloudWatch allarmi per eseguire un'azione quando una di queste metriche supera un valore specificato per un determinato periodo di tempo.

Metrica	Description
`dynamodb-pipeline.BlockingBuffer.bufferUsage.value`	Indica la quantità di buffer utilizzata.
`dynamodb-pipeline.dynamodb.activeExportS3ObjectConsumers.value`	Mostra il numero totale di oggetti Amazon S3 OCUs che stanno elaborando attivamente per l'esportazione.
`dynamodb-pipeline.dynamodb.bytesProcessed.count`	Numero di byte elaborati dal sorgente DynamoDB.
`dynamodb-pipeline.dynamodb.changeEventsProcessed.count`	Numero di eventi di modifica elaborati dal flusso DynamoDB.

Metrica	Description
dynamodb-pipeline.dynamodb.changeEventsProcessingErrors.count	Numero di errori derivanti da eventi di modifica elaborati da DynamoDB.
dynamodb-pipeline.dynamodb.exportJobFailure.count	Numero di tentativi di invio di lavori di esportazione non riusciti.
dynamodb-pipeline.dynamodb.exportJobSuccess.count	Numero di lavori di esportazione che sono stati inviati con successo.
dynamodb-pipeline.dynamodb.exportRecordsProcessed.count	Numero totale di record elaborati dall'esportazione.
dynamodb-pipeline.dynamodb.exportRecordsTotal.count	Numero totale di record esportati da DynamoDB, essenziale per tenere traccia dei volumi di esportazione dei dati.
dynamodb-pipeline.dynamodb.exportS3ObjectsProcessed.count	Numero totale di file di dati di esportazione che sono stati elaborati correttamente da Amazon S3.
dynamodb-pipeline.opensearch.bulkBadRequestErrors.count	Numero di errori durante le richieste in blocco a causa di una richiesta non valida.
dynamodb-pipeline.opensearch.bulkRequestLatency.avg	Latenza media per le richieste di scrittura in blocco effettuate a. OpenSearch
dynamodb-pipeline.opensearch.bulkRequestNotFoundErrors.count	Numero di richieste in blocco non riuscite perché non è stato possibile trovare i dati di destinazione.
dynamodb-pipeline.opensearch.bulkRequestNumberOfRetries.count	Numero di tentativi di scrittura del cluster da parte delle pipeline di OpenSearch Ingestion. OpenSearch
dynamodb-pipeline.opensearch.bulkRequestSizeBytes.sum	Dimensione totale in byte di tutte le richieste in blocco effettuate a. OpenSearch

Metrica	Description
dynamodb-pipeline.opensearch.documentErrors.count	Numero di errori durante l'invio di documenti a. OpenSearch I documenti che causano gli errori verranno inviati a DLQ.
dynamodb-pipeline.opensearch.documentsSuccess.count	Numero di documenti scritti correttamente in un OpenSearch cluster o in una raccolta.
dynamodb-pipeline.opensearch.documentsSuccessFirstAttempt.count	Numero di documenti indicizzati correttamente OpenSearch al primo tentativo.
dynamodb-pipeline.opensearch.documentsVersionConflictErrors.count	Numero di errori dovuti a conflitti di versione nei documenti durante l'elaborazione.
dynamodb-pipeline.opensearch.PipelineLatency.avg	Latenza media della pipeline di OpenSearch Ingestion per l'elaborazione dei dati dalla lettura dall'origine alla scrittura fino alla destinazione.
dynamodb-pipeline.opensearch.PipelineLatency.max	Latenza massima della pipeline di OpenSearch Ingestion per elaborare i dati leggendo dall'origine alla scrittura della destinazione.
dynamodb-pipeline.opensearch.recordsIn.count	Numero di record in cui sono stati inseriti con successo. OpenSearch Questa metrica è essenziale per tenere traccia del volume di dati elaborati e archiviati.
dynamodb-pipeline.opensearch.s3.dlqS3RecordsFailed.count	Numero di record che non sono stati scritti su DLQ.
dynamodb-pipeline.opensearch.s3.dlqS3RecordsSuccess.count	Numero di record scritti su DLQ.
dynamodb-pipeline.opensearch.s3.dlqS3RequestLatency.count	Conteggio delle misurazioni della latenza per le richieste alla coda di lettere morte di Amazon S3.

Metrica	Description
`dynamodb-pipeline.opensearch.s3.dlqS3RequestLatency.sum`	Latenza totale per tutte le richieste alla coda di lettera morta di Amazon S3
`dynamodb-pipeline.opensearch.s3.dlqS3RequestSizeBytes.sum`	Dimensione totale in byte di tutte le richieste effettuate alla coda di lettere morte di Amazon S3.
`dynamodb-pipeline.recordsProcessed.count`	Numero totale di record elaborati nella pipeline, una metrica chiave per il throughput complessivo.
`dynamodb.changeEventsProcessed.count`	Nessun record viene raccolto dai flussi DynamoDB. Ciò potrebbe essere dovuto all'assenza di attività sulla tabella, all'esportazione in corso o a un problema di accesso ai flussi DynamoDB.
`dynamodb.exportJobFailure.count`	Il tentativo di attivare un'esportazione in S3 non è riuscito.
`dynamodb-pipeline.opensearch.bulkRequestInvalidInputErrors.count`	Numero di errori nelle richieste di massa OpenSearch dovuti a input non validi, fondamentale per il monitoraggio della qualità dei dati e i problemi operativi.
`opensearch.EndToEndLatency.avg`	La latenza end-to-end è superiore a quella desiderata per la lettura dai flussi DynamoDB. Ciò potrebbe essere dovuto a un OpenSearch cluster sottodimensionato o a una capacità OCU massima della pipeline troppo bassa per il throughput WCU sulla tabella DynamoDB. Questa latenza end-to-end sarà elevata dopo un'esportazione e dovrebbe diminuire nel tempo man mano che si aggiornerà ai flussi DynamoDB più recenti.

Utilizzo di una pipeline OpenSearch di ingestione con Amazon DocumentDB

Puoi utilizzare il plug-in DocumentDB per trasmettere le modifiche ai documenti, come creazioni, aggiornamenti ed eliminazioni, ad Amazon Service. OpenSearch La pipeline supporta l'acquisizione dei dati di modifica (CDC), se disponibile, o il polling delle API per lo streaming su larga scala e a bassa latenza.

È possibile elaborare i dati con o senza un'istantanea iniziale completa. Uno snapshot completo acquisisce un'intera raccolta Amazon DocumentDB e la carica su Amazon S3. La pipeline invia quindi i dati a uno o più indici. OpenSearch Dopo aver acquisito l'istantanea, la pipeline sincronizza le modifiche in corso per mantenere la coerenza e alla fine recupera gli aggiornamenti quasi in tempo reale.

Se disponi già di un'istantanea completa da un'altra fonte o devi solo elaborare nuovi eventi, puoi eseguire lo streaming senza un'istantanea. In questo caso, la pipeline legge direttamente i flussi di modifiche di Amazon DocumentDB senza un carico iniziale di massa.

Se abiliti lo streaming, devi abilitare un flusso di modifiche sulla tua raccolta Amazon DocumentDB. Tuttavia, se esegui solo un caricamento o un'esportazione completi, non è necessario un flusso di modifiche.

Prerequisiti

Prima di creare la pipeline OpenSearch di Ingestion, effettuate le seguenti operazioni:

1. Crea un cluster Amazon DocumentDB con l'autorizzazione a leggere i dati seguendo i passaggi descritti in Creare un cluster Amazon DocumentDB nella Amazon DocumentDB Developer Guide. Se utilizzi l'infrastruttura CDC, configura il cluster Amazon DocumentDB per pubblicare flussi di modifiche.
2. Abilitare TLS sul cluster Amazon DocumentDB.
3. Configura un CIDR VPC di uno spazio di indirizzi privato da utilizzare con Ingestion. OpenSearch
4. Configura l'autenticazione sul tuo cluster Amazon DocumentDB con. Gestione dei segreti AWS Abilita la rotazione dei segreti seguendo i passaggi descritti in Rotazione automatica delle password per Amazon DocumentDB. Per ulteriori informazioni, consulta Accesso al database utilizzando il controllo e la sicurezza degli accessi basati sui ruoli in Amazon DocumentDB.
5. Se utilizzi un flusso di modifiche per sottoscrivere le modifiche ai dati sulla tua raccolta Amazon DocumentDB, evita la perdita di dati estendendo il periodo di conservazione fino a 7 giorni

utilizzando il parametro. `change_stream_log_retention_duration` Gli eventi Change Streams vengono archiviati per impostazione predefinita per 3 ore dopo la registrazione dell'evento, tempo non sufficiente per raccolte di grandi dimensioni. Per modificare il periodo di conservazione del log del flusso di modifica, consulta Modifica della durata di conservazione del registro del flusso di modifica.

6. Crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless. Per ulteriori informazioni, consultare the section called “Creazione OpenSearch di domini di servizio” e the section called “Creazione di raccolte”.
7. Allega una policy basata sulle risorse al tuo dominio o una policy di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal tuo cluster Amazon DocumentDB al tuo dominio o alla tua raccolta.

   Il seguente esempio di policy di accesso al dominio consente al ruolo pipeline, creato nel passaggio successivo, di scrivere dati su un dominio. Assicurati di aggiornarlo `resource` con il tuo ARN.

   JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::444455556666:role/pipeline-role"
      },
      "Action": [
        "es:DescribeDomain",
        "es:ESHttp*"
      ],
      "Resource": [
        "arn:aws:es:us-east-1:111122223333:domain/domain-name"
      ]
    }
  ]
}
```

   Per creare un ruolo IAM con le autorizzazioni corrette per accedere ai dati di scrittura nella raccolta o nel dominio, consulta. the section called “Configurazione di ruoli e utenti”

Fase 1: Configurare il ruolo della pipeline

Dopo aver impostato i prerequisiti della pipeline di Amazon DocumentDB, configura il ruolo pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi le seguenti autorizzazioni Amazon DocumentDB nel ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "allowS3ListObjectAccess",
            "Effect": "Allow",
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::s3-bucket"
            ],
            "Condition": {
                "StringLike": {
                    "s3:prefix": "s3-prefix/*"
                }
            }
        },
        {
            "Sid": "allowReadAndWriteToS3ForExportStream",
            "Effect": "Allow",
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
                "s3:DeleteObject"
            ],
            "Resource": [
                "arn:aws:s3:::s3-bucket/s3-prefix/*"
            ]
        },
        {
            "Sid": "SecretsManagerReadAccess",
            "Effect": "Allow",
            "Action": [
```

```
                "secretsmanager:GetSecretValue"
            ],
            "Resource": [
                "arn:aws:secretsmanager:us-east-1:111122223333:secret:secret-
name"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:AttachNetworkInterface",
                "ec2:CreateNetworkInterface",
                "ec2:CreateNetworkInterfacePermission",
                "ec2:DeleteNetworkInterface",
                "ec2:DeleteNetworkInterfacePermission",
                "ec2:DetachNetworkInterface",
                "ec2:DescribeNetworkInterfaces"
            ],
            "Resource": [
                "arn:aws:ec2:*:111122223333:network-interface/*",
                "arn:aws:ec2:*:111122223333:subnet/*",
                "arn:aws:ec2:*:111122223333:security-group/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeDhcpOptions",
                "ec2:DescribeRouteTables",
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeSubnets",
                "ec2:DescribeVpcs",
                "ec2:Describe*"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:CreateTags"
            ],
            "Resource": "arn:aws:ec2:*:*:network-interface/*",
            "Condition": {
                "StringEquals": {
```

```
                    "aws:RequestTag/OSISManaged": "true"
                }
            }
        }
    ]
}
```

Devi fornire le autorizzazioni Amazon EC2 di cui sopra sul ruolo IAM che usi per creare la pipeline di OpenSearch ingestione, poiché la pipeline utilizza queste autorizzazioni per creare ed eliminare un'interfaccia di rete nel tuo VPC. La pipeline può accedere al cluster Amazon DocumentDB solo tramite questa interfaccia di rete.

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestion come la seguente, che specifica Amazon DocumentDB come origine. Tieni presente che per compilare il nome dell'indice, la `getMetadata` funzione utilizza come chiave di metadati. *`documentdb_collection`* Se si desidera utilizzare un nome di indice diverso senza il `getMetadata` metodo, è possibile utilizzare la configurazione. `index: "`*`my_index_name`*`"`

```
version: "2"
documentdb-pipeline:
  source:
    documentdb:
      acknowledgments: true
      host: "https://docdb-cluster-id.us-east-1.docdb.amazonaws.com"
      port: 27017
      authentication:
        username: ${aws_secrets:secret:username}
        password: ${aws_secrets:secret:password}
      aws:
      s3_bucket: "bucket-name"
      s3_region: "bucket-region"
      s3_prefix: "path" #optional path for storing the temporary data
      collections:
        - collection: "dbname.collection"
          export: true
          stream: true
  sink:
  - opensearch:
      hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
```

```
      index: "${getMetadata(\"documentdb_collection\")}"
      index_type: custom
      document_id: "${getMetadata(\"primary_key\")}"
      action: "${getMetadata(\"opensearch_action\")}"
      document_version: "${getMetadata(\"document_version\")}"
      document_version_type: "external"
extension:
  aws:
    secrets:
      secret:
        secret_id: "my-docdb-secret"
        region: "us-east-1"
        refresh_interval: PT1H
```

Puoi utilizzare un blueprint Amazon DocumentDB preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Se utilizzi il Console di gestione AWS per creare la tua pipeline, devi anche collegare la pipeline al tuo VPC per utilizzare Amazon DocumentDB come sorgente. Per farlo, trova la sezione Opzioni di rete di origine, seleziona la casella di controllo Collega a VPC e scegli il tuo CIDR da una delle opzioni predefinite fornite. È possibile utilizzare qualsiasi CIDR da uno spazio di indirizzi privato come definito nella Best Current Practice della RFC 1918.

Per fornire un CIDR personalizzato, seleziona Altro dal menu a discesa. Per evitare una collisione negli indirizzi IP tra OpenSearch Ingestion e Amazon DocumentDB, assicurati che il CIDR VPC di Amazon DocumentDB sia diverso dal CIDR per Ingestion. OpenSearch

Per ulteriori informazioni, consulta Configurazione dell'accesso VPC per una pipeline.

Coerenza dei dati

La pipeline garantisce la coerenza dei dati effettuando continuamente polling o ricevendo modifiche dal cluster Amazon DocumentDB e aggiornando i documenti corrispondenti nell'indice. OpenSearch

OpenSearch Ingestion supporta il end-to-end riconoscimento per garantire la durabilità dei dati. Quando una pipeline legge istantanee o flussi, crea dinamicamente partizioni per l'elaborazione parallela. La pipeline contrassegna una partizione come completa quando riceve un riconoscimento dopo aver acquisito tutti i record nel dominio o nella raccolta. OpenSearch

Se desideri importare in una raccolta di ricerca OpenSearch Serverless, puoi generare un ID di documento nella pipeline. Se desideri importare in una raccolta di serie temporali OpenSearch Serverless, tieni presente che la pipeline non genera un ID di documento, quindi devi

`document_id: "${getMetadata(\"primary_key\")}"` ometterlo nella configurazione del pipeline sink.

Una pipeline OpenSearch di Ingestion mappa anche le azioni degli eventi in entrata nelle corrispondenti azioni di indicizzazione di massa per facilitare l'acquisizione dei documenti. Ciò mantiene i dati coerenti, in modo che ogni modifica dei dati in Amazon DocumentDB venga riconciliata con le corrispondenti modifiche al documento. OpenSearch

Mappatura dei tipi di dati

OpenSearch Il servizio mappa dinamicamente i tipi di dati in ogni documento in entrata al tipo di dati corrispondente in Amazon DocumentDB. La tabella seguente mostra come OpenSearch Service mappa automaticamente vari tipi di dati.

Tipo di dati	OpenSearch	Amazon DocumentDB
Numero intero	OpenSearch mappa automaticamente i valori interi di Amazon DocumentDB su numeri interi. OpenSearch OpenSearch mappa dinamicamente il campo in base al primo documento inviato. Se disponi di una combinazione di tipi di dati per lo stesso attributo in Amazon DocumentDB, la mappatura automatica potrebbe non riuscire. Ad esempio, se il primo documento ha un attributo lungo e un documento successivo ha lo stesso attributo come numero intero, OpenSearch non riesce a importare il secondo documento. In questi casi, è necessario fornire un modello di mappatura esplicito che scelga il tipo di numero più flessibile, come il seguente: `{`	Amazon DocumentDB supporta numeri interi.

Tipo di dati	OpenSearch	Amazon DocumentDB

	<pre>"template": {
	"mappings": {
	"properties": {
	"MixedNumberField": {
	"type": "float"
	}
	}
	}
	}
	}</pre>	

Tipo di dati	OpenSearch	Amazon DocumentDB
Long	OpenSearch mappa automaticamente i valori long di Amazon DocumentDB su OpenSearch long. OpenSearch mappa dinamicamente il campo in base al primo documento inviato. Se disponi di una combinazione di tipi di dati per lo stesso attributo in Amazon DocumentDB, la mappatura automatica potrebbe non riuscire. Ad esempio, se il primo documento ha un attributo lungo e un documento successivo ha lo stesso attributo come numero intero, OpenSearch non riesce a importare il secondo documento. In questi casi, è necessario fornire un modello di mappatura esplicito che scelga il tipo di numero più flessibile, come il seguente: `{ "template": { "mappings": { "properties": { "MixedNumberField": { "type": "float" } } } } }`	Amazon DocumentDB supporta i file long.

Tipo di dati	OpenSearch	Amazon DocumentDB
Stringa	OpenSearch mappa automaticamente i valori delle stringhe come testo. In alcune situazioni, come i valori enumerati, è possibile eseguire il mapping al tipo di parola chiave. L'esempio seguente mostra come mappare un attributo Amazon DocumentDB denominato `PartType` a una OpenSearch parola chiave. `{ "template": { "mappings": { "properties": { "PartType": { "type": "keyword" } } } } }`	Amazon DocumentDB supporta le stringhe.

Tipo di dati	OpenSearch	Amazon DocumentDB
Double	OpenSearch mappa automaticamente i valori doppi di Amazon DocumentDB a OpenSearch valori doppi. OpenSearch mappa dinamicamente il campo in base al primo documento inviato. Se disponi di una combinazione di tipi di dati per lo stesso attributo in Amazon DocumentDB, la mappatura automatica potrebbe non riuscire. Ad esempio, se il primo documento ha un attributo lungo e un documento successivo ha lo stesso attributo come numero intero, OpenSearch non riesce a importare il secondo documento. In questi casi, è necessario fornire un modello di mappatura esplicito che scelga il tipo di numero più flessibile, come il seguente: `{ "template": { "mappings": { "properties": { "MixedNumberField": { "type": "float" } } } } }`	Amazon DocumentDB supporta il doppio.

Tipo di dati	OpenSearch	Amazon DocumentDB
Data	Per impostazione predefinita, la data corrisponde a un numero intero in. OpenSearch È possibile definire un modello di mappatura personalizzato per associare una data a una OpenSearch data. `{ "template": { "mappings": { "properties": { "myDateField": { "type": "date", "format": "epoch_second" } } } } }`	Amazon DocumentDB supporta le date.

Tipo di dati	OpenSearch	Amazon DocumentDB
Time stamp	Per impostazione predefinita, il timestamp corrisponde a un numero intero in. OpenSearch È possibile definire un modello di mappatura personalizzato per associare una data a una data. OpenSearch `{ "template": { "mappings": { "properties": { "myTimestampField": { "type": "date", "format": "epoch_second" } } } } }`	Amazon DocumentDB supporta i timestamp.
Booleano	OpenSearch mappa un tipo booleano Amazon DocumentDB in un tipo booleano. OpenSearch	Amazon DocumentDB supporta gli attributi di tipo booleano.

Tipo di dati	OpenSearch	Amazon DocumentDB
Decimale	OpenSearch mappa gli attributi delle mappe di Amazon DocumentDB ai campi annidati. Le stesse mappature si applicano all'interno di un campo nidificato. L'esempio seguente mappa una stringa in un campo nidificato a un tipo di parola chiave in: OpenSearch `{ "template": { "mappings": { "properties": { "myDecimalField": { "type": "double" } } } } }` Con questa mappatura personalizzata, puoi interrogare e aggregare il campo con una precisione a doppio livello. Il valore originale mantiene la massima precisione nella `_source` proprietà del documento. OpenSearch Senza questa mappatura, OpenSearch utilizza il testo per impostazione predefinita.	Amazon DocumentDB supporta i decimali.
Espressione regolare	Il tipo regex crea campi annidati. Questi includono *<myFieldName>* `.pattern` e. *<myFieldName>* `.options`	Amazon DocumentDB supporta le espressioni regolari.

Tipo di dati	OpenSearch	Amazon DocumentDB
Dati binari	OpenSearch mappa automaticamente i dati binari di Amazon DocumentDB su OpenSearch testo. Puoi fornire una mappatura in cui scriverli come campi binari. OpenSearch L'esempio seguente mostra come mappare un campo Amazon DocumentDB denominato `imageData` a un campo OpenSearch binario. `{ "template": { "mappings": { "properties": { "imageData": { "type": "binary" } } } } }`	Amazon DocumentDB supporta campi di dati binari.
ObjectId	I campi con un tipo di ObjectID vengono mappati ai campi di OpenSearch testo. Il valore sarà la rappresentazione in formato stringa di ObjectID.	Amazon DocumentDB supporta gli ObjectID.

Tipo di dati	OpenSearch	Amazon DocumentDB
Null	OpenSearch può importare documenti con il tipo null di Amazon DocumentDB. Salva il valore come valore nullo nel documento. Non esiste alcuna mappatura per questo tipo e questo campo non è indicizzato o ricercabile. Se lo stesso nome di attributo viene utilizzato per un tipo nullo e successivamente passa a un tipo diverso, ad esempio string, OpenSearch crea una mappatura dinamica per il primo valore non nullo. I valori successivi possono comunque essere valori null di Amazon DocumentDB.	Amazon DocumentDB supporta campi di tipo null.
Undefined	OpenSearch può importare documenti con il tipo non definito di Amazon DocumentDB. Salva il valore come valore nullo nel documento. Non esiste alcuna mappatura per questo tipo e questo campo non è indicizzato o ricercabile. Se lo stesso nome di campo viene utilizzato per un tipo non definito e successivamente passa a un tipo diverso, ad esempio string, OpenSearch crea una mappatura dinamica per il primo valore non definito. I valori successivi possono ancora essere valori non definiti di Amazon DocumentDB.	Amazon DocumentDB supporta campi di tipo non definito.

Tipo di dati	OpenSearch	Amazon DocumentDB
MinKey	OpenSearch può importare documenti con il tipo Amazon DocumentDB MinKey. Salva il valore come valore nullo nel documento. Non esiste alcuna mappatura per questo tipo e questo campo non è indicizzato o ricercabile. Se lo stesso nome di campo viene utilizzato per un tipo MinKey e successivamente passa a un tipo diverso, ad esempio string, OpenSearch crea una mappatura dinamica per il primo valore non MinKey. I valori successivi possono ancora essere valori MinKey di Amazon DocumentDB.	Amazon DocumentDB supporta i campi di tipo MinKey.
MaxKey	OpenSearch può importare documenti con il tipo Amazon DocumentDB MaxKey. Salva il valore come valore nullo nel documento. Non esiste alcuna mappatura per questo tipo e questo campo non è indicizzato o ricercabile. Se lo stesso nome di campo viene utilizzato per un tipo di MaxKey e successivamente passa a un tipo diverso, ad esempio string, OpenSearch crea una mappatura dinamica per il primo valore non MaxKey. I valori successivi possono ancora essere valori MaxKey di Amazon DocumentDB.	Amazon DocumentDB supporta i campi di tipo MaxKey.

Ti consigliamo di configurare la dead-letter queue (DLQ) nella pipeline di Ingestion. OpenSearch Se hai configurato la coda, OpenSearch Service invia tutti i documenti non riusciti che non possono essere importati a causa di errori di mappatura dinamica sulla coda.

Nel caso in cui le mappature automatiche falliscano, puoi utilizzare `template_type` e `template_content` nella configurazione della pipeline per definire regole di mappatura esplicite. In alternativa, puoi creare modelli di mappatura direttamente nel tuo dominio di ricerca o nella tua raccolta prima di avviare la pipeline.

Limitazioni

Considera le seguenti limitazioni quando configuri una pipeline di OpenSearch ingestione per Amazon DocumentDB:

- L'integrazione OpenSearch di Ingestion con Amazon DocumentDB attualmente non supporta l'ingestione tra regioni. Il cluster Amazon DocumentDB e la pipeline OpenSearch di Ingestion devono trovarsi nello stesso ambiente. Regione AWS
- L'integrazione OpenSearch di Ingestion con Amazon DocumentDB attualmente non supporta l'ingestione tra account. Il cluster Amazon DocumentDB e la pipeline OpenSearch di Ingestion devono trovarsi nello stesso ambiente. Account AWS
- Una pipeline OpenSearch di ingestion supporta solo un cluster Amazon DocumentDB come origine.
- L'integrazione OpenSearch di Ingestion con Amazon DocumentDB supporta specificamente i cluster basati su istanze di Amazon DocumentDB. Non supporta i cluster elastici di Amazon DocumentDB.
- L'integrazione OpenSearch di Ingestion supporta solo Gestione dei segreti AWS come meccanismo di autenticazione per il tuo cluster Amazon DocumentDB.
- Non puoi aggiornare la configurazione della pipeline esistente per importare dati da un database o una raccolta diversi. È invece necessario creare una nuova pipeline.

Allarmi consigliati CloudWatch

Per prestazioni ottimali, ti consigliamo di utilizzare i seguenti CloudWatch allarmi quando crei una pipeline di OpenSearch ingestion per accedere a un cluster Amazon DocumentDB come origine.

CloudWatch Allarme	Description
<pipeline-name> .doucmentdb.Credenziali modificate	Questa metrica indica la frequenza con cui i segreti vengono ruotati. AWS
<pipeline-name> .documentdb. executorR efreshErrors	Questa metrica indica gli errori nell'aggi ornamento AWS dei segreti.
<pipeline-name> .documentdb. exportRec ordsTotal	Questa metrica indica il numero di record esportati da Amazon DocumentDB.
<pipeline-name> .documentdb. exportRec ordsProcessed	Questa metrica indica il numero di record elaborati dalla pipeline di OpenSearch ingestion e.
<pipeline-name> .documentdb. exportRec ordProcessingErrori	Questa metrica indica il numero di errori di elaborazione in una pipeline di OpenSearc h ingestione durante la lettura dei dati da un cluster Amazon DocumentDB.
<pipeline-name> .documentdb. exportRec ordsSuccessTotale	Questa metrica indica il numero totale di record di esportazione elaborati correttamente.
<pipeline-name> .documentdb. exportRec ordsFailedTotale	Questa metrica indica il numero totale di record di esportazione che non sono stati elaborati.
<pipeline-name> .DocumentDb. Byte ricevuti	Questa metrica indica il numero totale di byte ricevuti da una pipeline di ingestione. OpenSearch
<pipeline-name> .DocumentDB. Byte elaborati	Questa metrica indica il numero totale di byte elaborati da una pipeline di ingestione. OpenSearch
<pipeline-name> .documentdb. exportPar titionQueryTotale	Questa metrica indica il totale della partizione di esportazione.
<pipeline-name> .documentdb. streamRec ordsSuccessTotale	Questa metrica indica il numero di record elaborati con successo dallo stream.

CloudWatch Allarme	Description
<pipeline-name> .documentdb. streamRecordsFailedTotale	Questa metrica indica il numero totale di record non elaborati dallo stream.

Utilizzo di una pipeline OpenSearch di ingestione con Confluent Cloud Kafka

Puoi utilizzare una pipeline di OpenSearch ingestion per lo streaming di dati dai cluster Confluent Cloud Kafka ai domini Amazon Service e alle raccolte Serverless. OpenSearch OpenSearch OpenSearch Ingestion supporta configurazioni di rete pubbliche e private per lo streaming di dati dai cluster Confluent Cloud Kafka a domini o raccolte gestiti da Service o Serverless. OpenSearch OpenSearch

Connettività ai cluster Kafka pubblici di Confluent Cloud

Puoi utilizzare le pipeline di OpenSearch Ingestion per migrare i dati da un cluster Confluent Cloud Kafka con una configurazione pubblica, il che significa che il nome DNS del dominio può essere risolto pubblicamente. A tale scopo, configura una pipeline di OpenSearch Ingestion con il cluster Kafka pubblico di Confluent Cloud come origine e Service o Serverless come destinazione. OpenSearch OpenSearch Questo elabora i dati di streaming da un cluster di origine autogestito a un dominio o raccolta di destinazione gestito. AWS

Prerequisiti

Prima di creare la pipeline OpenSearch di Ingestion, effettuate le seguenti operazioni:

1. Crea un cluster di cluster Confluent Cloud Kafka che funge da fonte. Il cluster deve contenere i dati che desideri importare in Service. OpenSearch
2. Crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless verso cui migrare i dati. Per ulteriori informazioni, consultare the section called “Creazione OpenSearch di domini di servizio” e the section called “Creazione di raccolte”.
3. Configura l'autenticazione sul tuo cluster Confluent Cloud Kafka con. Gestione dei segreti AWSAbilita la rotazione dei segreti seguendo i passaggi in Ruota i segreti. Gestione dei segreti AWS

4. Allega una politica basata sulle risorse al tuo dominio o una politica di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal cluster autogestito al dominio o alla raccolta.

   Il seguente esempio di policy di accesso al dominio consente al ruolo pipeline, creato nel passaggio successivo, di scrivere dati su un dominio. Assicurati di aggiornarlo `resource` con il tuo ARN.

   JSON

   ```
   {
     "Version":"2012-10-17",
     "Statement": [
       {
         "Effect": "Allow",
         "Principal": {
           "AWS": "arn:aws:iam::444455556666:role/pipeline-role"
         },
         "Action": [
           "es:DescribeDomain",
           "es:ESHttp*"
         ],
         "Resource": [
           "arn:aws:es:us-east-1:111122223333:domain/domain-name"
         ]
       }
     ]
   }
   ```

   Per creare un ruolo IAM con le autorizzazioni corrette per accedere ai dati di scrittura nella raccolta o nel dominio, consulta. the section called “Configurazione di ruoli e utenti”

Fase 1: Configurare il ruolo della pipeline

Dopo aver configurato i prerequisiti della pipeline del cluster Confluent Cloud Kafka, configura il ruolo della pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi l'autorizzazione alla scrittura su un dominio di OpenSearch servizio o una raccolta OpenSearch Serverless, nonché l'autorizzazione a leggere segreti da Secrets Manager.

Per gestire l'interfaccia di rete sono necessarie le seguenti autorizzazioni:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ec2:AttachNetworkInterface",
                "ec2:CreateNetworkInterface",
                "ec2:CreateNetworkInterfacePermission",
                "ec2:DeleteNetworkInterface",
                "ec2:DeleteNetworkInterfacePermission",
                "ec2:DetachNetworkInterface",
                "ec2:DescribeNetworkInterfaces"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1:111122223333:network-interface/*",
                "arn:aws:ec2:us-east-1:111122223333:subnet/*",
                "arn:aws:ec2:us-east-1:111122223333:security-group/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeDhcpOptions",
                "ec2:DescribeRouteTables",
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeSubnets",
                "ec2:DescribeVpcs",
                "ec2:Describe*"
            ],
            "Resource": "arn:aws:ec2:us-east-1:111122223333:subnet/*"
        },
        {
            "Effect": "Allow",
            "Action": [ "ec2:CreateTags" ],
            "Resource": "arn:aws:ec2:us-east-1:111122223333:network-interface/*",
            "Condition": {
               "StringEquals": { "aws:RequestTag/OSISManaged": "true" }
            }
        }
    ]
```

```
}
```

Le seguenti sono le autorizzazioni necessarie per leggere i segreti del Gestione dei segreti AWS servizio:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "SecretsManagerReadAccess",
            "Effect": "Allow",
            "Action": ["secretsmanager:GetSecretValue"],
            "Resource": ["arn:aws:secretsmanager:us-
east-1:111122223333:secret:,secret-name"]
        }
    ]
}
```

Per scrivere su un dominio Amazon OpenSearch Service sono necessarie le seguenti autorizzazioni:

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::account-id:role/pipeline-role"
      },
      "Action": ["es:DescribeDomain", "es:ESHttp*"],
      "Resource": "arn:aws:es:region:account-id:domain/domain-name/*"
    }
  ]
}
```

Fase 2: Creare la pipeline

Puoi quindi configurare una pipeline di OpenSearch Ingestion come la seguente, che specifica Confluent Cloud Kafka come origine.

Puoi specificare più OpenSearch domini di servizio come destinazioni per i tuoi dati. Questa funzionalità consente il routing o la replica condizionale dei dati in entrata in più domini di servizio. OpenSearch

Puoi anche migrare i dati da un cluster Confluent Kafka di origine a una raccolta VPC serverless. OpenSearch Assicurati di fornire una politica di accesso alla rete all'interno della configurazione della pipeline. È possibile utilizzare un registro di schemi Confluent per definire uno schema Confluent.

```
version: "2"
kafka-pipeline:
  source:
    kafka:
      encryption:
        type: "ssl"
      topics:
        - name: "topic-name"
          group_id: "group-id"
      bootstrap_servers:
        - "bootstrap-server.us-east-1.aws.private.confluent.cloud:9092"
      authentication:
        sasl:
          plain:
            username: ${aws_secrets:confluent-kafka-secret:username}
            password: ${aws_secrets:confluent-kafka-secret:password}
      schema:
        type: confluent
        registry_url: https://my-registry.us-east-1.aws.confluent.cloud
        api_key: "${{aws_secrets:schema-secret:schema_registry_api_key}}"
        api_secret: "${{aws_secrets:schema-secret:schema_registry_api_secret}}"
        basic_auth_credentials_source: "USER_INFO"
  sink:
  - opensearch:
      hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
      aws:
          region: "us-east-1"
  aws:
    secrets:
      confluent-kafka-secret:
        secret_id: "my-kafka-secret"
        region: "us-east-1"
      schema-secret:
        secret_id: "my-self-managed-kafka-schema"
```

```
        region: "us-east-1"
```

È possibile utilizzare un blueprint preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Connettività ai cluster Confluent Cloud Kafka in un VPC

Puoi anche utilizzare le pipeline di OpenSearch Ingestion per migrare i dati da un cluster Confluent Cloud Kafka in esecuzione in un VPC. A tale scopo, configura una pipeline di OpenSearch ingestion con un cluster Confluent Cloud Kafka come origine e Service o Serverless come destinazione. OpenSearch OpenSearch Questo elabora i dati di streaming da un cluster di origine Confluent Cloud Kafka a un dominio o una raccolta di destinazione gestiti. AWS

OpenSearch Ingestion supporta i cluster Confluent Cloud Kafka configurati in tutte le modalità di rete supportate in Confluent. Le seguenti modalità di configurazione di rete sono supportate come origine in Ingestion: OpenSearch

- AWS Peering VPC
- AWS PrivateLink per cluster dedicati
- AWS PrivateLink per cluster aziendali
- AWS Transit Gateway

Prerequisiti

Prima di creare la pipeline OpenSearch di Ingestion, effettuate le seguenti operazioni:

1. Crea un cluster Confluent Cloud Kafka con una configurazione di rete VPC che contenga i dati che desideri inserire nel servizio. OpenSearch
2. Crea un dominio di OpenSearch servizio o una raccolta OpenSearch Serverless verso cui migrare i dati. Per ulteriori informazioni, vedere Per ulteriori informazioni, vedere the section called “Creazione OpenSearch di domini di servizio” e. the section called “Creazione di raccolte”
3. Configura l'autenticazione sul tuo cluster Confluent Cloud Kafka con. Gestione dei segreti AWSAbilita la rotazione dei segreti seguendo i passaggi in Ruota i segreti. Gestione dei segreti AWS
4. Ottieni l'ID del VPC che ha accesso al cluster Confluent Cloud Kafka. Scegli il VPC CIDR da utilizzare con Ingestion. OpenSearch

> **Note**
>
> Se utilizzi il Console di gestione AWS per creare la tua pipeline, devi anche collegare la pipeline OpenSearch Ingestion al tuo VPC per utilizzare il cluster Confluent Cloud Kafka. Per farlo, trova la sezione Configurazione di rete, seleziona la casella di controllo Collega a VPC e scegli il tuo CIDR da una delle opzioni predefinite fornite oppure selezionane una personalizzata. È possibile utilizzare qualsiasi CIDR da uno spazio di indirizzi privato come definito nella RFC 1918 Best Current Practice.
> Per fornire un CIDR personalizzato, seleziona Altro dal menu a discesa. Per evitare una collisione negli indirizzi IP tra OpenSearch Ingestion e Self-managed OpenSearch, assicurati che il CIDR OpenSearch VPC autogestito sia diverso dal CIDR per Ingestion. OpenSearch

5. Allega una policy basata sulle risorse al tuo dominio o una policy di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal cluster autogestito al dominio o alla raccolta.

> **Note**
>
> Se utilizzi AWS PrivateLink per connettere Confluent Cloud Kafka, dovrai configurare le opzioni DHCP VPC. I nomi host DNS e la risoluzione DNS devono essere abilitati.
> In particolare, utilizzate i seguenti valori del set di opzioni:
> Cluster aziendali:
>
> ```
> domain-name: aws.private.confluent.cloud
> domain-name-servers: AmazonProvidedDNS
> ```
>
> Cluster dedicati:
>
> ```
> domain-name: aws.confluent.cloud
> domain-name-servers: AmazonProvidedDNS
> ```
>
> Questa modifica garantisce che la risoluzione DNS per l' PrivateLink endpoint Confluent funzioni correttamente all'interno del VPC.

Il seguente esempio di policy di accesso al dominio consente al ruolo pipeline, creato nel passaggio successivo, di scrivere dati su un dominio. Assicurati di aggiornarlo `resource` con il tuo ARN.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::444455556666:role/pipeline-role"
      },
      "Action": [
        "es:DescribeDomain",
        "es:ESHttp*"
      ],
      "Resource": [
        "arn:aws:es:us-east-1:111122223333:domain/domain-name"
      ]
    }
  ]
}
```

Per creare un ruolo IAM con le autorizzazioni corrette per accedere ai dati di scrittura nella raccolta o nel dominio, consulta. the section called “Configurazione di ruoli e utenti”

Fase 1: Configurare il ruolo della pipeline

Dopo aver impostato i prerequisiti della pipeline, configura il ruolo della pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi le seguenti autorizzazioni nel ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
```

```
            "Sid": "SecretsManagerReadAccess",
            "Effect": "Allow",
            "Action": [
                "secretsmanager:GetSecretValue"
            ],
            "Resource": ["arn:aws:secretsmanager:us-
east-1:111122223333:secret:secret-name"]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:AttachNetworkInterface",
                "ec2:CreateNetworkInterface",
                "ec2:CreateNetworkInterfacePermission",
                "ec2:DeleteNetworkInterface",
                "ec2:DeleteNetworkInterfacePermission",
                "ec2:DetachNetworkInterface",
                "ec2:DescribeNetworkInterfaces"
            ],
            "Resource": [
                "arn:aws:ec2:*:*:network-interface/*",
                "arn:aws:ec2:*:*:subnet/*",
                "arn:aws:ec2:*:*:security-group/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeDhcpOptions",
                "ec2:DescribeRouteTables",
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeSubnets",
                "ec2:DescribeVpcs",
                "ec2:Describe*"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:CreateTags"
            ],
            "Resource": "arn:aws:ec2:*:*:network-interface/*",
            "Condition": {
```

```
                "StringEquals":
                    {
                        "aws:RequestTag/OSISManaged": "true"
                    }
            }
        }
    ]
}
```

Devi fornire le autorizzazioni Amazon EC2 di cui sopra sul ruolo IAM che usi per creare la pipeline di OpenSearch ingestione, poiché la pipeline utilizza queste autorizzazioni per creare ed eliminare un'interfaccia di rete nel tuo VPC. La pipeline può accedere al cluster Kafka solo tramite questa interfaccia di rete.

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestione come la seguente, che specifica Kafka come origine.

È possibile specificare più domini di OpenSearch servizio come destinazioni per i dati. Questa funzionalità consente il routing o la replica condizionale dei dati in entrata in più domini di servizio. OpenSearch

Puoi anche migrare i dati da un cluster Confluent Kafka di origine a una raccolta VPC serverless. OpenSearch Assicurati di fornire una politica di accesso alla rete all'interno della configurazione della pipeline. È possibile utilizzare un registro di schemi Confluent per definire uno schema Confluent.

```
 version: "2"
kafka-pipeline:
  source:
    kafka:
      encryption:
        type: "ssl"
      topics:
        - name: "topic-name"
          group_id: "group-id"
      bootstrap_servers:
        - "bootstrap-server.us-east-1.aws.private.confluent.cloud:9092"
      authentication:
        sasl:
          plain:
```

```
            username: ${aws_secrets:confluent-kafka-secret:username}
            password: ${aws_secrets:confluent-kafka-secret:password}
      schema:
        type: confluent
        registry_url: https://my-registry.us-east-1.aws.confluent.cloud
        api_key: "${{aws_secrets:schema-secret:schema_registry_api_key}}"
        api_secret: "${{aws_secrets:schema-secret:schema_registry_api_secret}}"
        basic_auth_credentials_source: "USER_INFO"
  sink:
  - opensearch:
      hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
      aws:
          region: "us-east-1"
      index: "confluent-index"
extension:
  aws:
    secrets:
      confluent-kafka-secret:
        secret_id: "my-kafka-secret"
        region: "us-east-1"
      schema-secret:
        secret_id: "my-self-managed-kafka-schema"
        region: "us-east-2"
```

Utilizzo di una pipeline di OpenSearch ingestione con Amazon Streaming gestito per Apache Kafka

Puoi utilizzare il plug-in Kafka per inserire dati da Amazon Managed Streaming for Apache Kafka (Amazon MSK) nella tua pipeline di Ingestion. OpenSearch Con Amazon MSK, puoi creare ed eseguire applicazioni che utilizzano Apache Kafka per elaborare dati in streaming. OpenSearch Ingestion utilizza AWS PrivateLink per connettersi ad Amazon MSK. Puoi importare dati da cluster Amazon MSK e Amazon MSK Serverless. L'unica differenza tra i due processi sono i passaggi preliminari da eseguire prima di configurare la pipeline.

Argomenti

- Prerequisiti Amazon MSK forniti
- Prerequisiti per Amazon MSK Serverless
- Fase 1: Configurare un ruolo della pipeline
- Fase 2: Creare la pipeline

- Fase 3: (Facoltativo) Usa il registro degli schemi AWS Glue
- Fase 4: (Facoltativo) Configurazione delle unità di calcolo consigliate (OCUs) per la pipeline Amazon MSK

Prerequisiti Amazon MSK forniti

Prima di creare la pipeline di OpenSearch Ingestion, esegui i seguenti passaggi:

1. Crea un cluster con provisioning Amazon MSK seguendo i passaggi descritti nella sezione Creazione di un cluster nella Amazon Managed Streaming for Apache Kafka Developer Guide. Per il tipo di broker, scegli qualsiasi opzione tranne i `t3` tipi, poiché questi non sono supportati da Ingestion. OpenSearch
2. Dopo che il cluster ha uno stato Attivo, segui i passaggi in Attivare la connettività multi-VPC.
3. Segui i passaggi riportati in Allegare una politica del cluster al cluster MSK per allegare una delle seguenti politiche, a seconda che il cluster e la pipeline coincidano. Account AWS Questa policy consente a OpenSearch Ingestion di creare una AWS PrivateLink connessione al tuo cluster Amazon MSK e leggere dati da argomenti di Kafka. Assicurati di aggiornarlo `resource` con il tuo ARN.

   Le seguenti politiche si applicano quando il cluster e la pipeline coincidono: Account AWS

   JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "osis.amazonaws.com"
      },
      "Action": [
        "kafka:CreateVpcConnection",
        "kafka:DescribeClusterV2"
      ],
      "Resource": "arn:aws:kafka:us-east-1:111122223333:cluster/cluster-
name/cluster-id"
    },
    {
      "Effect": "Allow",
```

```
      "Principal": {
        "Service": "osis-pipelines.amazonaws.com"
      },
      "Action": [
        "kafka:CreateVpcConnection",
        "kafka:GetBootstrapBrokers",
        "kafka:DescribeClusterV2"
      ],
      "Resource": "arn:aws:kafka:us-east-1:111122223333:cluster/cluster-
name/cluster-id"
    }
  ]
}
```

Se il tuo cluster Amazon MSK si trova in una pipeline Account AWS diversa dalla tua, allega invece la seguente policy. Tieni presente che l'accesso tra account è possibile solo con i cluster Amazon MSK forniti e non con i cluster Amazon MSK Serverless. L'ARN per il AWS `principal` dovrebbe essere l'ARN per lo stesso ruolo di pipeline fornito alla configurazione della pipeline:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "osis.amazonaws.com"
      },
      "Action": [
        "kafka:CreateVpcConnection",
        "kafka:DescribeClusterV2"
      ],
      "Resource": "arn:aws:kafka:us-east-1:111122223333:cluster/cluster-
name/cluster-id"
    },
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "osis-pipelines.amazonaws.com"
      },
      "Action": [
        "kafka:CreateVpcConnection",
```

```
        "kafka:GetBootstrapBrokers",
        "kafka:DescribeClusterV2"
      ],
      "Resource": "arn:aws:kafka:us-east-1:111122223333:cluster/cluster-
name/cluster-id"
    },
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::444455556666:role/pipeline-role"
      },
      "Action": [
        "kafka-cluster:*",
        "kafka:*"
      ],
      "Resource": [
        "arn:aws:kafka:us-east-1:111122223333:cluster/cluster-name/cluster-id",
        "arn:aws:kafka:us-east-1:111122223333:topic/cluster-name/cluster-id/*",
        "arn:aws:kafka:us-east-1:111122223333:group/cluster-name/*"
      ]
    }
  ]
}
```

4. Crea un argomento di Kafka seguendo la procedura descritta in Creare un argomento. Assicurati che *BootstrapServerString* sia uno dei bootstrap privati dell'endpoint (Single-VPC). URLs Il valore per `--replication-factor` dovrebbe essere 2 o3, in base al numero di zone del tuo cluster Amazon MSK. Il valore di `--partitions` deve essere almeno`10`.
5. Produci e consuma dati seguendo i passaggi descritti in Produrre e consumare dati. Ancora una volta, assicurati che *BootstrapServerString* sia uno dei tuoi endpoint privati (Single-VPC) bootstrap. URLs

Prerequisiti per Amazon MSK Serverless

Prima di creare la pipeline di OpenSearch Ingestion, esegui i seguenti passaggi:

1. Crea un cluster Serverless Amazon MSK seguendo i passaggi descritti in Creare un cluster Serverless MSK nella Amazon Managed Streaming for Apache Kafka Developer Guide.
2. Dopo che il cluster ha raggiunto lo stato Attivo, segui i passaggi in Allegare una politica del cluster al cluster MSK per allegare la seguente politica. Assicurati di aggiornarlo `resource` con il tuo ARN.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "osis.amazonaws.com"
      },
      "Action": [
        "kafka:CreateVpcConnection",
        "kafka:DescribeClusterV2"
      ],
      "Resource": "arn:aws:kafka:us-east-1:111122223333:cluster/cluster-
name/cluster-id"
    },
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "osis-pipelines.amazonaws.com"
      },
      "Action": [
        "kafka:CreateVpcConnection",
        "kafka:GetBootstrapBrokers",
        "kafka:DescribeClusterV2"
      ],
      "Resource": "arn:aws:kafka:us-east-1:111122223333:cluster/cluster-
name/cluster-id"
    }
  ]
}
```

Questa policy consente a OpenSearch Ingestion di creare una AWS PrivateLink connessione al tuo cluster Amazon MSK Serverless e leggere dati da argomenti di Kafka. Questa politica si applica quando il cluster e la pipeline coincidono Account AWS, il che deve essere vero in quanto Amazon MSK Serverless non supporta l'accesso tra account.

3. Crea un argomento di Kafka seguendo i passaggi in Creare un argomento. Assicurati che *BootstrapServerString* sia uno dei tuoi bootstrap IAM Simple Authentication and Security Layer (SASL). URLs Il valore per `--replication-factor` dovrebbe essere 2 o3, in base al

numero di zone del tuo cluster Amazon MSK Serverless. Il valore di `--partitions` deve essere almeno `10`

4. Produci e consuma dati seguendo i passaggi descritti in Produrre e consumare dati. Ancora una volta, assicurati che *`BootstrapServerString`* sia uno dei tuoi bootstrap URLs IAM Simple Authentication and Security Layer (SASL).

Fase 1: Configurare un ruolo della pipeline

Dopo aver configurato il cluster Amazon MSK con provisioning o serverless, aggiungi le seguenti autorizzazioni Kafka nel ruolo pipeline che desideri utilizzare nella configurazione della pipeline:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:Connect",
                "kafka-cluster:AlterCluster",
                "kafka-cluster:DescribeCluster",
                "kafka:DescribeClusterV2",
                "kafka:GetBootstrapBrokers"
            ],
            "Resource": [
                "arn:aws:kafka:us-east-1:111122223333:cluster/cluster-
name/cluster-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:*Topic*",
                "kafka-cluster:ReadData"
            ],
            "Resource": [
                "arn:aws:kafka:us-east-1:111122223333:topic/cluster-name/cluster-
id/topic-name"
            ]
        },
```

```
        {
            "Effect": "Allow",
            "Action": [
                "kafka-cluster:AlterGroup",
                "kafka-cluster:DescribeGroup"
            ],
            "Resource": [
                "arn:aws:kafka:us-east-1:111122223333:group/cluster-name/*"
            ]
        }
    ]
}
```

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestione come la seguente, che specifica Kafka come origine:

```
version: "2"
log-pipeline:
  source:
    kafka:
      acknowledgements: true
      topics:
      - name: "topic-name"
        group_id: "grouplambd-id"
      aws:
        msk:
          arn: "arn:aws:kafka:region:account-id:cluster/cluster-name/cluster-id"
        region: "us-west-2"
  processor:
  - grok:
      match:
        message:
        - "%{COMMONAPACHELOG}"
  - date:
      destination: "@timestamp"
      from_time_received: true
  sink:
  - opensearch:
      hosts: ["https://search-domain-endpoint.us-east-1es.amazonaws.com"]
      index: "index_name"
```

```
      aws_region: "region"
      aws_sigv4: true
```

Puoi utilizzare un blueprint Amazon MSK preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Fase 3: (Facoltativo) Usa il registro degli schemi AWS Glue

Quando usi OpenSearch Ingestion con Amazon MSK, puoi utilizzare il formato di dati AVRO per gli schemi ospitati nel registro degli schemi. AWS Glue Con lo AWS Glue Schema Registry, puoi scoprire, controllare ed evolvere centralmente gli schemi dei flussi di dati.

Per utilizzare questa opzione, abilita lo schema `type` nella configurazione della pipeline:

```
schema:
  type: "aws_glue"
```

Devi inoltre fornire i permessi AWS Glue di accesso in lettura nel tuo ruolo di pipeline. È possibile utilizzare la politica AWS gestita denominata. AWSGlueSchemaRegistryReadonlyAccess Inoltre, il registro deve trovarsi nella stessa Account AWS area geografica della pipeline di OpenSearch Ingestion.

Fase 4: (Facoltativo) Configurazione delle unità di calcolo consigliate (OCUs) per la pipeline Amazon MSK

Ogni unità di elaborazione ha un utente per argomento. I broker bilanciano le partizioni tra questi consumatori per un determinato argomento. Tuttavia, quando il numero di partizioni è maggiore del numero di consumatori, Amazon MSK ospita più partizioni per ogni consumatore. OpenSearch Ingestion dispone della scalabilità automatica integrata per aumentare o ridurre in base all'utilizzo della CPU o al numero di record in sospeso nella pipeline.

Per prestazioni ottimali, distribuisci le partizioni su più unità di calcolo per l'elaborazione parallela. Se gli argomenti hanno un numero elevato di partizioni (ad esempio, più di 96, che è il numero massimo OCUs per pipeline), ti consigliamo di configurare una pipeline con 1—96. OCUs Questo perché si ridimensionerà automaticamente in base alle esigenze. Se un argomento ha un numero basso di partizioni (ad esempio, meno di 96), mantieni l'unità di calcolo massima uguale al numero di partizioni.

Quando una pipeline ha più di un argomento, scegliete l'argomento con il maggior numero di partizioni come riferimento per configurare il numero massimo di unità di calcolo. Aggiungendo

un'altra pipeline con un nuovo set allo stesso argomento e OCUs allo stesso gruppo di consumatori, puoi scalare il throughput in modo quasi lineare.

Utilizzo di una pipeline OpenSearch di ingestione con Amazon RDS

Puoi utilizzare una pipeline OpenSearch di importazione con Amazon RDS per esportare dati esistenti e trasmettere le modifiche (ad esempio creazione, aggiornamento ed eliminazione) verso domini e raccolte di Amazon OpenSearch Service. La pipeline OpenSearch Ingestion incorpora l'infrastruttura Change Data Capture (CDC) per fornire un modo su larga scala e a bassa latenza per lo streaming continuo di dati da Amazon RDS. Sono supportati RDS per MySQL e RDS per PostgreSQL.

Esistono due modi per utilizzare Amazon RDS come origine per elaborare i dati: con o senza uno snapshot iniziale completo. Uno snapshot iniziale completo è uno snapshot di tabelle specificate e questo snapshot viene esportato in Amazon S3. Da lì, una pipeline di OpenSearch ingestione lo invia a un indice in un dominio o lo partiziona in più indici di un dominio. Per mantenere OpenSearch coerenti i dati in Amazon RDS, la pipeline sincronizza tutti gli eventi di creazione, aggiornamento ed eliminazione nelle tabelle delle istanze Amazon RDS con i documenti salvati nell'indice o negli indici. OpenSearch

Quando usi uno snapshot iniziale completo, la tua pipeline di OpenSearch Ingestion prima lo inserisce e poi inizia a leggere i dati dai flussi di modifiche di Amazon RDS. Alla fine recupera e mantiene la coerenza dei dati quasi in tempo reale tra Amazon RDS e. OpenSearch

Puoi anche utilizzare l'integrazione di OpenSearch Ingestion con Amazon RDS per tenere traccia delle modifiche, acquisire i dati e importare tutti gli aggiornamenti in Aurora to. OpenSearch Scegli questa opzione se disponi già di uno snapshot completo di un altro meccanismo o se desideri semplicemente acquisire tutte le modifiche ai dati in un'istanza Amazon RDS.

Quando scegli questa opzione, devi configurare la registrazione binaria di Amazon RDS for MySQL o configurare la replica logica per l'istanza database Amazon RDS for PostgressSQL.

Argomenti

- RDS per MySQL
- RDS per PostgreSQL

RDS per MySQL

Completa i seguenti passaggi per configurare una pipeline di OpenSearch ingestione con Amazon RDS for RDS for MySQL.

Argomenti

- Prerequisiti RDS per MySQL
- Fase 1: Configurare il ruolo della pipeline
- Fase 2: Creare la pipeline
- Coerenza dei dati
- Mappatura dei tipi di dati
- Limitazioni
- CloudWatch Allarmi consigliati

Prerequisiti RDS per MySQL

Prima di creare la pipeline di OpenSearch Ingestion, effettuate le seguenti operazioni:

1. Crea un gruppo di parametri DB personalizzato in Amazon RDS per configurare la registrazione binaria e impostare i seguenti parametri.

   ```
   binlog_format=ROW
   binlog_row_image=full
   binlog_row_metadata=FULL
   ```

   Inoltre, assicurati che il `binlog_row_value_options` parametro non sia impostato su. `PARTIAL_JSON`

   Per ulteriori informazioni, vedere Configurazione di RDS per la registrazione binaria MySQL.
2. Seleziona o crea un'istanza DB RDS for MySQL e associa il gruppo di parametri creato nel passaggio precedente all'istanza DB.
3. Verifica che i backup automatici siano abilitati sul database. Per ulteriori informazioni, vedere Abilitazione dei backup automatici.
4. Configura la conservazione dei log binari con un tempo sufficiente per la replica, ad esempio 24 ore. Per ulteriori informazioni, consulta Impostazione e visualizzazione della configurazione del log binario nella Amazon RDS User Guide.

5. Configura l'autenticazione di nome utente e password sulla tua istanza Amazon RDS utilizzando la gestione delle password con Amazon RDS e. Gestione dei segreti AWS Puoi anche creare una username/password combinazione creando un segreto di Secrets Manager.
6. Se utilizzi la funzionalità di snapshot iniziale completa, crea un ruolo AWS KMS key e un ruolo IAM per l'esportazione dei dati da Amazon RDS ad Amazon S3.

   Il ruolo IAM deve avere la seguente politica di autorizzazione:

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "ExportPolicy",
            "Effect": "Allow",
            "Action": [
                "s3:PutObject*",
                "s3:ListBucket",
                "s3:GetObject*",
                "s3:DeleteObject*",
                "s3:GetBucketLocation"
            ],
            "Resource": [
                "arn:aws:s3:::s3-bucket-used-in-pipeline",
                "arn:aws:s3:::s3-bucket-used-in-pipeline/*"
            ]
        }
    ]
}
```

   Il ruolo dovrebbe inoltre avere le seguenti relazioni di fiducia:

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "export.rds.amazonaws.com"
```

```
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

7. Seleziona o crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless. Per ulteriori informazioni, consulta Creazione di domini OpenSearch di servizio e Creazione di raccolte.
8. Allega una politica basata sulle risorse al tuo dominio o una politica di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dalla tua istanza DB Amazon RDS al tuo dominio o alla tua raccolta.

Fase 1: Configurare il ruolo della pipeline

Dopo aver configurato i prerequisiti della pipeline Amazon RDS, configura il ruolo della pipeline da utilizzare nella configurazione della pipeline. Aggiungi anche le seguenti autorizzazioni per il codice sorgente Amazon RDS al ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
    {
    "Sid": "allowReadingFromS3Buckets",
    "Effect": "Allow",
    "Action": [
    "s3:GetObject",
    "s3:DeleteObject",
    "s3:GetBucketLocation",
    "s3:ListBucket",
    "s3:PutObject"
    ],
    "Resource": [
    "arn:aws:s3:::s3_bucket",
    "arn:aws:s3:::s3_bucket/*"
    ]
    },
    {
    "Sid": "allowNetworkInterfacesActions",
    "Effect": "Allow",
```

```
    "Action": [
    "ec2:AttachNetworkInterface",
    "ec2:CreateNetworkInterface",
    "ec2:CreateNetworkInterfacePermission",
    "ec2:DeleteNetworkInterface",
    "ec2:DeleteNetworkInterfacePermission",
    "ec2:DetachNetworkInterface",
    "ec2:DescribeNetworkInterfaces"
    ],
    "Resource": [
    "arn:aws:ec2:*:111122223333:network-interface/*",
    "arn:aws:ec2:*:111122223333:subnet/*",
    "arn:aws:ec2:*:111122223333:security-group/*"
    ]
    },
    {
    "Sid": "allowDescribeEC2",
    "Effect": "Allow",
    "Action": [
    "ec2:Describe*"
    ],
    "Resource": "*"
    },
    {
    "Sid": "allowTagCreation",
    "Effect": "Allow",
    "Action": [
    "ec2:CreateTags"
    ],
    "Resource": "arn:aws:ec2:*:111122223333:network-interface/*",
    "Condition": {
    "StringEquals": {
    "aws:RequestTag/OSISManaged": "true"
    }
    }
    },
    {
    "Sid": "AllowDescribeInstances",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBInstances"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:db:*"
```

```
    ]
    },
    {
    "Sid": "AllowSnapshots",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBSnapshots",
    "rds:CreateDBSnapshot",
    "rds:AddTagsToResource"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:db:DB-id",
    "arn:aws:rds:us-east-2:111122223333:snapshot:DB-id*"
    ]
    },
    {
    "Sid": "AllowExport",
    "Effect": "Allow",
    "Action": [
    "rds:StartExportTask"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:snapshot:DB-id*"
    ]
    },
    {
    "Sid": "AllowDescribeExports",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeExportTasks"
    ],
    "Resource": "*",
    "Condition": {
    "StringEquals": {
    "aws:RequestedRegion": "us-east-2",
    "aws:ResourceAccount": "111122223333"
    }
    }
    },
    {
    "Sid": "AllowAccessToKmsForExport",
    "Effect": "Allow",
    "Action": [
    "kms:Decrypt",
```

```
    "kms:Encrypt",
    "kms:DescribeKey",
    "kms:RetireGrant",
    "kms:CreateGrant",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*"
    ],
    "Resource": [
    "arn:aws:kms:us-east-2:111122223333:key/export-key-id"
    ]
    },
    {
    "Sid": "AllowPassingExportRole",
    "Effect": "Allow",
    "Action": "iam:PassRole",
    "Resource": [
    "arn:aws:iam::111122223333:role/export-role"
    ]
    },
    {
    "Sid": "SecretsManagerReadAccess",
    "Effect": "Allow",
    "Action": [
    "secretsmanager:GetSecretValue"
    ],
    "Resource": [
    "arn:aws:secretsmanager:*:111122223333:secret:*"
    ]
    }
    ]
    }
```

Fase 2: Creare la pipeline

Configurate una pipeline OpenSearch di ingestione simile alla seguente. La pipeline di esempio specifica un'istanza Amazon RDS come origine.

```
version: "2"
rds-mysql-pipeline:
  source:
    rds:
      db_identifier: "instance-id"
```

```
      engine: mysql
      database: "database-name"
      tables:
        include:
          - "table1"
          - "table2"
      s3_bucket: "bucket-name"
      s3_region: "bucket-region"
      s3_prefix: "prefix-name"
      export:
        kms_key_id: "kms-key-id"
        iam_role_arn: "export-role-arn"
      stream: true
      aws:
        sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        region: "us-east-1"
      authentication:
        username: ${{aws_secrets:secret:username}}
        password: ${{aws_secrets:secret:password}}
  sink:
    - opensearch:
        hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
        index: "${getMetadata(\"table_name\")}"
        index_type: custom
        document_id: "${getMetadata(\"primary_key\")}"
        action: "${getMetadata(\"opensearch_action\")}"
        document_version: "${getMetadata(\"document_version\")}"
        document_version_type: "external"
        aws:
          sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
          region: "us-east-1"
extension:
  aws:
    secrets:
      secret:
        secret_id: "rds-secret-id"
        region: "us-east-1"
        sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        refresh_interval: PT1H
```

Puoi utilizzare un blueprint Amazon RDS preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta Lavorare con i progetti.

Per utilizzare Amazon Aurora come sorgente, devi configurare l'accesso VPC per la pipeline. Il VPC scelto deve essere lo stesso VPC utilizzato dalla sorgente Amazon Aurora. Quindi scegli una o più sottoreti e uno o più gruppi di sicurezza VPC. Tieni presente che la pipeline richiede l'accesso di rete a un database Aurora MySQL, quindi dovresti anche verificare che il cluster Aurora sia configurato con un gruppo di sicurezza VPC che consenta il traffico in entrata dal gruppo di sicurezza VPC della pipeline alla porta del database. Per ulteriori informazioni, consulta Controllo dell'accesso con gruppi di sicurezza.

Se utilizzi il Console di gestione AWS per creare la tua pipeline, devi anche collegare la pipeline al tuo VPC per utilizzare Amazon Aurora come sorgente. Per fare ciò, trova la sezione Configurazione di rete, scegli Collega a VPC e scegli il tuo CIDR da una delle opzioni predefinite fornite, oppure seleziona la tua. È possibile utilizzare qualsiasi CIDR da uno spazio di indirizzi privato come definito nella Best Current Practice della RFC 1918.

Per fornire un CIDR personalizzato, seleziona Altro dal menu a discesa. Per evitare una collisione tra gli indirizzi IP tra OpenSearch Ingestion e Amazon RDS, assicurati che il CIDR VPC di Amazon RDS sia diverso dal CIDR per Ingestion. OpenSearch

Per ulteriori informazioni, consulta Configurazione dell'accesso VPC per una pipeline.

Coerenza dei dati

La pipeline garantisce la coerenza dei dati effettuando continuamente polling o ricevendo modifiche dall'istanza Amazon RDS e aggiornando i documenti corrispondenti nell'indice. OpenSearch

OpenSearch Ingestion supporta il end-to-end riconoscimento per garantire la durabilità dei dati. Quando una pipeline legge istantanee o flussi, crea dinamicamente partizioni per l'elaborazione parallela. La pipeline contrassegna una partizione come completa quando riceve un riconoscimento dopo aver acquisito tutti i record nel dominio o nella raccolta. OpenSearch Se desideri importare in una raccolta di ricerca OpenSearch Serverless, puoi generare un ID di documento nella pipeline. Se desideri importare in una raccolta di serie temporali OpenSearch Serverless, tieni presente che la pipeline non genera un ID di documento, quindi devi `document_id: "${getMetadata(\"primary_key\")}"` ometterlo nella configurazione del pipeline sink.

Una pipeline OpenSearch di Ingestion mappa anche le azioni degli eventi in entrata nelle corrispondenti azioni di indicizzazione di massa per facilitare l'acquisizione dei documenti. Ciò mantiene i dati coerenti, in modo che ogni modifica dei dati in Amazon RDS venga riconciliata con le corrispondenti modifiche del documento. OpenSearch

Mappatura dei tipi di dati

OpenSearch La pipeline di ingestione mappa i tipi di dati MySQL su rappresentazioni adatte all'utilizzo da parte di domini o raccolte di servizi. OpenSearch Se non è definito alcun modello di mappatura OpenSearch, determina OpenSearch automaticamente i tipi di campo con mappatura dinamica in base al primo documento inviato. È inoltre possibile definire in modo esplicito i tipi di campo più adatti alle proprie esigenze OpenSearch tramite un modello di mappatura.

La tabella seguente elenca i tipi di dati MySQL e OpenSearch i tipi di campo corrispondenti. La colonna Tipo di OpenSearch campo predefinito mostra il tipo di campo corrispondente OpenSearch se non è definita alcuna mappatura esplicita. In questo caso, determina OpenSearch automaticamente i tipi di campo con mappatura dinamica. La colonna Tipo di OpenSearch campo consigliato è il tipo di campo corrispondente che si consiglia di specificare esplicitamente in un modello di mappatura. Questi tipi di campo sono più allineati ai tipi di dati in MySQL e di solito possono abilitare migliori funzionalità di ricerca disponibili in. OpenSearch

Tipi di dati MySQL	Tipo di campo predefinito OpenSearch	Tipo di OpenSearch campo consigliato
BIGINT	Long	Long
BIGINT UNSIGNED	Long	lungo senza segno
BIT	Long	byte, corto, intero o lungo a seconda del numero di bit
DECIMAL	testo	doppio o parola chiave
DOUBLE	virgola mobile	virgola mobile a doppia precisione
FLOAT	virgola mobile	virgola mobile
INT	Long	intero
INT UNSIGNED	Long	Long
MEDIUMIN	Long	intero

Tipi di dati MySQL	Tipo di campo predefinito OpenSearch	Tipo di OpenSearch campo consigliato
MEDIUMINT UNSIGNED	Long	intero
NUMERIC	testo	doppio o parola chiave
SMALLINT	Long	short
SMALLINT UNSIGNED	Long	intero
TINYINT	Long	byte
TINYINT UNSIGNED	Long	short
BINARY	testo	binary
BLOB	testo	binary
CHAR	testo	testo
ENUM	testo	Parola chiave
LONGBLOB	testo	binary
LONGTEXT	testo	testo
MEDIUMBLOB	testo	binary
MEDIUMTEXT	testo	testo
SET	testo	Parola chiave
TEXT	testo	testo
TINYBLOB	testo	binary

Tipi di dati MySQL	Tipo di campo predefinito OpenSearch	Tipo di OpenSearch campo consigliato
TINYTEXT	testo	testo
VARBINAR	testo	binary
VARCHAR	testo	testo
DATE	lungo (in millisecondi di epoca)	data
DATETIME	long (in millisecondi epocali)	data
TIME	long (in millisecondi epocali)	data
TIMESTAM	long (in millisecondi epocali)	data
ANNO	long (in millisecondi epocali)	data
GEOMETR	testo (in formato WKT)	geo_shape
GEOMETR OLLECTIO	testo (in formato WKT)	geo_shape
LINESTRIN G	testo (in formato WKT)	geo_shape
MULTILINE STRING	testo (in formato WKT)	geo_shape
MULTIPOI T	testo (in formato WKT)	geo_shape
MULTIPOL GON	testo (in formato WKT)	geo_shape
POINT	testo (in formato WKT)	geo_point o geo_shape
POLYGON	testo (in formato WKT)	geo_shape
JSON	testo	oggetto

Ti consigliamo di configurare la coda di lettere morte (DLQ) nella pipeline di Ingestion. OpenSearch Se hai configurato la coda, OpenSearch Service invia tutti i documenti non riusciti che non possono essere importati a causa di errori di mappatura dinamica sulla coda.

Se le mappature automatiche falliscono, puoi utilizzare `template_type` e `template_content` nella configurazione della pipeline per definire regole di mappatura esplicite. In alternativa, puoi creare modelli di mappatura direttamente nel tuo dominio di ricerca o nella tua raccolta prima di avviare la pipeline.

Limitazioni

Considerate le seguenti limitazioni quando impostate una pipeline di OpenSearch ingestione per RDS for MySQL:

- L'integrazione supporta solo un database MySQL per pipeline.
- Al momento l'integrazione non supporta l'inserimento di dati tra regioni; l'istanza e il OpenSearch dominio Amazon RDS devono trovarsi nello stesso posto. Regione AWS
- L'integrazione attualmente non supporta l'ingestione di dati tra account; l'istanza Amazon RDS e la pipeline di OpenSearch Ingestion devono trovarsi nella stessa posizione. Account AWS
- Assicurati che l'istanza Amazon RDS abbia l'autenticazione abilitata utilizzando Secrets Manager, che è l'unico meccanismo di autenticazione supportato.
- La configurazione della pipeline esistente non può essere aggiornata per inserire dati da un database and/or diverso in una tabella diversa. Per aggiornare il nome del database e/o della tabella di una pipeline, è necessario creare una nuova pipeline.
- Le istruzioni DDL (Data Definition Language) in genere non sono supportate. La coerenza dei dati non verrà mantenuta se:
  - Le chiavi primarie vengono modificate (add/delete/rename).
  - Le tabelle vengono eliminate/troncate.
  - I nomi delle colonne o i tipi di dati vengono modificati.
- Se le tabelle MySQL da sincronizzare non hanno chiavi primarie definite, la coerenza dei dati non è garantita. Dovrai definire correttamente `document_id` l'opzione personalizzata nella configurazione del OpenSearch sink per poterla updates/deletes sincronizzare con. OpenSearch
- I riferimenti a chiavi esterne con azioni di eliminazione a cascata non sono supportati e possono causare incongruenze dei dati tra RDS for MySQL e. OpenSearch
- I cluster DB con zone di disponibilità multipla di Amazon RDS non sono supportati.
- Versioni supportate: MySQL versione 8.0 e successive.

CloudWatch Allarmi consigliati

Le seguenti CloudWatch metriche sono consigliate per monitorare le prestazioni della pipeline di ingestione. Queste metriche possono aiutarti a identificare la quantità di dati elaborati dalle esportazioni, il numero di eventi elaborati dagli stream, gli errori nell'elaborazione delle esportazioni e degli eventi di flusso e il numero di documenti scritti nella destinazione. È possibile impostare CloudWatch allarmi per eseguire un'azione quando una di queste metriche supera un valore specificato per un determinato periodo di tempo.

Metrica	Description
pipeline-name .rds.credenziali modificate	Questa metrica indica la frequenza con cui i segreti vengono ruotati. AWS
pipeline-name .rds.executorRefreshErrors	Questa metrica indica gli errori nell'aggiornamento AWS dei segreti.
pipeline-name .rds.exportRecordsTotal	Questa metrica indica il numero di record esportati da Amazon Aurora.
pipeline-name .rds.exportRecordsProcessed	Questa metrica indica il numero di record elaborati dalla pipeline di OpenSearch ingestione.
pipeline-name .rds.exportRecordProcessingErrori	Questa metrica indica il numero di errori di elaborazione in una pipeline di OpenSearch ingestione durante la lettura dei dati da un cluster Amazon Aurora.
pipeline-name .rds.exportRecordsSuccessTotale	Questa metrica indica il numero totale di record di esportazione elaborati correttamente.

Metrica	Description
pipeline-name .rds. exportRecordsFailedTotale	Questa metrica indica il numero totale di record di esportazione che non sono stati elaborati.
pipeline-name .RDS.byte ricevuti	Questa metrica indica il numero totale di byte ricevuti da una pipeline di ingestione. OpenSearch
pipeline-name .rds.bytes processati	Questa metrica indica il numero totale di byte elaborati da una pipeline di ingestione. OpenSearch
pipeline-name .rds. streamRecordsSuccessTotale	Questa metrica indica il numero di record elaborati con successo dallo stream.
pipeline-name .rds. streamRecordsFailedTotale	Questa metrica indica il numero totale di record non elaborati dallo stream.

RDS per PostgreSQL

Completa i seguenti passaggi per configurare una pipeline di OpenSearch ingestione con Amazon RDS for RDS per PostgreSQL.

Argomenti

- Prerequisiti RDS per PostgreSQL
- Fase 1: Configurare il ruolo della pipeline
- Fase 2: Creare la pipeline
- Coerenza dei dati
- Mappatura dei tipi di dati

- Limitazioni
- CloudWatch Allarmi consigliati

Prerequisiti RDS per PostgreSQL

Prima di creare la pipeline di OpenSearch Ingestion, effettuate le seguenti operazioni:

1. Crea un gruppo di parametri DB personalizzato in Amazon RDS per configurare la replica logica.

   ```
   rds.logical_replication=1
   ```

   Per ulteriori informazioni, consulta Esecuzione della replica logica per Amazon RDS for PostgreSQL.
2. Seleziona o crea un'istanza DB RDS per PostgreSQL e associa il gruppo di parametri creato nel passaggio 1 all'istanza DB.
3. Configura l'autenticazione di nome utente e password sulla tua istanza Amazon RDS utilizzando la gestione delle password con Gestione dei segreti AWS Aurora e. Puoi anche creare una username/password combinazione creando un segreto di Secrets Manager.
4. Se utilizzi la funzionalità di snapshot iniziale completa, crea un ruolo AWS KMS key e un ruolo IAM per l'esportazione dei dati da Amazon RDS ad Amazon S3.

   Il ruolo IAM deve avere la seguente politica di autorizzazione:

   JSON

   ```
   {
       "Version":"2012-10-17",
       "Statement": [
           {
               "Sid": "ExportPolicy",
               "Effect": "Allow",
               "Action": [
                   "s3:PutObject*",
                   "s3:ListBucket",
                   "s3:GetObject*",
                   "s3:DeleteObject*",
                   "s3:GetBucketLocation"
               ],
               "Resource": [
                   "arn:aws:s3:::s3-bucket-used-in-pipeline",
   ```

```
                "arn:aws:s3:::s3-bucket-used-in-pipeline/*"
            ]
        }
    ]
}
```

Il ruolo dovrebbe inoltre avere le seguenti relazioni di fiducia:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "export.rds.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

5. Seleziona o crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless. Per ulteriori informazioni, consulta Creazione di domini OpenSearch di servizio e Creazione di raccolte.
6. Allega una politica basata sulle risorse al tuo dominio o una politica di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dalla tua istanza DB Amazon RDS al tuo dominio o alla tua raccolta.

Fase 1: Configurare il ruolo della pipeline

Dopo aver configurato i prerequisiti della pipeline Amazon RDS, configura il ruolo della pipeline da utilizzare nella configurazione della pipeline. Aggiungi anche le seguenti autorizzazioni per il codice sorgente Amazon RDS al ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
```

```
    {
    "Sid": "allowReadingFromS3Buckets",
    "Effect": "Allow",
    "Action": [
    "s3:GetObject",
    "s3:DeleteObject",
    "s3:GetBucketLocation",
    "s3:ListBucket",
    "s3:PutObject"
    ],
    "Resource": [
    "arn:aws:s3:::s3_bucket",
    "arn:aws:s3:::s3_bucket/*"
    ]
    },
    {
    "Sid": "allowNetworkInterfacesActions",
    "Effect": "Allow",
    "Action": [
    "ec2:AttachNetworkInterface",
    "ec2:CreateNetworkInterface",
    "ec2:CreateNetworkInterfacePermission",
    "ec2:DeleteNetworkInterface",
    "ec2:DeleteNetworkInterfacePermission",
    "ec2:DetachNetworkInterface",
    "ec2:DescribeNetworkInterfaces"
    ],
    "Resource": [
    "arn:aws:ec2:*:111122223333:network-interface/*",
    "arn:aws:ec2:*:111122223333:subnet/*",
    "arn:aws:ec2:*:111122223333:security-group/*"
    ]
    },
    {
    "Sid": "allowDescribeEC2",
    "Effect": "Allow",
    "Action": [
    "ec2:Describe*"
    ],
    "Resource": "*"
    },
    {
    "Sid": "allowTagCreation",
    "Effect": "Allow",
```

```
    "Action": [
    "ec2:CreateTags"
    ],
    "Resource": "arn:aws:ec2:*:111122223333:network-interface/*",
    "Condition": {
    "StringEquals": {
    "aws:RequestTag/OSISManaged": "true"
    }
    }
    },
    {
    "Sid": "AllowDescribeInstances",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBInstances"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:db:*"
    ]
    },
    {
    "Sid": "AllowSnapshots",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeDBSnapshots",
    "rds:CreateDBSnapshot",
    "rds:AddTagsToResource"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:db:DB-id",
    "arn:aws:rds:us-east-2:111122223333:snapshot:DB-id*"
    ]
    },
    {
    "Sid": "AllowExport",
    "Effect": "Allow",
    "Action": [
    "rds:StartExportTask"
    ],
    "Resource": [
    "arn:aws:rds:us-east-2:111122223333:snapshot:DB-id*"
    ]
    },
    {
```

```
    "Sid": "AllowDescribeExports",
    "Effect": "Allow",
    "Action": [
    "rds:DescribeExportTasks"
    ],
    "Resource": "*",
    "Condition": {
    "StringEquals": {
    "aws:RequestedRegion": "us-east-2",
    "aws:ResourceAccount": "111122223333"
    }
    }
    },
    {
    "Sid": "AllowAccessToKmsForExport",
    "Effect": "Allow",
    "Action": [
    "kms:Decrypt",
    "kms:Encrypt",
    "kms:DescribeKey",
    "kms:RetireGrant",
    "kms:CreateGrant",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*"
    ],
    "Resource": [
    "arn:aws:kms:us-east-2:111122223333:key/export-key-id"
    ]
    },
    {
    "Sid": "AllowPassingExportRole",
    "Effect": "Allow",
    "Action": "iam:PassRole",
    "Resource": [
    "arn:aws:iam::111122223333:role/export-role"
    ]
    },
    {
    "Sid": "SecretsManagerReadAccess",
    "Effect": "Allow",
    "Action": [
    "secretsmanager:GetSecretValue"
    ],
    "Resource": [
```

```
      "arn:aws:secretsmanager:*:111122223333:secret:*"
      ]
      }
      ]
      }
```

Fase 2: Creare la pipeline

Configura una pipeline di OpenSearch ingestione come la seguente, che specifica un'istanza RDS per PostgreSQL come origine.

```
version: "2"
rds-postgres-pipeline:
  source:
    rds:
      db_identifier: "instance-id"
      engine: postgresql
      database: "database-name"
      tables:
        include:
          - "schema1.table1"
          - "schema2.table2"
      s3_bucket: "bucket-name"
      s3_region: "bucket-region"
      s3_prefix: "prefix-name"
      export:
        kms_key_id: "kms-key-id"
        iam_role_arn: "export-role-arn"
      stream: true
      aws:
        sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        region: "us-east-1"
      authentication:
        username: ${{aws_secrets:secret:username}}
        password: ${{aws_secrets:secret:password}}
  sink:
    - opensearch:
        hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
        index: "${getMetadata(\"table_name\")}"
        index_type: custom
        document_id: "${getMetadata(\"primary_key\")}"
        action: "${getMetadata(\"opensearch_action\")}"
```

```
        document_version: "${getMetadata(\"document_version\")}"
        document_version_type: "external"
        aws:
          sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
          region: "us-east-1"
extension:
  aws:
    secrets:
      secret:
        secret_id: "rds-secret-id"
        region: "us-east-1"
        sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        refresh_interval: PT1H
```

> **Note**
>
> Puoi utilizzare un blueprint Amazon RDS preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta Lavorare con i progetti.

Per utilizzare Amazon Aurora come sorgente, devi configurare l'accesso VPC per la pipeline. Il VPC scelto deve essere lo stesso VPC utilizzato dalla sorgente Amazon Aurora. Quindi scegli una o più sottoreti e uno o più gruppi di sicurezza VPC. Tieni presente che la pipeline richiede l'accesso di rete a un database Aurora MySQL, quindi dovresti anche verificare che il cluster Aurora sia configurato con un gruppo di sicurezza VPC che consenta il traffico in entrata dal gruppo di sicurezza VPC della pipeline alla porta del database. Per ulteriori informazioni, consulta Controllo dell'accesso con gruppi di sicurezza.

Se utilizzi il Console di gestione AWS per creare la tua pipeline, devi anche collegare la pipeline al tuo VPC per utilizzare Amazon Aurora come sorgente. Per fare ciò, trova la sezione Configurazione di rete, scegli Collega a VPC e scegli il tuo CIDR da una delle opzioni predefinite fornite, oppure seleziona la tua. È possibile utilizzare qualsiasi CIDR da uno spazio di indirizzi privato come definito nella Best Current Practice della RFC 1918.

Per fornire un CIDR personalizzato, seleziona Altro dal menu a discesa. Per evitare una collisione tra gli indirizzi IP tra OpenSearch Ingestion e Amazon RDS, assicurati che il CIDR VPC di Amazon Aurora sia diverso dal CIDR per Ingestion. OpenSearch

Per ulteriori informazioni, consulta Configurazione dell'accesso VPC per una pipeline.

Coerenza dei dati

La pipeline garantisce la coerenza dei dati effettuando continuamente polling o ricevendo modifiche dall'istanza Amazon RDS e aggiornando i documenti corrispondenti nell'indice. OpenSearch

OpenSearch Ingestion supporta il end-to-end riconoscimento per garantire la durabilità dei dati. Quando una pipeline legge istantanee o flussi, crea dinamicamente partizioni per l'elaborazione parallela. La pipeline contrassegna una partizione come completa quando riceve un riconoscimento dopo aver acquisito tutti i record nel dominio o nella raccolta. OpenSearch Se desideri importare in una raccolta di ricerca OpenSearch Serverless, puoi generare un ID di documento nella pipeline. Se desideri importare in una raccolta di serie temporali OpenSearch Serverless, tieni presente che la pipeline non genera un ID di documento, quindi devi `document_id: "${getMetadata(\"primary_key\")}"` ometterlo nella configurazione del pipeline sink.

Una pipeline OpenSearch di Ingestion mappa anche le azioni degli eventi in entrata nelle corrispondenti azioni di indicizzazione di massa per facilitare l'acquisizione dei documenti. Ciò mantiene i dati coerenti, in modo che ogni modifica dei dati in Amazon RDS venga riconciliata con le corrispondenti modifiche del documento. OpenSearch

Mappatura dei tipi di dati

OpenSearch La pipeline di ingestione mappa i tipi di dati PostgreSQL su rappresentazioni adatte all'uso da parte di domini o raccolte di servizi. OpenSearch Se non è definito alcun modello di mappatura OpenSearch, determina OpenSearch automaticamente i tipi di campo con una mappatura dinamica basata sul primo documento inviato. Puoi anche definire in modo esplicito i tipi di campo più adatti alle tue esigenze OpenSearch tramite un modello di mappatura.

La tabella seguente elenca i tipi di dati RDS per PostgreSQL e i tipi di campo corrispondenti. OpenSearch La colonna Tipo di OpenSearch campo predefinito mostra il tipo di campo corrispondente OpenSearch se non è definita alcuna mappatura esplicita. In questo caso, determina OpenSearch automaticamente i tipi di campo con mappatura dinamica. La colonna Tipo di OpenSearch campo consigliato è il tipo di campo consigliato corrispondente da specificare esplicitamente in un modello di mappatura. Questi tipi di campo sono più allineati ai tipi di dati di RDS per PostgreSQL e in genere possono abilitare migliori funzionalità di ricerca disponibili in. OpenSearch

Tipo di dati RDS per PostgreSQL	Tipo OpenSearch di campo predefinito	Tipo di OpenSearch campo consigliato
smallint	Long	short
intero	Long	intero
bigint	Long	Long
decimal	testo	doppio o parola chiave
numerico [(p, s)]	testo	doppio o parola chiave
real	virgola mobile	virgola mobile
double precision	virgola mobile	virgola mobile a doppia precisione
smallserial	Long	short
seriale	Long	intero
bigserial	Long	Long
money	oggetto	oggetto
carattere variabile (n)	testo	testo
varchar(n)	testo	testo
carattere (n)	testo	testo

Tipo di dati RDS per PostgreSQL	Tipo OpenSearch di campo predefinito	Tipo di OpenSearch campo consigliato
char(n)	testo	testo
bpchar (n)	testo	testo
bpchar	testo	testo
testo	testo	testo
enum	testo	testo
bytea	testo	binary
timestamp [(p)] [senza fuso orario]	long (in millisecondi di epoca)	data
timestamp [(p)] con fuso orario	long (in millisecondi di epoca)	data
data	long (in millisecondi epocali)	data
ora [(p)] [senza fuso orario]	long (in millisecondi di epoca)	data
ora [(p)] con fuso orario	lungo (in millisecondi di epoca)	data

Tipo di dati RDS per PostgreSQL	Tipo OpenSearch di campo predefinito	Tipo di OpenSearch campo consigliato
intervallo [campi] [(p)]	testo (ISO8601 formato)	testo
booleano	Booleano	Booleano
point	testo (in formato WKT)	geo_shape
linea	testo (in formato WKT)	geo_shape
gamba	testo (in formato WKT)	geo_shape
scatola	testo (in formato WKT)	geo_shape
path	testo (in formato WKT)	geo_shape
poligono	testo (in formato WKT)	geo_shape
cerchio	oggetto	oggetto
cidr	testo	testo
inet	testo	testo
macaddr	testo	testo
macaddr8	testo	testo
bit(n)	Long	byte, corto, intero o lungo (a seconda del numero di bit)
bit varying(n)	Long	byte, corto, intero o lungo (a seconda del numero di bit)
json	oggetto	oggetto

Tipo di dati RDS per PostgreSQL	Tipo OpenSearch di campo predefinito	Tipo di OpenSearch campo consigliato
jsonb	oggetto	oggetto
jsonpath	testo	testo

Ti consigliamo di configurare la dead-letter queue (DLQ) nella pipeline di Ingestion. OpenSearch Se hai configurato la coda, OpenSearch Service invia tutti i documenti non riusciti che non possono essere importati a causa di errori di mappatura dinamica sulla coda.

Nel caso in cui le mappature automatiche falliscano, puoi utilizzare `template_type` e `template_content` nella configurazione della pipeline per definire regole di mappatura esplicite. In alternativa, puoi creare modelli di mappatura direttamente nel tuo dominio di ricerca o nella tua raccolta prima di avviare la pipeline.

Limitazioni

Considerate le seguenti limitazioni quando impostate una pipeline di OpenSearch ingestione per RDS per PostgreSQL:

- L'integrazione supporta solo un database PostgreSQL per pipeline.
- Al momento l'integrazione non supporta l'inserimento di dati tra regioni; l'istanza e il OpenSearch dominio Amazon RDS devono trovarsi nello stesso posto. Regione AWS
- L'integrazione attualmente non supporta l'ingestione di dati tra account; l'istanza Amazon RDS e la pipeline di OpenSearch Ingestion devono trovarsi nella stessa posizione. Account AWS
- Assicurati che sull'istanza Amazon RDS sia abilitata l'autenticazione Gestione dei segreti AWS, che è l'unico meccanismo di autenticazione supportato.
- La configurazione della pipeline esistente non può essere aggiornata per inserire dati da un database and/or diverso in una tabella diversa. Per aggiornare il nome del database e/o della tabella di una pipeline, è necessario arrestare la pipeline e riavviarla con una configurazione aggiornata oppure creare una nuova pipeline.
- Le istruzioni DDL (Data Definition Language) in genere non sono supportate. La coerenza dei dati non verrà mantenuta se:

  - Le chiavi primarie vengono modificate (add/delete/rename).
  - Le tabelle vengono eliminate/troncate.
  - I nomi delle colonne o i tipi di dati vengono modificati.
- Se le tabelle PostgreSQL da sincronizzare non hanno chiavi primarie definite, la coerenza dei dati non è garantita. Dovrai definire, personalizzare l'`document_id`opzione OpenSearch e configurare correttamente il sink per poterti sincronizzare con. updates/deletes OpenSearch
- I cluster DB RDS Multi-AZ non sono supportati.
- Versioni supportate: PostgreSQL 16 e versioni successive.

CloudWatch Allarmi consigliati

Le seguenti CloudWatch metriche sono consigliate per monitorare le prestazioni della pipeline di ingestione. Queste metriche possono aiutarti a identificare la quantità di dati elaborati dalle esportazioni, il numero di eventi elaborati dagli stream, gli errori nell'elaborazione delle esportazioni e degli eventi di flusso e il numero di documenti scritti nella destinazione. È possibile impostare CloudWatch allarmi per eseguire un'azione quando una di queste metriche supera un valore specificato per un determinato periodo di tempo.

Metrica	Description
pipeline-name .rds.credenziali modificate	Questa metrica indica la frequenza con cui i segreti vengono ruotati. AWS
pipeline-name .rds.executorRefreshErrors	Questa metrica indica gli errori nell'aggiornamento AWS dei segreti.
pipeline-name .rds.exportRecordsTotal	Questa metrica indica il numero di record esportati da Amazon Aurora.
pipeline-name .rds.	Questa metrica indica il numero di record elaborati dalla pipeline di OpenSearch ingestione.

Metrica	Description
exportRecordsProcessed	
pipeline-name .rds.exportRecordProcessingErrori	Questa metrica indica il numero di errori di elaborazione in una pipeline di OpenSearch ingestione durante la lettura dei dati da un cluster Amazon Aurora.
pipeline-name .rds.exportRecordsSuccessTotale	Questa metrica indica il numero totale di record di esportazione elaborati correttamente.
pipeline-name .rds.exportRecordsFailedTotale	Questa metrica indica il numero totale di record di esportazione che non sono stati elaborati.
pipeline-name .RDS.byte ricevuti	Questa metrica indica il numero totale di byte ricevuti da una pipeline di ingestione. OpenSearch
pipeline-name .rds.bytes processati	Questa metrica indica il numero totale di byte elaborati da una pipeline di ingestione. OpenSearch
pipeline-name .rds.streamRecordsSuccessTotale	Questa metrica indica il numero di record elaborati con successo dallo stream.
pipeline-name .rds.streamRecordsFailedTotale	Questa metrica indica il numero totale di record non elaborati dallo stream.

Utilizzo di una pipeline OpenSearch di ingestione con Amazon S3

Con OpenSearch Ingestion, puoi usare Amazon S3 come origine o come destinazione. Quando usi Amazon S3 come fonte, invii dati a una pipeline di OpenSearch ingestione. Quando usi Amazon S3 come destinazione, scrivi dati da una pipeline di importazione OpenSearch su uno o più bucket S3.

Argomenti

- Amazon S3 come fonte
- Amazon S3 come destinazione
- Account multiplo Amazon S3 come fonte

Amazon S3 come fonte

Esistono due modi per utilizzare Amazon S3 come origine per elaborare i dati: con l'elaborazione S3-SQS e con le scansioni pianificate.

Utilizza l'elaborazione S3-SQS quando è necessaria una scansione quasi in tempo reale dei file dopo la scrittura su S3. Puoi configurare i bucket Amazon S3 per generare un evento ogni volta che un oggetto viene archiviato o modificato all'interno del bucket. Utilizza una scansione pianificata una tantum o ricorrente per elaborare in batch i dati in un bucket S3.

Argomenti

- Prerequisiti
- Fase 1: Configurare il ruolo della pipeline
- Fase 2: Creare la pipeline

Prerequisiti

Per utilizzare Amazon S3 come origine per una pipeline di importazione sia per una scansione pianificata che per l' OpenSearch elaborazione S3-SQS, devi prima creare un bucket S3.

> **Note**
>
> Se il bucket S3 utilizzato come origine nella pipeline di OpenSearch Ingestion si trova in un altro, devi anche abilitare le autorizzazioni di lettura tra account sul bucket. Account AWS Ciò consente alla pipeline di leggere ed elaborare i dati. Per abilitare le autorizzazioni per più

account, consulta la sezione relativa alla concessione delle autorizzazioni per i bucket tra più account nella Guida per l'utente di Amazon S3.
Se i tuoi bucket S3 si trovano in più account, usa una mappa. `bucket_owners` Per un esempio, vedi Cross-account S3 access nella documentazione. OpenSearch

Per configurare l'elaborazione S3-SQS, è inoltre necessario eseguire le seguenti operazioni:

1. Crea una coda Amazon SQS.
2. Abilita le notifiche degli eventi sul bucket S3 con la coda SQS come destinazione.

Fase 1: Configurare il ruolo della pipeline

A differenza di altri plugin di origine che inviano dati a una pipeline, il plug-in di origine S3 ha un'architettura basata sulla lettura in cui la pipeline estrae i dati dalla sorgente.

Pertanto, affinché una pipeline possa leggere da S3, è necessario specificare un ruolo all'interno della configurazione di origine S3 della pipeline che abbia accesso sia al bucket S3 che alla coda Amazon SQS. La pipeline assumerà questo ruolo per leggere i dati dalla coda.

> **Note**
>
> Il ruolo specificato nella configurazione di origine di S3 deve essere il ruolo della pipeline.
> Pertanto, il ruolo della pipeline deve contenere due policy di autorizzazione separate, una per la scrittura su un sink e l'altra da estrarre dal codice sorgente S3. È necessario utilizzare lo stesso in tutti i componenti della pipeline. `sts_role_arn`

La seguente politica di esempio mostra le autorizzazioni necessarie per l'utilizzo di S3 come fonte:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action":[
```

```
            "s3:ListBucket",
            "s3:GetBucketLocation",
            "s3:GetObject"
         ],
        "Resource": "arn:aws:s3:::amzn-s3-demo-bucket/*"
      },
      {
         "Effect":"Allow",
         "Action":"s3:ListAllMyBuckets",
         "Resource":"arn:aws:s3:::*"
      },
      {
        "Effect": "Allow",
        "Action": [
          "sqs:DeleteMessage",
          "sqs:ReceiveMessage",
          "sqs:ChangeMessageVisibility"
        ],
        "Resource": "arn:aws:sqs:us-east-1:111122223333:MyS3EventSqsQueue"
      }
    ]
}
```

È necessario associare queste autorizzazioni al ruolo IAM specificato nell'`sts_role_arn`opzione all'interno della configurazione del plug-in di origine S3:

```
version: "2"
source:
  s3:
    ...
    aws:
      ...
processor:
  ...
sink:
  - opensearch:
      ...
```

Fase 2: Creare la pipeline

Dopo aver impostato le autorizzazioni, puoi configurare una pipeline di OpenSearch ingestione in base al tuo caso d'uso di Amazon S3.

Elaborazione S3-SQS

Per configurare l'elaborazione S3-SQS, configura la pipeline per specificare S3 come origine e configura le notifiche di Amazon SQS:

```
version: "2"
s3-pipeline:
  source:
    s3:
      notification_type: "sqs"
      codec:
        newline: null
      sqs:
        queue_url: "https://sqs.us-east-1amazonaws.com/account-id/ingestion-queue"
      compression: "none"
      aws:
        region: "region"
  processor:
  - grok:
      match:
        message:
        - "%{COMMONAPACHELOG}"
  - date:
      destination: "@timestamp"
      from_time_received: true
  sink:
  - opensearch:
      hosts: ["https://search-domain-endpoint.us-east-1es.amazonaws.com"]
      index: "index-name"
      aws:
        region: "region"
```

Se riscontri un basso utilizzo della CPU durante l'elaborazione di file di piccole dimensioni su Amazon S3, valuta la possibilità di aumentare la velocità effettiva modificando il valore dell'opzione. `workers` Per ulteriori informazioni, consulta le opzioni di configurazione del plug-in S3.

Scansione pianificata

Per configurare una scansione pianificata, configura la pipeline con una pianificazione a livello di scansione applicabile a tutti i bucket S3 o a livello di bucket. Una pianificazione a livello di bucket o una configurazione a intervalli di scansione sovrascrive sempre una configurazione a livello di scansione.

È possibile configurare le scansioni pianificate con una scansione singola, ideale per la migrazione dei dati, o una scansione ricorrente, ideale per l'elaborazione in batch.

Per configurare la pipeline per la lettura da Amazon S3, utilizza i blueprint Amazon S3 preconfigurati. Puoi modificare la `scan` parte della configurazione della pipeline per soddisfare le tue esigenze di pianificazione. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Scansione una tantum

Una scansione pianificata una tantum viene eseguita una sola volta. Nella configurazione della pipeline, è possibile utilizzare un `start_time` e `end_time` per specificare quando si desidera che gli oggetti nel bucket vengano scansionati. In alternativa, è possibile utilizzare `range` per specificare l'intervallo di tempo relativo all'ora corrente in cui si desidera che gli oggetti nel bucket vengano scansionati.

Ad esempio, un intervallo impostato per PT4H scansionare tutti i file creati nelle ultime quattro ore. Per configurare una scansione singola da eseguire una seconda volta, è necessario arrestare e riavviare la pipeline. Se non hai configurato un intervallo, devi anche aggiornare l'ora di inizio e di fine.

La seguente configurazione imposta una scansione unica per tutti i bucket e tutti gli oggetti in tali bucket:

```
version: "2"
log-pipeline:
  source:
    s3:
      codec:
        csv:
      compression: "none"
      aws:
        region: "region"
      acknowledgments: true
      scan:
        buckets:
          - bucket:
              name: my-bucket
              filter:
                include_prefix:
                  - Objects1/
                exclude_suffix:
                  - .jpeg
```

```
                - .png
        - bucket:
            name: my-bucket-2
            key_prefix:
              include:
                - Objects2/
              exclude_suffix:
                - .jpeg
                - .png
    delete_s3_objects_on_read: false
  processor:
    - date:
        destination: "@timestamp"
        from_time_received: true
  sink:
    - opensearch:
        hosts: ["https://search-domain-endpoint.us-east-1es.amazonaws.com"]
        index: "index-name"
        aws:
          region: "region"
        dlq:
          s3:
            bucket: "dlq-bucket"
            region: "us-east-1"
```

La seguente configurazione imposta una scansione unica per tutti i bucket durante una finestra temporale specificata. Ciò significa che S3 elabora solo gli oggetti con tempi di creazione che rientrano in questa finestra.

```
scan:
  start_time: 2023-01-21T18:00:00.000Z
  end_time: 2023-04-21T18:00:00.000Z
  buckets:
    - bucket:
        name: my-bucket-1
        filter:
          include:
            - Objects1/
          exclude_suffix:
            - .jpeg
            - .png
    - bucket:
        name: my-bucket-2
```

```
        filter:
          include:
            - Objects2/
          exclude_suffix:
            - .jpeg
            - .png
```

La seguente configurazione imposta una scansione una tantum sia a livello di scansione che a livello di bucket. Gli orari di inizio e fine a livello di bucket sostituiscono gli orari di inizio e fine a livello di scansione.

```
scan:
  start_time: 2023-01-21T18:00:00.000Z
  end_time: 2023-04-21T18:00:00.000Z
  buckets:
    - bucket:
        start_time: 2023-01-21T18:00:00.000Z
        end_time: 2023-04-21T18:00:00.000Z
        name: my-bucket-1
        filter:
          include:
            - Objects1/
          exclude_suffix:
            - .jpeg
            - .png
    - bucket:
        start_time: 2023-01-21T18:00:00.000Z
        end_time: 2023-04-21T18:00:00.000Z
        name: my-bucket-2
        filter:
          include:
            - Objects2/
          exclude_suffix:
            - .jpeg
            - .png
```

L'arresto di una pipeline rimuove qualsiasi riferimento preesistente a quali oggetti sono stati scansionati dalla pipeline prima dell'arresto. Se una singola pipeline di scansione viene interrotta, eseguirà nuovamente la scansione di tutti gli oggetti dopo l'avvio, anche se erano già stati scansionati. Se è necessario interrompere una singola pipeline di scansione, si consiglia di modificare la finestra temporale prima di riavviare la pipeline.

Se è necessario filtrare gli oggetti in base all'ora di inizio e all'ora di fine, l'unica opzione è interrompere e avviare la pipeline. Se non è necessario filtrare per ora di inizio e ora di fine, è possibile filtrare gli oggetti per nome. Il filtraggio per nome non richiede di interrompere e avviare la pipeline. Per fare ciò, usa e. `include_prefix` `exclude_suffix`

Scansione ricorrente

Una scansione pianificata ricorrente esegue una scansione dei bucket S3 specificati a intervalli regolari e pianificati. Puoi configurare questi intervalli solo a livello di scansione perché le configurazioni a livello di bucket individuali non sono supportate.

Nella configurazione della pipeline, `interval` specifica la frequenza della scansione ricorrente e può essere compresa tra 30 secondi e 365 giorni. La prima di queste scansioni si verifica sempre quando si crea la pipeline. `count`definisce il numero totale di istanze di scansione.

La seguente configurazione imposta una scansione ricorrente, con un ritardo di 12 ore tra le scansioni:

```
scan:
  scheduling:
    interval: PT12H
    count: 4
  buckets:
    - bucket:
        name: my-bucket-1
        filter:
          include:
            - Objects1/
          exclude_suffix:
            - .jpeg
            - .png
    - bucket:
        name: my-bucket-2
        filter:
          include:
            - Objects2/
          exclude_suffix:
            - .jpeg
            - .png
```

Amazon S3 come destinazione

Per scrivere dati da una pipeline di OpenSearch ingestione a un bucket S3, usa il blueprint S3 preconfigurato per creare una pipeline con un sink S3. Questa pipeline indirizza i dati selettivi verso un sink e invia simultaneamente tutti i dati per l'archiviazione in S3. OpenSearch Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Quando crei il tuo sink S3, puoi specificare la formattazione preferita tra una varietà di codec sink. Ad esempio, se desideri scrivere dati in formato colonnare, scegli il codec Parquet o Avro. Se preferisci un formato basato su righe, scegli JSON o NDJSON. Per scrivere dati su S3 in uno schema specificato, puoi anche definire uno schema in linea all'interno dei codec sink usando il formato Avro.

L'esempio seguente definisce uno schema in linea in un sink S3:

```
- s3:
  codec:
    parquet:
      schema: >
        {
           "type" : "record",
           "namespace" : "org.vpcFlowLog.examples",
           "name" : "VpcFlowLog",
           "fields" : [
             { "name" : "version", "type" : "string"},
             { "name" : "srcport", "type": "int"},
             { "name" : "dstport", "type": "int"},
             { "name" : "start", "type": "int"},
             { "name" : "end", "type": "int"},
             { "name" : "protocol", "type": "int"},
             { "name" : "packets", "type": "int"},
             { "name" : "bytes", "type": "int"},
             { "name" : "action", "type": "string"},
             { "name" : "logStatus", "type" : "string"}
           ]
         }
```

Quando definisci questo schema, specifica un superset di tutte le chiavi che potrebbero essere presenti nei diversi tipi di eventi che la pipeline invia a un sink.

Ad esempio, se in un evento è possibile che manchi una chiave, aggiungete quella chiave allo schema con un `null` valore. Le dichiarazioni di valori nulli consentono allo schema di elaborare dati

non uniformi (laddove alcuni eventi abbiano queste chiavi e altri no). Quando negli eventi in entrata sono presenti queste chiavi, i relativi valori vengono scritti nei sink.

Questa definizione dello schema funge da filtro che consente solo l'invio di chiavi definite ai sink e rimuove le chiavi non definite dagli eventi in arrivo.

Puoi anche utilizzare `include_keys` e `exclude_keys` nel tuo sink per filtrare i dati che vengono indirizzati ad altri sink. Questi due filtri si escludono a vicenda, quindi puoi utilizzarne solo uno alla volta nello schema. Inoltre, non è possibile utilizzarli all'interno di schemi definiti dall'utente.

Per creare pipeline con tali filtri, utilizzate il blueprint del filtro sink preconfigurato. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Account multiplo Amazon S3 come fonte

Puoi concedere l'accesso a più account con Amazon S3 in modo che le pipeline di OpenSearch Ingestion possano accedere ai bucket S3 in un altro account come fonte. Per abilitare l'accesso su più account, consulta la sezione relativa alla concessione delle autorizzazioni per i bucket tra più account nella Amazon S3 User Guide. Dopo aver concesso l'accesso, assicurati che il tuo ruolo di pipeline disponga delle autorizzazioni richieste.

Quindi, puoi creare una pipeline utilizzando `bucket_owners` per abilitare l'accesso tra account a un bucket Amazon S3 come origine:

```
s3-pipeline:
 source:
  s3:
   notification_type: "sqs"
   codec:
    csv:
     delimiter: ","
     quote_character: "\""
     detect_header: True
   sqs:
    queue_url: "https://sqs.ap-northeast-1.amazonaws.com/401447383613/test-s3-queue"
   bucket_owners:
    my-bucket-01: 123456789012
    my-bucket-02: 999999999999
   compression: "gzip"
```

Utilizzo di una pipeline OpenSearch di ingestione con Amazon Security Lake

Puoi utilizzare il plug-in sorgente S3 per importare dati da Amazon Security Lake nella tua pipeline di OpenSearch ingestione. Security Lake centralizza automaticamente i dati di sicurezza provenienti da AWS ambienti, ambienti locali e provider SaaS in un data lake appositamente progettato. È possibile creare un abbonamento che replica i dati da Security Lake alla pipeline di OpenSearch Ingestion, che poi li scrive nel dominio di servizio o nella raccolta Serverless. OpenSearch OpenSearch

Per configurare la pipeline in modo che legga da Security Lake, utilizza il blueprint Security Lake preconfigurato. Il blueprint include una configurazione predefinita per l'importazione di file parquet Open Cybersecurity Schema Framework (OCSF) da Security Lake. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Argomenti

- Utilizzo di una pipeline OpenSearch di ingestione con Amazon Security Lake come fonte
- Utilizzo di una pipeline OpenSearch di ingestione con Amazon Security Lake come sink

Utilizzo di una pipeline OpenSearch di ingestione con Amazon Security Lake come fonte

Puoi utilizzare il plug-in sorgente Amazon S3 all'interno della tua pipeline OpenSearch Ingestion per importare dati da Amazon Security Lake. Security Lake centralizza automaticamente i dati di sicurezza provenienti da AWS ambienti, sistemi locali e provider SaaS in un data lake appositamente progettato.

Amazon Security Lake ha i seguenti attributi di metadati all'interno di una pipeline:

- `bucket_name`: il nome del bucket Amazon S3 creato da Security Lake per l'archiviazione dei dati di sicurezza.
- `path_prefix`: il nome sorgente personalizzato definito nella policy del ruolo IAM di Security Lake.
- `region`: La posizione Regione AWS in cui si trova il bucket Security Lake S3.
- `accountID`: L' Account AWS ID in cui è abilitato Security Lake.
- `sts_role_arn`: L'ARN del ruolo IAM destinato all'uso con Security Lake.

Prerequisiti

Prima di creare la pipeline OpenSearch di Ingestion, effettuate le seguenti operazioni:

- Abilita Security Lake.
- Crea un abbonato in Security Lake.
  - Scegli le fonti che desideri inserire nella tua pipeline.
  - Per le credenziali dell'abbonato, aggiungi l'ID del Account AWS luogo in cui intendi creare la pipeline. Per l'ID esterno, specificare. `OpenSearchIngestion-{accountid}`
  - Per il metodo di accesso ai dati, scegli S3.
  - Per i dettagli della notifica, scegli SQS queue.

Quando crei un abbonato, Security Lake crea automaticamente due policy di autorizzazione in linea, una per S3 e una per SQS. Le politiche hanno il seguente formato: e. `AmazonSecurityLake-amzn-s3-demo-bucket-S3` `AmazonSecurityLake-AWS Demo-SQS` Per consentire alla tua pipeline di accedere alle fonti degli abbonati, devi associare le autorizzazioni richieste al tuo ruolo di pipeline.

Configura il ruolo della pipeline

Crea una nuova politica di autorizzazioni in IAM che combini solo le autorizzazioni richieste delle due policy create automaticamente da Security Lake. La seguente policy di esempio mostra il privilegio minimo richiesto a una pipeline di OpenSearch Ingestion per leggere i dati da più fonti Security Lake:

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Action":[
            "s3:GetObject"
         ],
         "Resource":[
            "arn:aws:s3:::aws-security-data-lake-us-east-1-abcde/aws/
LAMBDA_EXECUTION/1.0/*",
            "arn:aws:s3:::aws-security-data-lake-us-east-1-abcde/aws/S3_DATA/1.0/
*",
```

```
            "arn:aws:s3:::aws-security-data-lake-us-east-1-abcde/aws/
VPC_FLOW/1.0/*",
            "arn:aws:s3:::aws-security-data-lake-us-east-1-abcde/aws/ROUTE53/1.0/
*",
            "arn:aws:s3:::aws-security-data-lake-us-east-1-abcde/aws/
SH_FINDINGS/1.0/*"
         ]
      },
      {
         "Effect":"Allow",
         "Action":[
            "sqs:ReceiveMessage",
            "sqs:DeleteMessage"
         ],
         "Resource":[
            "arn:aws:sqs:us-east-1:111122223333:AmazonSecurityLake-abcde-Main-
Queue"
         ]
      }
   ]
}
```

> **Important**
>
> Security Lake non gestisce al posto tuo la policy relativa ai ruoli della pipeline. Se aggiungi o rimuovi fonti dal tuo abbonamento a Security Lake, devi aggiornare manualmente la politica. Security Lake crea partizioni per ogni origine di registro, quindi è necessario aggiungere o rimuovere manualmente le autorizzazioni nel ruolo della pipeline.

È necessario associare queste autorizzazioni al ruolo IAM specificato nell'`sts_role_arn`opzione all'interno della configurazione del plug-in di origine S3, sotto. `sqs`

```
version: "2"
source:
  s3:
    ...
    sqs:
      queue_url: "https://sqs.us-east-1amazonaws.com/account-id/
AmazonSecurityLake-amzn-s3-demo-bucket-Main-Queue"
    aws:
```

```
      ...
processor:
  ...
sink:
  - opensearch:
      ...
```

Crea la pipeline

Dopo aver aggiunto le autorizzazioni al ruolo della pipeline, utilizza il blueprint Security Lake preconfigurato per creare la pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

È necessario specificare l'`queue_url`opzione all'interno della configurazione di `s3` origine, che è l'URL della coda di Amazon SQS da cui leggere. Per formattare l'URL, individua l'endpoint Subscription nella configurazione del sottoscrittore e passa a. `arn:aws:` `https://` Ad esempio, `https://sqs.`*`us-east-1`*`amazonaws.com/`*`account-id`*`/AmazonSecurityLake-`*`AWS Demo`*`-Main-Queue`.

`sts_role_arn`Quello che specifichi all'interno della configurazione di origine S3 deve essere l'ARN del ruolo della pipeline.

Utilizzo di una pipeline OpenSearch di ingestione con Amazon Security Lake come sink

Utilizza il plug-in sink Amazon S3 in OpenSearch Ingestion per inviare dati da qualsiasi fonte supportata ad Amazon Security Lake. Security Lake raccoglie e archivia i dati di sicurezza provenienti da AWS ambienti locali e provider SaaS in un data lake dedicato.

Per configurare la pipeline per scrivere i dati di registro su Security Lake, utilizza il blueprint preconfigurato Firewall Traffic Logs. Il modello include una configurazione predefinita per il recupero dei log di sicurezza non elaborati o di altri dati archiviati in un bucket Amazon S3, l'elaborazione dei record e la loro normalizzazione. Quindi mappa i dati su Open Cybersecurity Schema Framework (OCSF) e invia i dati trasformati conformi a OCSF a Security Lake.

La pipeline ha i seguenti attributi di metadati:

- `bucket_name`: il nome del bucket Amazon S3 creato da Security Lake per l'archiviazione dei dati di sicurezza.
- `path_prefix`: il nome sorgente personalizzato definito nella policy del ruolo IAM di Security Lake.

- `region`: La posizione Regione AWS in cui si trova il bucket Security Lake S3.
- `accountID`: L' Account AWS ID in cui è abilitato Security Lake.
- `sts_role_arn`: L'ARN del ruolo IAM destinato all'uso con Security Lake.

Prerequisiti

Prima di creare una pipeline per inviare dati a Security Lake, effettuate le seguenti operazioni:

- Abilita e configura Amazon Security Lake: configura Amazon Security Lake per centralizzare i dati di sicurezza da varie fonti. Per istruzioni, consulta Abilitazione di Security Lake tramite la console.

  Quando selezioni una fonte, scegli Inserisci AWS fonti specifiche e seleziona una o più fonti di log ed eventi che desideri importare.
- Configura le autorizzazioni: configura il ruolo della pipeline con le autorizzazioni necessarie per scrivere dati su Security Lake. Per ulteriori informazioni, consulta il ruolo Pipeline.

Crea la pipeline

Usa il blueprint Security Lake preconfigurato per creare la pipeline. Per ulteriori informazioni, vedere Utilizzo dei blueprint per creare una pipeline.

Utilizzo di una pipeline di OpenSearch ingestione con Fluent Bit

Questo file di configurazione Fluent Bit di esempio invia i dati di registro da Fluent Bit a una pipeline di ingestione. OpenSearch Per ulteriori informazioni sull'acquisizione dei dati di registro, consulta Log Analytics nella documentazione di Data Prepper.

Tenere presente quanto segue:

- Il `host` valore deve essere l'endpoint della pipeline. Ad esempio, *`pipeline-endpoint`*`.`*`us-east-1`*`osis.amazonaws.com`.
- Il valore `aws_service` deve essere `osis`.
- Il `aws_role_arn` valore è l'ARN del ruolo AWS IAM che il client deve assumere e utilizzare per l'autenticazione Signature Version 4.

```
[INPUT]
  name                  tail
```

```
  refresh_interval      5
  path                  /var/log/test.log
  read_from_head        true

[OUTPUT]
  Name http
  Match *
  Host pipeline-endpoint.us-east-1osis.amazonaws.com
  Port 443
  URI /log/ingest
  Format json
  aws_auth true
  aws_region region
  aws_service osis
  aws_role_arn arn:aws:iam::account-id:role/ingestion-role
  Log_Level trace
  tls On
```

È quindi possibile configurare una pipeline di OpenSearch ingestion come la seguente, che ha HTTP come origine:

```
version: "2"
unaggregated-log-pipeline:
  source:
    http:
      path: "/log/ingest"
  processor:
    - grok:
        match:
          log:
            - "%{TIMESTAMP_ISO8601:timestamp} %{NOTSPACE:network_node}
 %{NOTSPACE:network_host} %{IPORHOST:source_ip}:%{NUMBER:source_port:int} ->
 %{IPORHOST:destination_ip}:%{NUMBER:destination_port:int} %{GREEDYDATA:details}"
    - grok:
        match:
          details:
            - "'%{NOTSPACE:http_method} %{NOTSPACE:http_uri}' %{NOTSPACE:protocol}"
            - "TLS%{NOTSPACE:tls_version} %{GREEDYDATA:encryption}"
            - "%{NUMBER:status_code:int} %{NUMBER:response_size:int}"
    - delete_entries:
        with_keys: ["details", "log"]

  sink:
```

```
    - opensearch:
        hosts: ["https://search-domain-endpoint.us-east-1es.amazonaws.com"]
        index: "index_name"
        index_type: custom
        bulk_size: 20
        aws:
          region: "region"
```

Utilizzo di una pipeline di OpenSearch ingestione con Fluentd

Fluentd è un ecosistema di raccolta dati open source che fornisce diverse lingue e sottoprogetti come Fluent SDKs Bit. Questo file di configurazione Fluentd di esempio invia i dati di registro da Fluentd a una pipeline di Ingestion. OpenSearch Per ulteriori informazioni sull'acquisizione dei dati di registro, consulta Log Analytics nella documentazione di Data Prepper.

Tenere presente quanto segue:

- Il `endpoint` valore deve essere l'endpoint della pipeline. Ad esempio, *`pipeline-endpoint`*`.us-east-1`osis.amazonaws.com/apache-log-pipeline/logs.
- Il valore `aws_service` deve essere `osis`.
- Il `aws_role_arn` valore è l'ARN del ruolo AWS IAM che il client deve assumere e utilizzare per l'autenticazione Signature Version 4.

```
<source>
  @type tail
  path logs/sample.log
  path_key log
  tag apache
  <parse>
    @type none
  </parse>
</source>

<filter apache>
  @type record_transformer
  <record>
    log ${record["message"]}
  </record>
</filter>
```

```
<filter apache>
  @type record_transformer
  remove_keys message
</filter>

<match apache>
  @type http
  endpoint pipeline-endpoint.us-east-1osis.amazonaws.com/apache-log-pipeline/logs
  json_array true

  <auth>
    method aws_sigv4
    aws_service osis
    aws_region region
    aws_role_arn arn:aws:iam::account-id:role/ingestion-role
  </auth>

  <format>
    @type json
  </format>

  <buffer>
    flush_interval 1s
  </buffer>
</match>
```

È quindi possibile configurare una pipeline di OpenSearch ingestion come la seguente, che ha HTTP come origine:

```
version: "2"
apache-log-pipeline:
  source:
    http:
      path: "/${pipelineName}/logs"
  processor:
    - grok:
        match:
          log:
            - "%{TIMESTAMP_ISO8601:timestamp} %{NOTSPACE:network_node}
 %{NOTSPACE:network_host} %{IPORHOST:source_ip}:%{NUMBER:source_port:int} ->
 %{IPORHOST:destination_ip}:%{NUMBER:destination_port:int} %{GREEDYDATA:details}"
  sink:
    - opensearch:
```

```
        hosts: ["https://search-domain-endpoint.us-east-1es.amazonaws.com"]
        index: "index_name"
        aws_region: "region"
        aws_sigv4: true
```

Utilizzo di una pipeline OpenSearch di ingestione con inferenza batch offline di machine learning

Le pipeline Amazon OpenSearch Ingestion (OSI) supportano l'elaborazione di inferenze in batch offline di machine learning (ML) per arricchire in modo efficiente grandi volumi di dati a basso costo. Usa l'inferenza in batch offline ogni volta che hai set di dati di grandi dimensioni che possono essere elaborati in modo asincrono. L'inferenza batch offline funziona con Amazon Bedrock e SageMaker i modelli. Questa funzionalità è disponibile in tutti i domini Regioni AWS che supportano OpenSearch Ingestion with OpenSearch Service 2.17+.

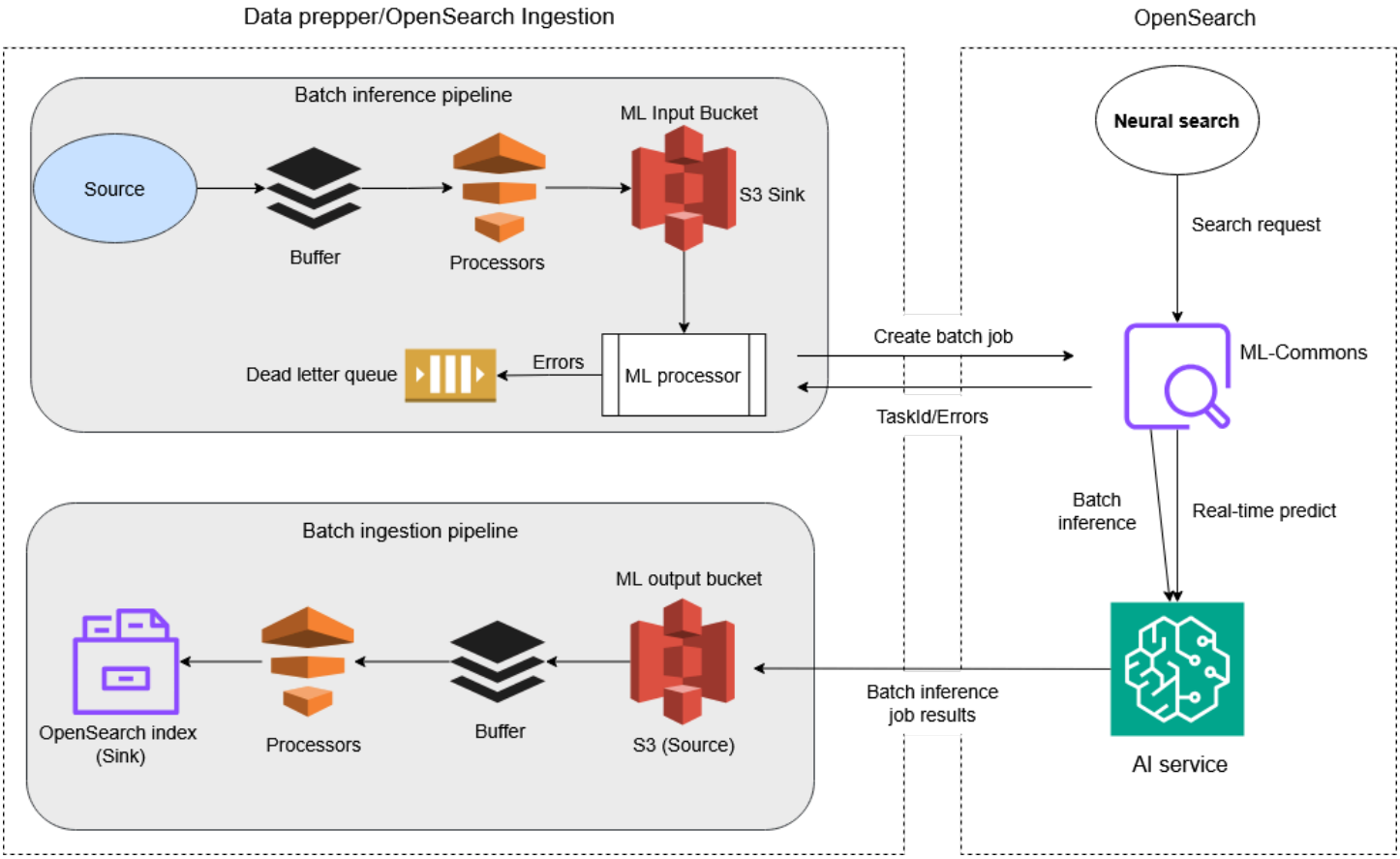
> **Note**
>
> Per l'elaborazione dell'inferenza in tempo reale, usa. Connettori Amazon OpenSearch Service ML per piattaforme di terze parti

L'elaborazione offline dell'inferenza in batch sfrutta una funzionalità OpenSearch chiamata ML Commons. ML Commons fornisce algoritmi ML tramite chiamate API di trasporto e REST. Queste chiamate scelgono i nodi e le risorse giusti per ogni richiesta ML e monitorano le attività ML per garantire l'operatività. In questo modo, ML Commons consente di sfruttare gli algoritmi ML open source esistenti e di ridurre lo sforzo richiesto per sviluppare nuove funzionalità di machine learning. Per ulteriori informazioni su ML Commons, consulta Machine learning nella documentazione.org. OpenSearch

Come funziona

È possibile creare una pipeline di inferenza batch offline su OpenSearch Ingestion aggiungendo un processore di inferenza di machine learning a una pipeline. Questo processore consente alla pipeline di connettersi a servizi di intelligenza artificiale, ad esempio eseguire lavori di inferenza in batch. SageMaker Puoi configurare il processore per connettersi al servizio di intelligenza artificiale desiderato tramite i connettori AI (con supporto batch_predict) in esecuzione sul dominio di destinazione.

OpenSearch Ingestion utilizza il `ml_inference` processore con ML Commons per creare lavori di inferenza batch offline. ML Commons utilizza quindi l'API batch_predict, che esegue l'inferenza su set di dati di grandi dimensioni in modalità asincrona offline utilizzando un modello distribuito su server di modelli esterni in Amazon Bedrock, Amazon, Cohere e OpenAI. SageMaker Il diagramma seguente mostra una pipeline di ingestione che orchestra più componenti per eseguire questo processo dall'inizio alla fine OpenSearch :



I componenti della pipeline funzionano come segue:

Pipeline 1 (preparazione e trasformazione dei dati) *:

- Fonte: i dati vengono scansionati dalla fonte esterna supportata da OpenSearch Ingestion.
- Processori di dati: i dati grezzi vengono elaborati e trasformati nel formato corretto per l'inferenza in batch sul servizio di intelligenza artificiale integrato.
- S3 (Sink): i dati elaborati vengono archiviati in un bucket Amazon S3 pronto a fungere da input per l'esecuzione di processi di inferenza in batch sul servizio di intelligenza artificiale integrato.

Pipeline 2 (Trigger ML batch_inference):

- Fonte: rilevamento automatico degli eventi S3 di nuovi file creati dall'output di Pipeline 1.
- Processore ML_Inference: processore che genera inferenze ML tramite un processo batch asincrono. Si connette ai servizi di intelligenza artificiale tramite il connettore AI configurato in esecuzione sul dominio di destinazione.
- ID attività: ogni processo batch è associato a un ID attività in ml-commons per il monitoraggio e la gestione.
- OpenSearch ML Commons: ML Commons, che ospita il modello per la ricerca neurale in tempo reale, gestisce i connettori ai server AI remoti e serve per l'inferenza in batch e la APIs gestione dei lavori.
- Servizi di intelligenza artificiale: OpenSearch ML Commons interagisce con servizi di intelligenza artificiale come Amazon Bedrock e Amazon SageMaker per eseguire inferenze in batch sui dati, producendo previsioni o approfondimenti. I risultati vengono salvati in modo asincrono in un file S3 separato.

Pipeline 3 (ingestione in blocco):

- S3 (fonte): i risultati dei processi batch vengono archiviati in S3, che è l'origine di questa pipeline.
- Processori di trasformazione dei dati: ulteriori elaborazioni e trasformazioni vengono applicate all'output di inferenza in batch prima dell'ingestione. Ciò garantisce che i dati siano mappati correttamente nell'indice. OpenSearch
- OpenSearch index (Sink): i risultati elaborati vengono indicizzati OpenSearch per l'archiviazione, la ricerca e ulteriori analisi.

> **Note**
>
> *Il processo descritto da Pipeline 1 è facoltativo. Se preferisci, puoi saltare questo processo e caricare semplicemente i dati preparati nel sink S3 per creare lavori in batch.

Informazioni sul processore ml_inference

OpenSearch Ingestion utilizza un'integrazione specializzata tra la sorgente S3 Scan e il processore di inferenza ML per l'elaborazione in batch. S3 Scan funziona in modalità solo metadati per raccogliere in modo efficiente le informazioni sui file S3 senza leggere il contenuto effettivo del file. Il `ml_inference` processore utilizza il file S3 per coordinarsi con ML Commons URLs per

l'elaborazione in batch. Questo design ottimizza il flusso di lavoro di inferenza in batch riducendo al minimo il trasferimento di dati non necessario durante la fase di scansione. Il `ml_inference` processore viene definito utilizzando i parametri. Ecco un esempio:

```
processor:
    - ml_inference:
        # The endpoint URL of your OpenSearch domain
        host: "https://AWS test-offlinebatch-123456789abcdefg.us-
west-2.es.amazonaws.com"

        # Type of inference operation:
        # - batch_predict: for batch processing
        # - predict: for real-time inference
        action_type: "batch_predict"

        # Remote ML model service provider (Amazon Bedrock or SageMaker)
        service_name: "bedrock"

        # Unique identifier for the ML model
        model_id: "AWS TestModelID123456789abcde"

        # S3 path where batch inference results will be stored
        output_path: "s3://amzn-s3-demo-bucket/"

        # Supports ISO_8601 notation strings like PT20.345S or PT15M
        # These settings control how long to keep your inputs in the processor for
 retry on throttling errors
        retry_time_window: "PT9M"

        # AWS configuration settings
        aws:
            # Regione AWS where the Lambda function is deployed
            region: "us-west-2"
            # IAM role ARN for Lambda function execution
            sts_role_arn: "arn:aws::iam::account_id:role/Admin"

        # Dead-letter queue settings for storing errors
        dlq:
          s3:
            region: us-west-2
            bucket: batch-inference-dlq
            key_path_prefix: bedrock-dlq
            sts_role_arn: arn:aws:iam::account_id:role/OSI-invoke-ml
```

```
        # Conditional expression that determines when to trigger the processor
        # In this case, only process when bucket matches "amzn-s3-demo-bucket"
        ml_when: /bucket == "amzn-s3-demo-bucket"
```

Miglioramenti delle prestazioni di ingestione utilizzando il processore ml_inference

Il `ml_inference` processore OpenSearch Ingestion migliora in modo significativo le prestazioni di inserimento dei dati per la ricerca basata su ML. Il processore è ideale per i casi d'uso che richiedono dati generati da modelli di machine learning, tra cui la ricerca semantica, la ricerca multimodale, l'arricchimento dei documenti e la comprensione delle interrogazioni. Nella ricerca semantica, il processore può accelerare la creazione e l'ingestione di vettori di grandi volumi e ad alta dimensione di un ordine di grandezza.

La capacità di inferenza in batch offline del processore offre vantaggi distinti rispetto all'invocazione di modelli in tempo reale. Sebbene l'elaborazione in tempo reale richieda un server modello live con limiti di capacità, l'inferenza in batch ridimensiona dinamicamente le risorse di calcolo su richiesta ed elabora i dati in parallelo. Ad esempio, quando la pipeline OpenSearch Ingestion riceve un miliardo di richieste di dati di origine, crea 100 file S3 per l'input di inferenza in batch ML. Il `ml_inference` processore avvia quindi un processo SageMaker in batch utilizzando 100 istanze `ml.m4.xlarge` Amazon Elastic Compute Cloud (Amazon EC2), completando la vettorizzazione di un miliardo di richieste in 14 ore, un'attività che sarebbe praticamente impossibile da eseguire in modalità in tempo reale.

Configura il processore ml_inference per inserire le richieste di dati per una ricerca semantica

Le seguenti procedure illustrano il processo di impostazione e configurazione del `ml_inference` processore OpenSearch Ingestion per importare un miliardo di richieste di dati per la ricerca semantica utilizzando un modello di incorporamento del testo.

Argomenti

- Fase 1: Creare connettori e registrare modelli in OpenSearch
- Fase 2: Creare una pipeline di OpenSearch ingestione per l'inferenza batch offline di ML
- Fase 3: Preparare i dati per l'ingestione
- Fase 4: Monitoraggio del processo di inferenza in batch
- Passaggio 5: Esegui la ricerca

Fase 1: Creare connettori e registrare modelli in OpenSearch

Per la procedura seguente, usa ML Commons batch_inference_sagemaker_connector_blueprint per creare un connettore e un modello in Amazon. SageMaker Se preferisci OpenSearch CloudFormation utilizzare modelli di (Procedura alternativa) Fase 1: Creazione di connettori e modelli utilizzando un modello di CloudFormation integrazione integrazione, consulta più avanti in questa sezione.

Per creare connettori e registrare modelli in OpenSearch

1. Crea un modello ML Deep Java Library (DJL) SageMaker per la trasformazione in batch. Per visualizzare altri modelli DJL, consulta Semantic_search_with_cfn_template_for_SageMaker su: GitHub

```
POST https://api.sagemaker.us-east-1.amazonaws.com/CreateModel
{
   "ExecutionRoleArn": "arn:aws:iam::123456789012:role/aos_ml_invoke_sagemaker",
   "ModelName": "DJL-Text-Embedding-Model-imageforjsonlines",
   "PrimaryContainer": {
      "Environment": {
         "SERVING_LOAD_MODELS" : "djl://ai.djl.huggingface.pytorch/sentence-
transformers/all-MiniLM-L6-v2"
      },
      "Image": "763104351884.dkr.ecr.us-east-1.amazonaws.com/djl-inference:0.29.0-
cpu-full"
   }
}
```

2. `action`Crea un `batch_predict` connettore con come nuovo tipo nel campo: `actions`

```
POST /_plugins/_ml/connectors/_create
{
  "name": "DJL Sagemaker Connector: all-MiniLM-L6-v2",
  "version": "1",
  "description": "The connector to sagemaker embedding model all-MiniLM-L6-v2",
  "protocol": "aws_sigv4",
  "credential": {
  "roleArn": "arn:aws:iam::111122223333:role/SageMakerRole"
},
  "parameters": {
    "region": "us-east-1",
    "service_name": "sagemaker",
    "DataProcessing": {
```

```
      "InputFilter": "$.text",
      "JoinSource": "Input",
      "OutputFilter": "$"
    },
    "MaxConcurrentTransforms": 100,
    "ModelName": "DJL-Text-Embedding-Model-imageforjsonlines",
    "TransformInput": {
      "ContentType": "application/json",
      "DataSource": {
        "S3DataSource": {
          "S3DataType": "S3Prefix",
          "S3Uri": "s3://offlinebatch/msmarcotests/"
        }
      },
      "SplitType": "Line"
    },
    "TransformJobName": "djl-batch-transform-1-billion",
    "TransformOutput": {
      "AssembleWith": "Line",
      "Accept": "application/json",
      "S3OutputPath": "s3://offlinebatch/msmarcotestsoutputs/"
    },
    "TransformResources": {
      "InstanceCount": 100,
      "InstanceType": "ml.m4.xlarge"
    },
    "BatchStrategy": "SingleRecord"
  },
  "actions": [
    {
      "action_type": "predict",
      "method": "POST",
      "headers": {
        "content-type": "application/json"
      },
      "url": "https://runtime.sagemaker.us-east-1.amazonaws.com/endpoints/
OpenSearch-sagemaker-060124023703/invocations",
      "request_body": "${parameters.input}",
      "pre_process_function": "connector.pre_process.default.embedding",
      "post_process_function": "connector.post_process.default.embedding"
    },
    {
      "action_type": "batch_predict",
      "method": "POST",
```

```
      "headers": {
        "content-type": "application/json"
      },
      "url": "https://api.sagemaker.us-east-1.amazonaws.com/CreateTransformJob",
      "request_body": """{ "BatchStrategy": "${parameters.BatchStrategy}",
 "ModelName": "${parameters.ModelName}", "DataProcessing" :
 ${parameters.DataProcessing}, "MaxConcurrentTransforms":
 ${parameters.MaxConcurrentTransforms}, "TransformInput":
 ${parameters.TransformInput}, "TransformJobName" :
 "${parameters.TransformJobName}", "TransformOutput" :
 ${parameters.TransformOutput}, "TransformResources" :
 ${parameters.TransformResources}}"""
    },
    {
      "action_type": "batch_predict_status",
      "method": "GET",
      "headers": {
        "content-type": "application/json"
      },
      "url": "https://api.sagemaker.us-east-1.amazonaws.com/DescribeTransformJob",
      "request_body": """{ "TransformJobName" :
 "${parameters.TransformJobName}"}"""
    },
    {
      "action_type": "cancel_batch_predict",
      "method": "POST",
      "headers": {
        "content-type": "application/json"
      },
      "url": "https://api.sagemaker.us-east-1.amazonaws.com/StopTransformJob",
      "request_body": """{ "TransformJobName" :
 "${parameters.TransformJobName}"}"""
    }
  ]
}
```

3. Utilizzate l'ID del connettore restituito per registrare il SageMaker modello:

```
POST /_plugins/_ml/models/_register
{
    "name": "SageMaker model for batch",
    "function_name": "remote",
    "description": "test model",
    "connector_id": "example123456789-abcde"
```

```
}
```

4. Invoca il modello con il tipo di `batch_predict` azione:

```
POST /_plugins/_ml/models/teHr3JABBiEvs-eod7sn/_batch_predict
{
  "parameters": {
    "TransformJobName": "SM-offline-batch-transform"
  }
}
```

La risposta contiene un ID attività per il processo batch:

```
{
 "task_id": "exampleIDabdcefd_1234567",
 "status": "CREATED"
}
```

5. Controlla lo stato del processo batch chiamando l'API Get Task utilizzando l'ID dell'attività:

```
GET /_plugins/_ml/tasks/exampleIDabdcefd_1234567
```

La risposta contiene lo stato dell'attività:

```
{
  "model_id": "nyWbv5EB_tT1A82ZCu-e",
  "task_type": "BATCH_PREDICTION",
  "function_name": "REMOTE",
  "state": "RUNNING",
  "input_type": "REMOTE",
  "worker_node": [
    "WDZnIMcbTrGtnR4Lq9jPDw"
  ],
  "create_time": 1725496527958,
  "last_update_time": 1725496527958,
  "is_async": false,
  "remote_job": {
    "TransformResources": {
      "InstanceCount": 1,
      "InstanceType": "ml.c5.xlarge"
    },
    "ModelName": "DJL-Text-Embedding-Model-imageforjsonlines",
```

```
    "TransformOutput": {
      "Accept": "application/json",
      "AssembleWith": "Line",
      "KmsKeyId": "",
      "S3OutputPath": "s3://offlinebatch/output"
    },
    "CreationTime": 1725496531.935,
    "TransformInput": {
      "CompressionType": "None",
      "ContentType": "application/json",
      "DataSource": {
        "S3DataSource": {
          "S3DataType": "S3Prefix",
          "S3Uri": "s3://offlinebatch/sagemaker_djl_batch_input.json"
        }
      },
      "SplitType": "Line"
    },
    "TransformJobArn": "arn:aws:sagemaker:us-east-1:111122223333:transform-job/SM-
offline-batch-transform15",
    "TransformJobStatus": "InProgress",
    "BatchStrategy": "SingleRecord",
    "TransformJobName": "SM-offline-batch-transform15",
    "DataProcessing": {
      "InputFilter": "$.content",
      "JoinSource": "Input",
      "OutputFilter": "$"
    }
  }
}
```

(Procedura alternativa) Fase 1: Creazione di connettori e modelli utilizzando un modello di CloudFormation integrazione

Se preferisci, puoi utilizzarlo AWS CloudFormation per creare automaticamente tutti i SageMaker connettori e i modelli Amazon necessari per l'inferenza ML. Questo approccio semplifica la configurazione utilizzando un modello preconfigurato disponibile nella console di Amazon OpenSearch Service. Per ulteriori informazioni, consulta Utilizzo CloudFormation per configurare l'inferenza remota per la ricerca semantica.

Per implementare uno CloudFormation stack che crei tutti i connettori e i modelli richiesti SageMaker

1. Apri la console Amazon OpenSearch Service.
2. Nel riquadro di navigazione, scegli Integrazioni.
3. Nel campo Cerca, inserisci**`SageMaker`**, quindi scegli Integrazione con modelli di incorporamento di testo tramite Amazon SageMaker.
4. Scegli Configura dominio, quindi scegli Configura dominio VPC o Configura dominio pubblico.
5. Inserisci le informazioni nei campi del modello. Per Enable Offline Batch Inference, scegli true per fornire risorse per l'elaborazione batch offline.
6. Scegliete Crea per creare lo CloudFormation stack.
7. Dopo aver creato lo stack, apri la scheda Outputs nella CloudFormation console Individua connector_id e model_id. Avrai bisogno di questi valori in seguito, quando configurerai la pipeline.

Fase 2: Creare una pipeline di OpenSearch ingestione per l'inferenza batch offline di ML

Utilizzate l'esempio seguente per creare una pipeline di OpenSearch ingestione per l'inferenza batch offline di ML. Per ulteriori informazioni sulla creazione di una pipeline per Ingestion, vedere. OpenSearch Creazione di pipeline Amazon OpenSearch Ingestion

Prima di iniziare

Nell'esempio seguente, si specifica un ARN del ruolo IAM per il `sts_role_arn` parametro. Utilizza la seguente procedura per verificare che questo ruolo sia mappato al ruolo di backend che ha accesso a ml-commons in. OpenSearch

1. Vai al plug-in OpenSearch Dashboards per il tuo dominio di servizio. OpenSearch Puoi trovare l'endpoint delle dashboard nella dashboard del tuo dominio nella console di servizio. OpenSearch
2. Dal menu principale scegli Sicurezza, Ruoli e seleziona il ruolo ml_full_access.
3. Scegliere Utenti mappati, Gestisci mappatura.
4. In Ruoli di backend, inserisci l'ARN del ruolo Lambda che richiede l'autorizzazione per chiamare il tuo dominio. Ecco un esempio: arn:aws:iam: ::role/ *`111122223333 lambda-role`*
5. Selezionare Mappa e confermare che l'utente o il ruolo venga visualizzato in Utenti mappati.

Esempio per creare una pipeline di ingestione per l'inferenza batch offline di ML OpenSearch

```
version: '2'
extension:
  osis_configuration_metadata:
    builder_type: visual
sagemaker-batch-job-pipeline:
  source:
    s3:
      acknowledgments: true
      delete_s3_objects_on_read: false
      scan:
        buckets:
          - bucket:
              name: name
              data_selection: metadata_only
              filter:
                include_prefix:
                  - sagemaker/sagemaker_djl_batch_input
                exclude_suffix:
                  - .manifest
          - bucket:
              name: name
              data_selection: data_only
              filter:
                include_prefix:
                  - sagemaker/output/
        scheduling:
          interval: PT6M
      aws:
        region: name
      default_bucket_owner: account_ID
      codec:
        ndjson:
          include_empty_objects: false
      compression: none
      workers: '1'
  processor:
    - ml_inference:
        host: "https://search-AWStest-offlinebatch-123456789abcdef.us-
west-2.es.amazonaws.com"
        aws_sigv4: true
        action_type: "batch_predict"
        service_name: "sagemaker"
        model_id: "model_ID"
```

```
        output_path: "s3://AWStest-offlinebatch/sagemaker/output"
        aws:
          region: "us-west-2"
          sts_role_arn: "arn:aws:iam::account_ID:role/Admin"
        ml_when: /bucket == "AWStest-offlinebatch"
        dlq:
          s3:
            region: us-west-2
            bucket: batch-inference-dlq
            key_path_prefix: bedrock-dlq
            sts_role_arn: arn:aws:iam::account_ID:role/OSI-invoke-ml
    - copy_values:
        entries:
          - from_key: /text
            to_key: chapter
          - from_key: /SageMakerOutput
            to_key: chapter_embedding
          - delete_entries:
            with_keys:
          - text
          - SageMakerOutput
  sink:
    - opensearch:
        hosts: ["https://search-AWStest-offlinebatch-123456789abcdef.us-
west-2.es.amazonaws.com"]
        aws:
          serverless: false
          region: us-west-2
        routes:
          - ml-ingest-route
        index_type: custom
        index: test-nlp-index
  routes:
    - ml-ingest-route: /chapter != null and /title != null
```

Fase 3: Preparare i dati per l'ingestione

Per preparare i dati per l'elaborazione offline dell'inferenza in batch di machine learning, preparate voi stessi i dati utilizzando i vostri strumenti o processi oppure utilizzate Data Prepper. OpenSearch Verifica che i dati siano organizzati nel formato corretto utilizzando una pipeline per utilizzare i dati dalla tua fonte di dati o creando un set di dati di apprendimento automatico.

L'esempio seguente utilizza il set di dati MS MARCO, che include una raccolta di query utente reali per attività di elaborazione del linguaggio naturale. Il set di dati è strutturato in formato JSONL, in cui ogni riga rappresenta una richiesta inviata al modello di incorporamento ML:

```
{"_id": "1185869", "text": ")what was the immediate impact of the Paris Peace Treaties
 of 1947?", "metadata": {"world war 2"}}
{"_id": "1185868", "text": "_________ justice is designed to repair the harm to victim,
 the community and the offender caused by the offender criminal act. question 19
 options:", "metadata": {"law"}}
{"_id": "597651", "text": "what is amber", "metadata": {"nothing"}}
{"_id": "403613", "text": "is autoimmune hepatitis a bile acid synthesis disorder",
 "metadata": {"self immune"}}
...
```

Per testare l'utilizzo del set di dati MS MARCO, immagina uno scenario in cui crei un miliardo di richieste di input distribuite su 100 file, ciascuno contenente 10 milioni di richieste. I file verrebbero archiviati in Amazon S3 con il prefisso s3://_djl_batch_input/. offlinebatch/sagemaker/sagemaker La pipeline OpenSearch Ingestion scansionerebbe questi 100 file contemporaneamente e avvierebbe un processo in SageMaker batch con 100 addetti per l'elaborazione parallela, consentendo la vettorizzazione e l'ingestione efficienti del miliardo di documenti in esso contenuti. OpenSearch

Negli ambienti di produzione, è possibile utilizzare una pipeline di OpenSearch Ingestion per generare file S3 per l'input di inferenza in batch. La pipeline supporta varie fonti di dati e opera secondo una pianificazione per trasformare continuamente i dati di origine in file S3. Questi file vengono quindi elaborati automaticamente dai server AI tramite processi batch offline pianificati, garantendo l'elaborazione e l'ingestione continue dei dati.

Fase 4: Monitoraggio del processo di inferenza in batch

È possibile monitorare i lavori di inferenza in batch utilizzando la SageMaker console o il. AWS CLI Puoi anche utilizzare l'API Get Task per monitorare i processi in batch:

```
GET /_plugins/_ml/tasks/_search
{
  "query": {
    "bool": {
      "filter": [
        {
          "term": {
            "state": "RUNNING"
          }
```

```
        }
      ]
    }
  },
  "_source": ["model_id", "state", "task_type", "create_time", "last_update_time"]
}
```

L'API restituisce un elenco di attività di lavoro in batch attive:

```
{
  "took": 2,
  "timed_out": false,
  "_shards": {
    "total": 5,
    "successful": 5,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": {
      "value": 3,
      "relation": "eq"
    },
    "max_score": 0.0,
    "hits": [
      {
        "_index": ".plugins-ml-task",
        "_id": "nyWbv5EB_tT1A82ZCu-e",
        "_score": 0.0,
        "_source": {
          "model_id": "nyWbv5EB_tT1A82ZCu-e",
          "state": "RUNNING",
          "task_type": "BATCH_PREDICTION",
          "create_time": 1725496527958,
          "last_update_time": 1725496527958
        }
      },
      {
        "_index": ".plugins-ml-task",
        "_id": "miKbv5EB_tT1A82ZCu-f",
        "_score": 0.0,
        "_source": {
          "model_id": "miKbv5EB_tT1A82ZCu-f",
```

```
          "state": "RUNNING",
          "task_type": "BATCH_PREDICTION",
          "create_time": 1725496528123,
          "last_update_time": 1725496528123
        }
      },
      {
        "_index": ".plugins-ml-task",
        "_id": "kiLbv5EB_tT1A82ZCu-g",
        "_score": 0.0,
        "_source": {
          "model_id": "kiLbv5EB_tT1A82ZCu-g",
          "state": "RUNNING",
          "task_type": "BATCH_PREDICTION",
          "create_time": 1725496529456,
          "last_update_time": 1725496529456
        }
      }
    ]
  }
}
```

Passaggio 5: Esegui la ricerca

Dopo aver monitorato il processo di inferenza in batch e averne confermato il completamento, puoi eseguire vari tipi di ricerche AI, tra cui semantiche, ibride, conversazionali (con RAG), neural sparse e multimodali. Per ulteriori informazioni sulle ricerche AI supportate dal Servizio, consulta Ricerca AI. OpenSearch

Per cercare vettori grezzi, usa il tipo di `knn` query, fornisci l'`vector`array come input e specifica il `k` numero di risultati restituiti:

```
GET /my-raw-vector-index/_search
{
  "query": {
    "knn": {
      "my_vector": {
        "vector": [0.1, 0.2, 0.3],
        "k": 2
      }
    }
  }
```

```
}
```

Per eseguire una ricerca basata sull'intelligenza artificiale, utilizza il `neural` tipo di query. Specificate l'`query_text`input, il modello `model_id` di incorporamento configurato nella pipeline di OpenSearch ingestione e il numero di risultati restituiti. k Per escludere gli incorporamenti dai risultati della ricerca, specificate il nome del campo di incorporamento nel parametro: `_source.excludes`

```
GET /my-ai-search-index/_search
{
  "_source": {
    "excludes": [
      "output_embedding"
    ]
  },
  "query": {
    "neural": {
      "output_embedding": {
        "query_text": "What is AI search?",
        "model_id": "mBGzipQB2gmRjlv_dOoB",
        "k": 2
      }
    }
  }
}
```

Utilizzo di una pipeline di OpenSearch ingestione con Collector OpenTelemetry

Puoi usare OpenTelemetry Collector per inserire log, tracce e metriche nelle pipeline di Ingestion. OpenSearch È possibile utilizzare una singola pipeline per importare tutti i log, le tracce e le metriche in diversi indici di un dominio o di una raccolta. Puoi anche utilizzare le pipeline per importare solo log, tracce o metriche singolarmente.

Argomenti

- Prerequisiti
- Fase 1: Configurare il ruolo della pipeline
- Fase 2: Creare la pipeline
- Connettività tra account

- Limitazioni
- Allarmi consigliati CloudWatch per le fonti OpenTelemetry

Prerequisiti

Durante la configurazione del file di OpenTelemetry configurazione, è necessario configurare quanto segue affinché avvenga l'ingestione:

- Il ruolo di importazione richiede l'`osis:Ingest`autorizzazione per interagire con la pipeline. Per ulteriori informazioni, consulta Ingestion role.
- Il valore dell'endpoint deve includere l'endpoint della pipeline. Ad esempio, `https://pipeline-endpoint.us-east-1.osis.amazonaws.com.`
- Il valore del servizio deve essere. `osis`
- L'opzione di compressione per OTLP/HTTP Exporter deve corrispondere all'opzione di compressione sulla sorgente selezionata della pipeline.

```
extensions:
    sigv4auth:
        region: "region"
        service: "osis"

exporters:
    otlphttp:
        logs_endpoint: "https://pipeline-endpoint.us-east-1.osis.amazonaws.com/v1/logs"
        metrics_endpoint: "https://pipeline-endpoint.us-east-1.osis.amazonaws.com/v1/
metrics"
        traces_endpoint: "https://pipeline-endpoint.us-east-1.osis.amazonaws.com/v1/
traces"
        auth:
            authenticator: sigv4auth
        compression: none

service:
    extensions: [sigv4auth]
    pipelines:
        traces:
        receivers: [jaeger]
        exporters: [otlphttp]
```

Fase 1: Configurare il ruolo della pipeline

Dopo aver impostato la configurazione del OpenTelemetry raccoglitore, imposta il ruolo della pipeline che desideri utilizzare nella configurazione della pipeline. Il ruolo pipeline non richiede autorizzazioni specifiche per l'origine OTLP, ma solo autorizzazioni per concedere alle pipeline l'accesso al dominio o alla raccolta. OpenSearch

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestione come la seguente, che specifica OTLP come origine. È inoltre possibile configurare OpenTelemetry log, metriche e tracce come fonti individuali.

Configurazione della pipeline di origine OTLP:

```
version: 2
otlp-pipeline:
    source:
        otlp:
            logs_path: /otlp-pipeline/v1/logs
            traces_path: /otlp-pipeline/v1/traces
            metrics_path: /otlp-pipeline/v1/metrics
    sink:
        - opensearch:
            hosts: ["https://search-mydomain.region.es.amazonaws.com"]
            index: "ss4o_metrics-otel-%{yyyy.MM.dd}"
            index_type: custom
            aws:
                region: "region"
```

OpenTelemetry Registra la configurazione della pipeline:

```
version: 2
otel-logs-pipeline:
  source:
    otel_logs_source:
        path: /otel-logs-pipeline/v1/logs
  sink:
    - opensearch:
        hosts: ["https://search-mydomain.region.es.amazonaws.com"]
```

```
        index: "ss4o_metrics-otel-%{yyyy.MM.dd}"
        index_type: custom
        aws:
            region: "region"
```

OpenTelemetry Configurazione della pipeline delle metriche:

```
version: 2
otel-metrics-pipeline:
  source:
    otel_metrics_source:
        path: /otel-metrics-pipeline/v1/metrics
  sink:
    - opensearch:
        hosts: ["https://search-mydomain.region.es.amazonaws.com"]
        index: "ss4o_metrics-otel-%{yyyy.MM.dd}"
        index_type: custom
        aws:
            region: "region"
```

OpenTelemetry Configurazione della pipeline di Traces:

```
version: 2
otel-trace-pipeline:
  source:
    otel_trace_source:
        path: /otel-traces-pipeline/v1/traces
  sink:
    - opensearch:
        hosts: ["https://search-mydomain.region.es.amazonaws.com"]
        index: "ss4o_metrics-otel-%{yyyy.MM.dd}"
        index_type: custom
        aws:
            region: "region"
```

È possibile utilizzare un blueprint preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Connettività tra account

OpenSearch Le pipeline di ingestione con OpenTelemetry sorgenti hanno una capacità di ingestione tra account. Amazon OpenSearch Ingestion ti consente di condividere le pipeline Account AWS da un cloud privato virtuale (VPC) a un endpoint di pipeline in un VPC separato. Per ulteriori informazioni, consulta the section called “Configurazione delle pipeline per l'ingestione tra account”.

Limitazioni

La pipeline OpenSearch Ingestion non può ricevere richieste superiori a 20 MB. Questo valore è configurato dall'utente nell'opzione. `max_request_length` Il valore predefinito di questa opzione è 10 MB.

Allarmi consigliati CloudWatch per le fonti OpenTelemetry

Le seguenti CloudWatch metriche sono consigliate per monitorare le prestazioni della pipeline di ingestione. Queste metriche possono aiutarti a identificare la quantità di dati elaborati dalle esportazioni, la quantità di eventi elaborati dagli stream, gli errori nell'elaborazione delle esportazioni e degli eventi di streaming e il numero di documenti scritti nella destinazione. È possibile impostare CloudWatch allarmi per eseguire un'azione quando una di queste metriche supera un valore specificato per un determinato periodo di tempo.

Le CloudWatch metriche per la sorgente OTLP sono formattate come. `{pipeline-name}.otlp.{logs | traces | metrics}.{metric-name}` Ad esempio, `otel-pipeline.otlp.metrics.requestTimeouts.count`.

Nel caso di utilizzo di una singola OpenTelemetry fonte, le metriche verranno formattate come. `{pipeline-name}.{source-name}.{metric-name}` Ad esempio, `trace-pipeline.otel_trace_source.requestTimeouts.count`.

Tutti e tre OpenTelemetry i tipi di dati avranno le stesse metriche, ma per brevità le metriche verranno elencate solo nella tabella seguente per i dati di tipo di registro di origine OTLP.

Metrica	Description
`otel-pipeline.BlockingBuffer.bufferUsage.value`	Indica la quantità di buffer utilizzata.
`otel-pipeline.otlp.logs.requestTimeouts.count`	Il numero di richieste scadute.

Metrica	Description
`otel-pipeline.otlp.logs.requestsReceived.count`	Il numero di richieste ricevute dal OpenTelemetry Collector.
`otel-pipeline.otlp.logs.badRequests.count`	Il numero di richieste non corrette ricevute dal Collector. OpenTelemetry
`otel-pipeline.otlp.logs.requestsTooLarge.count`	Il numero di richieste superiore al massimo di 20 MB ricevute dal Collector. OpenTelemetry
`otel-pipeline.otlp.logs.internalServerError.count`	Il numero di errori HTTP 500 ricevuti dal Collector. OpenTelemetry
`otel-pipeline.opensearch.bulkBadRequestErrors.count`	Numero di errori durante le richieste in blocco a causa di una richiesta non valida.
`otel-pipeline.opensearch.bulkRequestLatency.avg`	Latenza media per le richieste di scrittura in blocco effettuate a. OpenSearch
`otel-pipeline.opensearch.bulkRequestNotFoundErrors.count`	Numero di richieste in blocco non riuscite perché non è stato possibile trovare i dati di destinazione.
`otel-pipeline.opensearch.bulkRequestNumberOfRetries.count`	Numero di tentativi di scrittura del cluster da parte delle pipeline di OpenSearch Ingestion. OpenSearch
`otel-pipeline.opensearch.bulkRequestSizeBytes.sum`	Dimensione totale in byte di tutte le richieste in blocco effettuate a. OpenSearch
`otel-pipeline.opensearch.documentErrors.count`	Numero di errori durante l'invio di documenti a. OpenSearch I documenti che causano gli errori verranno inviati a DLQ.
`otel-pipeline.opensearch.documentsSuccess.count`	Numero di documenti scritti correttamente in un OpenSearch cluster o in una raccolta.

Metrica	Description
otel-pipeline.opensearch.documentsSuccessFirstAttempt.count	Numero di documenti indicizzati correttamente OpenSearch al primo tentativo.
otel-pipeline.opensearch.documentsVersionConflictErrors.count	Numero di errori dovuti a conflitti di versione nei documenti durante l'elaborazione.
otel-pipeline.opensearch.PipelineLatency.avg	Latenza media della pipeline di OpenSearch Ingestion per l'elaborazione dei dati dalla lettura dall'origine alla scrittura nella destinazione.
otel-pipeline.opensearch.PipelineLatency.max	Latenza massima della pipeline di OpenSearch Ingestion per elaborare i dati leggendo dall'origine alla scrittura della destinazione.
otel-pipeline.opensearch.recordsIn.count	Numero di record in cui sono stati inseriti con successo. OpenSearch Questa metrica è essenziale per tenere traccia del volume di dati elaborati e archiviati.
otel-pipeline.opensearch.s3.dlqS3RecordsFailed.count	Numero di record che non sono stati scritti su DLQ.
otel-pipeline.opensearch.s3.dlqS3RecordsSuccess.count	Numero di record scritti su DLQ.
otel-pipeline.opensearch.s3.dlqS3RequestLatency.count	Conteggio delle misurazioni della latenza per le richieste alla coda di lettere morte di Amazon S3.
otel-pipeline.opensearch.s3.dlqS3RequestLatency.sum	Latenza totale per tutte le richieste alla coda di lettera morta di Amazon S3
otel-pipeline.opensearch.s3.dlqS3RequestSizeBytes.sum	Dimensione totale in byte di tutte le richieste effettuate alla coda di lettere morte di Amazon S3.

Metrica	Description
`otel-pipeline.recordsProcessed.count`	Numero totale di record elaborati nella pipeline, una metrica chiave per il throughput complessivo.
`otel-pipeline.opensearch.bulkRequestInvalidInputErrors.count`	Conteggio degli errori nelle richieste di massa OpenSearch dovuti a input non validi, fondamentale per il monitoraggio della qualità dei dati e dei problemi operativi.

Utilizzo di una pipeline OpenSearch di ingestione con Amazon Managed Service for Prometheus

Puoi utilizzare Amazon Managed Service for Prometheus come destinazione per la OpenSearch tua pipeline di Ingestion per archiviare le metriche in formato serie temporale. Il sink Prometheus ti consente di OpenTelemetry inviare metriche o altre serie temporali dalla tua pipeline a un'area di lavoro di Amazon Managed Service for Prometheus per il monitoraggio, gli avvisi e l'analisi.

Il plug-in `prometheus` sink consente alle pipeline di OpenSearch Ingestion di scrivere dati metrici negli spazi di lavoro di Amazon Managed Service for Prometheus utilizzando il protocollo di scrittura remota Prometheus. Questa integrazione ti consente di:

- Archivia i dati delle metriche delle serie temporali in Amazon Managed Service for Prometheus
- Monitora e invia avvisi sui parametri utilizzando Amazon Managed Service for Prometheus e Amazon Managed Grafana
- Indirizza le metriche verso più destinazioni contemporaneamente (ad esempio, OpenSearch e Amazon Managed Service for Prometheus)
- Elabora le OpenTelemetry metriche degli agenti esterni o genera metriche all'interno della pipeline

Argomenti

- Prerequisiti
- Passaggio 1: configura il ruolo della pipeline
- Fase 2: Creare la pipeline
- Monitoraggio e risoluzione dei problemi

- Limitazioni
- Best practice

Prerequisiti

Prima di configurare il lavabo Prometheus, assicurati di avere quanto segue:

- Amazon Managed Service for Prometheus Workspace: crea uno spazio di lavoro nella stessa pipeline di Ingestion. Account AWS Regione AWS OpenSearch Per istruzioni, consulta Creazione di uno spazio di lavoro nella Guida per l'utente di Amazon Managed Service for Prometheus.
- Autorizzazioni IAM: configura un ruolo IAM con autorizzazioni per scrivere su Amazon Managed Service for Prometheus. Per ulteriori informazioni, consulta the section called “Passaggio 1: configura il ruolo della pipeline”.

> **Note**
>
> Le aree di lavoro di Amazon Managed Service for Prometheus devono utilizzare chiavi gestite dal servizio. AWS AWS KMS Le AWS KMS chiavi gestite dai clienti non sono attualmente supportate per i sink Amazon Managed Service for Prometheus nelle pipeline di ingestione. OpenSearch

Passaggio 1: configura il ruolo della pipeline

Il sink Prometheus eredita automaticamente le autorizzazioni IAM del ruolo pipeline per l'autenticazione, quindi non è richiesta alcuna configurazione aggiuntiva del ruolo (`sts_role_arn`come) nelle impostazioni del sink.

La seguente policy di esempio mostra le autorizzazioni richieste per l'utilizzo di Amazon Managed Service for Prometheus come sink:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AMPRemoteWrite",
      "Effect": "Allow",
      "Action": [
```

```
        "aps:RemoteWrite"
      ],
      "Resource": "arn:aws:aps:region:account-id:workspace/workspace-id"
    }
  ]
}
```

Sostituire i seguenti segnaposto:

- *region*: La tua Regione AWS (ad esempio,) `us-east-1`
- *account-id*: Il tuo Account AWS ID
- *workspace-id*: il tuo ID dell'area di lavoro Amazon Managed Service per Prometheus

Devi associare queste autorizzazioni al tuo ruolo di pipeline.

Assicurati che il tuo ruolo nella pipeline abbia un rapporto di fiducia che consenta a OpenSearch Ingestion di assumerlo:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "osis-pipelines.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

Fase 2: Creare la pipeline

Dopo aver impostato le autorizzazioni, puoi configurare una pipeline di OpenSearch Ingestion per utilizzare Amazon Managed Service for Prometheus come sink.

Configurazione di base

L'esempio seguente mostra una configurazione minima del sink Prometheus:

```
version: "2"
```

```
sink:
  - prometheus:
      url: "https://aps-workspaces.region.amazonaws.com/workspaces/workspace-id/api/v1/
remote_write"
      aws:
        region: "region"
```

È necessario specificare l'`url`opzione all'interno della configurazione `prometheus` sink, che è l'endpoint di scrittura remota Amazon Managed Service for Prometheus. Per formattare l'URL, individua l'ID del tuo spazio di lavoro nella console Amazon Managed Service for Prometheus e costruisci l'URL come segue:. `https://aps-workspaces.region.amazonaws.com/workspaces/workspace-id/api/v1/remote_write`

Opzioni di configurazione

Utilizza le seguenti opzioni per configurare il comportamento di batching e flushing per il sink Prometheus:

Opzioni di configurazione del lavello Prometheus

Opzione	Richiesto	Tipo	Description
`max_events`	No	Numero intero	Il numero massimo di eventi da accumular e prima di passare a Prometheus. Il valore predefinito è “`1000`”.
`max_reque st_size`	No	Numero di byte	La dimensione massima del payload della richiesta prima del flushing. Il valore predefinito è “`1mb`”.
`flush_interval`	No	Durata	Il tempo massimo di attesa prima degli eventi di flushing. Il valore predefinito è “`10s`”. Il valore massimo consentito è. `60s`

Pipeline di esempio

Esempio 1: OpenTelemetry metriche relative ad Amazon Managed Service for Prometheus

Questa pipeline riceve i OpenTelemetry parametri da un agente esterno e li scrive su Amazon Managed Service for Prometheus:

```
version: "2"
source:
  otel_metrics_source:
    path: "/v1/metrics"
    output_format: otel

sink:
  - prometheus:
      url: "https://aps-workspaces.us-east-1.amazonaws.com/workspaces/ws-
a1b2c3d4-5678-90ab-cdef-EXAMPLE11111/api/v1/remote_write"
      aws:
        region: "us-east-1"
```

Esempio 2: doppio lavabo OpenSearch e Amazon Managed Service per Prometheus

Questa pipeline indirizza le metriche sia ad OpenSearch Amazon Managed Service for Prometheus che ad Amazon:

```
version: "2"
source:
  otel_metrics_source:
    path: "/v1/metrics"
    output_format: otel

sink:
  - opensearch:
      hosts:
        - "https://search-domain-endpoint.us-east-1.es.amazonaws.com"
      index: "metrics-%{yyyy.MM.dd}"
      aws:
        region: "us-east-1"
        sts_role_arn: "arn:aws:iam::123456789012:role/OSI-Pipeline-Role"

  - prometheus:
      url: "https://aps-workspaces.us-east-1.amazonaws.com/workspaces/ws-
a1b2c3d4-5678-90ab-cdef-EXAMPLE11111/api/v1/remote_write"
      aws:
        region: "us-east-1"
```

Esempio 3: metriche con filtro

Questa pipeline filtra le metriche prima di inviarle ad Amazon Managed Service for Prometheus:

```
version: "2"
source:
  otel_metrics_source:
    path: "/v1/metrics"
    output_format: otel

processor:
  - drop_events:
      drop_when: '/name != "http.server.duration" and /name != "http.client.duration"'

sink:
  - prometheus:
      url: "https://aps-workspaces.us-east-1.amazonaws.com/workspaces/ws-
a1b2c3d4-5678-90ab-cdef-EXAMPLE11111/api/v1/remote_write"
      aws:
        region: "us-east-1"
```

Puoi utilizzare un blueprint Amazon Managed Service for Prometheus preconfigurato per creare queste pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Creazione di una pipeline con Amazon Managed Service for Prometheus sink

AWS Utilizzo della console

1. Vai alla console OpenSearch di servizio.
2. Scegli Pipelines in Ingestion.
3. Scegliere Create pipeline (Crea pipeline).
4. Seleziona Crea usando blueprint e scegli i OpenTelemetry parametri per il blueprint Amazon Prometheus.
5. Configura la pipeline:
   - Inserisci il tuo ID dello spazio di lavoro Amazon Managed Service for Prometheus
   - Specificare il ruolo della pipeline (ARN)
   - Configura le impostazioni di origine e processore in base alle esigenze
6. Rivedi e crea la pipeline.

Utilizzo della AWS CLI

Crea un file di configurazione della pipeline (ad esempio,`amp-pipeline.yaml`) con la configurazione desiderata, quindi esegui:

```
aws osis create-pipeline \
  --pipeline-name my-amp-pipeline \
  --min-units 2 \
  --max-units 4 \
  --pipeline-configuration-body file://amp-pipeline.yaml
```

Usando AWS CloudFormation

```
Resources:
  MyAMPPipeline:
    Type: AWS::OSIS::Pipeline
    Properties:
      PipelineName: my-amp-pipeline
      MinUnits: 2
      MaxUnits: 4
      PipelineConfigurationBody: |
        version: "2"
        source:
          otel_metrics_source:
            path: "/v1/metrics"
            output_format: otel
        sink:
          - prometheus:
              url: "https://aps-workspaces.us-east-1.amazonaws.com/workspaces/ws-
a1b2c3d4-5678-90ab-cdef-EXAMPLE11111/api/v1/remote_write"
              aws:
                region: "us-east-1"
```

Monitoraggio e risoluzione dei problemi

CloudWatch metriche

Monitora le prestazioni della tua pipeline utilizzando CloudWatch le metriche:

- `DocumentsWritten`: Numero di metriche scritte correttamente in Amazon Managed Service for Prometheus
- `DocumentsWriteFailed`: numero di metriche che non è stato possibile scrivere

- `RequestLatency`: Latenza delle richieste di scrittura remote

Problemi comuni

Problema: Pipeline non riesce a scrivere su Amazon Managed Service for Prometheus

Soluzioni:

- Verifica che l'ID dell'area di lavoro e la regione nell'URL siano corretti
- Assicurati che il ruolo della pipeline sia autorizzato `aps:RemoteWrite`
- Verifica che l'area di lavoro utilizzi chiavi gestite dal servizio AWS KMS
- Verifica che la pipeline e l'area di lavoro siano uguali Account AWS

Problema: errori di autenticazione

Soluzioni:

- Verifica che la relazione di fiducia `osis-pipelines.amazonaws.com` consenta di assumere il ruolo di pipeline
- Assicurati che il ruolo della pipeline disponga dell'autorizzazione richiesta `aps:RemoteWrite`

Problema: latenza o limitazione elevate

Soluzioni:

- Aumentare la capacità delle unità della pipeline
- Implementa il batching nel processore
- Consulta le quote dei servizi Amazon Managed Service per Prometheus

Limitazioni

Considera le seguenti limitazioni quando configuri una pipeline di OpenSearch ingestione per Amazon Managed Service for Prometheus:

- Le aree di lavoro di Amazon Managed Service for Prometheus devono utilizzare chiavi gestite dal servizio. AWS AWS KMS Le chiavi gestite dal cliente AWS KMS non sono attualmente supportate.

- La pipeline e l'area di lavoro di Amazon Managed Service for Prometheus devono coincidere. Account AWS

Best practice

- Usa lo stesso ruolo IAM: il sink Prometheus utilizza automaticamente il ruolo pipeline. Se vengono utilizzati altri sink, assicurati che `sts_role_arn` sia lo stesso ruolo della pipeline
- Monitora le metriche: imposta CloudWatch allarmi per scritture non riuscite e latenza elevata
- Implementazione del filtraggio: utilizza i processori per filtrare le metriche non necessarie prima di inviarle ad Amazon Managed Service for Prometheus
- Capacità della giusta dimensione: inizia con la capacità minima e la scalabilità in base al volume delle metriche
- Usa i blueprint: sfrutta i blueprint preconfigurati per i casi d'uso più comuni

Utilizzo di una pipeline OpenSearch di ingestione con Kafka

Puoi utilizzare il plug-in Kafka per trasmettere dati da cluster Kafka autogestiti a domini OpenSearch Amazon Service e raccolte Serverless. OpenSearch OpenSearch Ingestion supporta le connessioni dai cluster Kafka configurati con reti pubbliche o private (VPC). Questo argomento descrive i prerequisiti e i passaggi per configurare una pipeline di ingestione, inclusa la configurazione delle impostazioni di rete e dei metodi di autenticazione come Mutual TLS (mTLS), SASL/SCRAM o IAM.

Migrazione dei dati dai cluster Kafka pubblici

È possibile utilizzare le pipeline di OpenSearch Ingestion per migrare i dati da un cluster Kafka pubblico autogestito, il che significa che il nome DNS del dominio può essere risolto pubblicamente. A tale scopo, configura una pipeline di OpenSearch ingestion con Kafka autogestito come origine e Service o Serverless come destinazione. OpenSearch OpenSearch Questo elabora i dati di streaming da un cluster di origine autogestito a un dominio o raccolta di destinazione gestito. AWS

Prerequisiti

Prima di creare la pipeline OpenSearch di Ingestion, effettuate le seguenti operazioni:

1. Crea un cluster Kafka autogestito con una configurazione di rete pubblica. Il cluster deve contenere i dati che desideri inserire nel servizio. OpenSearch

2. Crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless verso cui migrare i dati. Per ulteriori informazioni, consultare the section called “Creazione OpenSearch di domini di servizio” e the section called “Creazione di raccolte”.
3. Configura l'autenticazione sul tuo cluster autogestito con. Gestione dei segreti AWS Abilita la rotazione dei segreti seguendo la procedura descritta in Ruota Gestione dei segreti AWS i segreti.
4. Allega una politica basata sulle risorse al tuo dominio o una politica di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal cluster autogestito al dominio o alla raccolta.

   Il seguente esempio di policy di accesso al dominio consente al ruolo pipeline, creato nel passaggio successivo, di scrivere dati su un dominio. Assicurati di aggiornarlo `resource` con il tuo ARN.

   JSON

   ```
   {
     "Version":"2012-10-17",
     "Statement": [
       {
         "Effect": "Allow",
         "Principal": {
           "AWS": "arn:aws:iam::444455556666:role/pipeline-role"
         },
         "Action": [
           "es:DescribeDomain",
           "es:ESHttp*"
         ],
         "Resource": [
           "arn:aws:es:us-east-1:111122223333:domain/domain-name"
         ]
       }
     ]
   }
   ```

   Per creare un ruolo IAM con le autorizzazioni corrette per accedere ai dati di scrittura nella raccolta o nel dominio, consulta. the section called “Configurazione di ruoli e utenti”

Fase 1: Configurare il ruolo della pipeline

Dopo aver impostato i prerequisiti della pipeline Kafka, configura il ruolo della pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi l'autorizzazione alla scrittura su un dominio di OpenSearch servizio o una raccolta OpenSearch Serverless, nonché l'autorizzazione a leggere segreti da Secrets Manager.

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestione come la seguente, che specifica Kafka come origine.

È possibile specificare più domini di OpenSearch servizio come destinazioni per i dati. Questa funzionalità consente il routing o la replica condizionale dei dati in entrata in più domini di servizio. OpenSearch

Puoi anche migrare i dati da un cluster Confluent Kafka di origine a una raccolta VPC serverless. OpenSearch Assicurati di fornire una politica di accesso alla rete all'interno della configurazione della pipeline. È possibile utilizzare un registro di schemi Confluent per definire uno schema Confluent.

```
version: "2"
kafka-pipeline:
  source:
    kafka:
      encryption:
        type: "ssl"
      topics:
        - name: "topic-name"
          group_id: "group-id"
      bootstrap_servers:
        - "bootstrap-server.us-east-1.aws.private.confluent.cloud:9092"
      authentication:
        sasl:
          plain:
            username: ${aws_secrets:confluent-kafka-secret:username}
            password: ${aws_secrets:confluent-kafka-secret:password}
      schema:
        type: confluent
        registry_url: https://my-registry.us-east-1.aws.confluent.cloud
        api_key: "${{aws_secrets:schema-secret:schema_registry_api_key}}"
        api_secret: "${{aws_secrets:schema-secret:schema_registry_api_secret}}"
        basic_auth_credentials_source: "USER_INFO"
```

```
  sink:
  - opensearch:
      hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
      aws:
          region: "us-east-1"
      index: "confluent-index"
extension:
  aws:
    secrets:
      confluent-kafka-secret:
        secret_id: "my-kafka-secret"
        region: "us-east-1"
      schema-secret:
        secret_id: "my-self-managed-kafka-schema"
        region: "us-east-1"
```

È possibile utilizzare un blueprint preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Migrazione dei dati dai cluster Kafka in un VPC

Puoi anche utilizzare le pipeline di OpenSearch Ingestion per migrare i dati da un cluster Kafka autogestito in esecuzione in un VPC. A tale scopo, configura una pipeline di OpenSearch ingestion con Kafka autogestito come origine e Service o Serverless come destinazione. OpenSearch OpenSearch Questo elabora i dati di streaming da un cluster di origine autogestito a un dominio o raccolta di destinazione gestito. AWS

Prerequisiti

Prima di creare la pipeline OpenSearch di Ingestion, effettuate le seguenti operazioni:

1. Crea un cluster Kafka autogestito con una configurazione di rete VPC che contenga i dati che desideri importare in Service. OpenSearch
2. Crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless verso cui migrare i dati. Per ulteriori informazioni, consulta Creazione di domini di OpenSearch servizio e Creazione di raccolte.
3. Configura l'autenticazione sul tuo cluster autogestito con. Gestione dei segreti AWS Abilita la rotazione dei segreti seguendo la procedura descritta in Ruota Gestione dei segreti AWS i segreti.
4. Ottieni l'ID del VPC che ha accesso a Kafka autogestito. Scegli il VPC CIDR da utilizzare con Ingestion. OpenSearch

> **Note**
>
> Se utilizzi il Console di gestione AWS per creare la tua pipeline, devi anche collegare la pipeline di OpenSearch Ingestion al tuo VPC per utilizzare Kafka autogestito. Per farlo, trova la sezione Configurazione di rete, seleziona la casella di controllo Collega a VPC e scegli il tuo CIDR da una delle opzioni predefinite fornite oppure selezionane una personalizzata. È possibile utilizzare qualsiasi CIDR da uno spazio di indirizzi privato come definito nella RFC 1918 Best Current Practice.
> Per fornire un CIDR personalizzato, seleziona Altro dal menu a discesa. Per evitare una collisione negli indirizzi IP tra OpenSearch Ingestion e Self-managed OpenSearch, assicurati che il CIDR OpenSearch VPC autogestito sia diverso dal CIDR per Ingestion. OpenSearch

5. Allega una policy basata sulle risorse al tuo dominio o una policy di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal cluster autogestito al dominio o alla raccolta.

   Il seguente esempio di policy di accesso al dominio consente al ruolo pipeline, creato nel passaggio successivo, di scrivere dati su un dominio. Assicurati di aggiornarlo `resource` con il tuo ARN.

   JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::444455556666:role/pipeline-role"
      },
      "Action": [
        "es:DescribeDomain",
        "es:ESHttp*"
      ],
      "Resource": [
        "arn:aws:es:us-east-1:111122223333:domain/domain-name"
      ]
    }
  ]
```

```
}
```

Per creare un ruolo IAM con le autorizzazioni corrette per accedere ai dati di scrittura nella raccolta o nel dominio, consulta. the section called “Configurazione di ruoli e utenti”

Fase 1: Configurare il ruolo della pipeline

Dopo aver impostato i prerequisiti della pipeline, configura il ruolo della pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi le seguenti autorizzazioni nel ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "SecretsManagerReadAccess",
            "Effect": "Allow",
            "Action": [
                "secretsmanager:GetSecretValue"
            ],
            "Resource": ["arn:aws:secretsmanager:us-
east-1:111122223333:secret:secret-name"]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:AttachNetworkInterface",
                "ec2:CreateNetworkInterface",
                "ec2:CreateNetworkInterfacePermission",
                "ec2:DeleteNetworkInterface",
                "ec2:DeleteNetworkInterfacePermission",
                "ec2:DetachNetworkInterface",
                "ec2:DescribeNetworkInterfaces"
            ],
            "Resource": [
                "arn:aws:ec2:*:*:network-interface/*",
                "arn:aws:ec2:*:*:subnet/*",
                "arn:aws:ec2:*:*:security-group/*"
            ]
        },
```

```
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeDhcpOptions",
                "ec2:DescribeRouteTables",
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeSubnets",
                "ec2:DescribeVpcs",
                "ec2:Describe*"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:CreateTags"
            ],
            "Resource": "arn:aws:ec2:*:*:network-interface/*",
            "Condition": {
               "StringEquals":
                    {
                        "aws:RequestTag/OSISManaged": "true"
                    }
            }
        }
    ]
}
```

Devi fornire le autorizzazioni Amazon EC2 di cui sopra sul ruolo IAM che usi per creare la pipeline di OpenSearch ingestione, poiché la pipeline utilizza queste autorizzazioni per creare ed eliminare un'interfaccia di rete nel tuo VPC. La pipeline può accedere al cluster Kafka solo tramite questa interfaccia di rete.

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestione come la seguente, che specifica Kafka come origine.

È possibile specificare più domini di OpenSearch servizio come destinazioni per i dati. Questa funzionalità consente il routing o la replica condizionale dei dati in entrata in più domini di servizio. OpenSearch

Puoi anche migrare i dati da un cluster Confluent Kafka di origine a una raccolta VPC serverless. OpenSearch Assicurati di fornire una politica di accesso alla rete all'interno della configurazione della pipeline. È possibile utilizzare un registro di schemi Confluent per definire uno schema Confluent.

```
 version: "2"
kafka-pipeline:
  source:
    kafka:
      encryption:
        type: "ssl"
      topics:
        - name: "topic-name"
          group_id: "group-id"
      bootstrap_servers:
        - "bootstrap-server.us-east-1.aws.private.confluent.cloud:9092"
      authentication:
        sasl:
          plain:
            username: ${aws_secrets:confluent-kafka-secret:username}
            password: ${aws_secrets:confluent-kafka-secret:password}
      schema:
        type: confluent
        registry_url: https://my-registry.us-east-1.aws.confluent.cloud
        api_key: "${{aws_secrets:schema-secret:schema_registry_api_key}}"
        api_secret: "${{aws_secrets:schema-secret:schema_registry_api_secret}}"
        basic_auth_credentials_source: "USER_INFO"
  sink:
  - opensearch:
      hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
      aws:
          region: "us-east-1"
      index: "confluent-index"
extension:
  aws:
    secrets:
      confluent-kafka-secret:
        secret_id: "my-kafka-secret"
        region: "us-east-1"
      schema-secret:
        secret_id: "my-self-managed-kafka-schema"
        region: "us-east-1"
```

È possibile utilizzare un blueprint preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Migrazione dei dati da OpenSearch cluster autogestiti con Amazon Ingestion OpenSearch

Puoi utilizzare una pipeline Amazon OpenSearch Ingestion con gestione automatica OpenSearch o Elasticsearch per migrare i dati verso domini Amazon Service e raccolte Serverless. OpenSearch OpenSearch OpenSearch Ingestion supporta configurazioni di rete pubbliche e private per la migrazione di dati da Elasticsearch e autogestiti. OpenSearch

Migrazione da cluster OpenSearch pubblici

Puoi utilizzare le pipeline di OpenSearch Ingestion per migrare i dati da un cluster autogestito OpenSearch o Elasticsearch con una configurazione pubblica, il che significa che il nome DNS del dominio può essere risolto pubblicamente. A tale scopo, configura una pipeline di OpenSearch ingestion con Elasticsearch OpenSearch o autogestita come origine e Service o Serverless come destinazione. OpenSearch OpenSearch Ciò consente di migrare efficacemente i dati da un cluster di origine autogestito a un dominio o una raccolta di destinazione gestiti. AWS

Prerequisiti

Prima di creare la pipeline di OpenSearch Ingestion, effettuate le seguenti operazioni:

1. Crea un cluster autogestito OpenSearch o Elastisearch che contenga i dati che desideri migrare e configura un nome DNS pubblico. Per istruzioni, consulta Creare un cluster nella documentazione. OpenSearch
2. Crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless verso cui migrare i dati. Per ulteriori informazioni, consultare the section called “Creazione OpenSearch di domini di servizio” e the section called “Creazione di raccolte”.
3. Configura l'autenticazione sul tuo cluster autogestito con. Gestione dei segreti AWS Abilita la rotazione dei segreti seguendo la procedura descritta in Ruota Gestione dei segreti AWS i segreti.
4. Allega una politica basata sulle risorse al tuo dominio o una politica di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal cluster autogestito al dominio o alla raccolta.

Il seguente esempio di policy di accesso al dominio consente al ruolo pipeline, creato nel passaggio successivo, di scrivere dati su un dominio. Assicurati di aggiornarlo `resource` con il tuo ARN.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::444455556666:role/pipeline-role"
      },
      "Action": [
        "es:DescribeDomain",
        "es:ESHttp*"
      ],
      "Resource": [
        "arn:aws:es:us-east-1:111122223333:domain/domain-name"
      ]
    }
  ]
}
```

Per creare un ruolo IAM con le autorizzazioni corrette per accedere ai dati di scrittura nella raccolta o nel dominio, consulta. the section called “Configurazione di ruoli e utenti”

Fase 1: Configurare il ruolo della pipeline

Dopo aver impostato i prerequisiti della OpenSearch pipeline, configura il ruolo della pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi l'autorizzazione alla scrittura su un dominio di OpenSearch servizio o una raccolta OpenSearch Serverless, nonché l'autorizzazione a leggere i segreti da Secrets Manager.

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestione come la seguente, che viene specificata OpenSearch come origine.

È possibile specificare più domini di OpenSearch servizio come destinazioni per i dati. Questa funzionalità consente il routing o la replica condizionale dei dati in entrata in più domini di servizio. OpenSearch

Puoi anche migrare i dati da un'origine OpenSearch o da un cluster Elasticsearch a una raccolta VPC serverless OpenSearch . Assicurati di fornire una politica di accesso alla rete all'interno della configurazione della pipeline.

```
version: "2"
opensearch-migration-pipeline:
  source:
    opensearch:
      acknowledgments: true
      host: [ "https://my-self-managed-cluster-name:9200" ]
      indices:
        include:
          - index_name_regex: "include-.*"
        exclude:
          - index_name_regex: '\..*'
      authentication:
        username: ${aws_secrets:secret:username}
        password: ${aws_secrets:secret:password}
        scheduling:
           interval: "PT2H"
           index_read_count: 3
           start_time: "2023-06-02T22:01:30.00Z"
  sink:
  - opensearch:
      hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
      aws:
          region: "us-east-1"
          #Uncomment the following lines if your destination is an OpenSearch
 Serverless collection
          #serverless: true
          # serverless_options:
          #     network_policy_name: "network-policy-name"
      index: "${getMetadata(\"opensearch-index\")}"
      document_id: "${getMetadata(\"opensearch-document_id\")}"
      enable_request_compression: true
      dlq:
        s3:
          bucket: "bucket-name"
          key_path_prefix: "apache-log-pipeline/logs/dlq"
```

```
          region: "us-east-1"
extension:
  aws:
    secrets:
      secret:
        secret_id: "my-opensearch-secret"
        region: "us-east-1"
        refresh_interval: PT1H
```

È possibile utilizzare un blueprint preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Migrazione dei dati dai OpenSearch cluster in un VPC

Puoi anche utilizzare le pipeline di OpenSearch Ingestion per migrare i dati da un cluster autogestito OpenSearch o Elasticsearch in esecuzione in un VPC. A tale scopo, configura una pipeline di OpenSearch ingestion con Elasticsearch OpenSearch o autogestita come origine e Service o Serverless come destinazione. OpenSearch OpenSearch Ciò consente di migrare efficacemente i dati da un cluster di origine autogestito a un dominio o una raccolta di destinazione gestiti. AWS

Prerequisiti

Prima di creare la pipeline di OpenSearch Ingestion, effettuate le seguenti operazioni:

1. Crea un cluster autogestito OpenSearch o Elastisearch con una configurazione di rete VPC che contenga i dati che desideri migrare.
2. Crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless verso cui migrare i dati. Per ulteriori informazioni, consulta Creazione di domini di OpenSearch servizio e Creazione di raccolte.
3. Configura l'autenticazione sul tuo cluster autogestito con. Gestione dei segreti AWS Abilita la rotazione dei segreti seguendo la procedura descritta in Ruota Gestione dei segreti AWS i segreti.
4. Ottieni l'ID del VPC che ha accesso a Elasticsearch o autogestito OpenSearch . Scegli il VPC CIDR da utilizzare con Ingestion. OpenSearch

> **Note**
>
> Se utilizzi il Console di gestione AWS per creare la tua pipeline, devi anche collegare la pipeline di OpenSearch Ingestion al tuo VPC per utilizzare Elasticsearch o autogestito. OpenSearch Per farlo, trova la sezione Opzioni di rete di origine, seleziona la casella di

controllo Collega a VPC e scegli il tuo CIDR da una delle opzioni predefinite fornite. È possibile utilizzare qualsiasi CIDR da uno spazio di indirizzi privato come definito nella Best Current Practice della RFC 1918.
Per fornire un CIDR personalizzato, seleziona Altro dal menu a discesa. Per evitare una collisione negli indirizzi IP tra OpenSearch Ingestion e Self-managed OpenSearch, assicurati che il CIDR OpenSearch VPC autogestito sia diverso dal CIDR per Ingestion. OpenSearch

5. Allega una policy basata sulle risorse al tuo dominio o una policy di accesso ai dati alla tua raccolta. Queste politiche di accesso consentono a OpenSearch Ingestion di scrivere dati dal cluster autogestito al dominio o alla raccolta.

   Il seguente esempio di policy di accesso al dominio consente al ruolo pipeline, creato nel passaggio successivo, di scrivere dati su un dominio. Assicurati di aggiornarlo `resource` con il tuo ARN.

   JSON

   ```
   {
     "Version":"2012-10-17",
     "Statement": [
       {
         "Effect": "Allow",
         "Principal": {
           "AWS": "arn:aws:iam::444455556666:role/pipeline-role"
         },
         "Action": [
           "es:DescribeDomain",
           "es:ESHttp*"
         ],
         "Resource": [
           "arn:aws:es:us-east-1:111122223333:domain/example.com"
         ]
       }
     ]
   }
   ```

   Per creare un ruolo IAM con le autorizzazioni corrette per accedere ai dati di scrittura nella raccolta o nel dominio, consulta. the section called “Configurazione di ruoli e utenti”

Fase 1: Configurare il ruolo della pipeline

Dopo aver impostato i prerequisiti della pipeline, configura il ruolo della pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi le seguenti autorizzazioni nel ruolo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "SecretsManagerReadAccess",
            "Effect": "Allow",
            "Action": [
                "secretsmanager:GetSecretValue"
            ],
            "Resource": ["arn:aws:secretsmanager:us-
east-1:111122223333:secret:secret-name"]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:AttachNetworkInterface",
                "ec2:CreateNetworkInterface",
                "ec2:CreateNetworkInterfacePermission",
                "ec2:DeleteNetworkInterface",
                "ec2:DeleteNetworkInterfacePermission",
                "ec2:DetachNetworkInterface",
                "ec2:DescribeNetworkInterfaces"
            ],
            "Resource": [
                "arn:aws:ec2:*:*:network-interface/*",
                "arn:aws:ec2:*:*:subnet/*",
                "arn:aws:ec2:*:*:security-group/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeDhcpOptions",
                "ec2:DescribeRouteTables",
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeSubnets",
```

```
                "ec2:DescribeVpcs",
                "ec2:Describe*"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:CreateTags"
            ],
            "Resource": "arn:aws:ec2:*:*:network-interface/*",
            "Condition": {
               "StringEquals":
                    {
                        "aws:RequestTag/OSISManaged": "true"
                    }
            }
        }
    ]
}
```

Devi fornire le autorizzazioni Amazon EC2 di cui sopra sul ruolo IAM che usi per creare la pipeline di OpenSearch ingestione, poiché la pipeline utilizza queste autorizzazioni per creare ed eliminare un'interfaccia di rete nel tuo VPC. La pipeline può accedere al cluster solo tramite questa interfaccia di rete. OpenSearch

Fase 2: Creare la pipeline

È quindi possibile configurare una pipeline OpenSearch di ingestione come la seguente, che viene specificata OpenSearch come origine.

È possibile specificare più domini di OpenSearch servizio come destinazioni per i dati. Questa funzionalità consente il routing o la replica condizionale dei dati in entrata in più domini di servizio. OpenSearch

Puoi anche migrare i dati da un'origine OpenSearch o da un cluster Elasticsearch a una raccolta VPC serverless OpenSearch . Assicurati di fornire una politica di accesso alla rete all'interno della configurazione della pipeline.

```
version: "2"
opensearch-migration-pipeline:
```

```
  source:
    opensearch:
      acknowledgments: true
      host: [ "https://my-self-managed-cluster-name:9200" ]
      indices:
        include:
          - index_name_regex: "include-.*"
        exclude:
          - index_name_regex: '\..*'
      authentication:
        username: ${aws_secrets:secret:username}
        password: ${aws_secrets:secret:password}
        scheduling:
           interval: "PT2H"
           index_read_count: 3
           start_time: "2023-06-02T22:01:30.00Z"
  sink:
  - opensearch:
      hosts: ["https://search-mydomain.us-east-1.es.amazonaws.com"]
      aws:
          region: "us-east-1"
          #Uncomment the following lines if your destination is an OpenSearch
 Serverless collection
          #serverless: true
          # serverless_options:
          #     network_policy_name: "network-policy-name"
      index: "${getMetadata(\"opensearch-index\")}"
      document_id: "${getMetadata(\"opensearch-document_id\")}"
      enable_request_compression: true
      dlq:
        s3:
          bucket: "bucket-name"
          key_path_prefix: "apache-log-pipeline/logs/dlq"
          region: "us-east-1"
extension:
  aws:
    secrets:
      secret:
        secret_id: "my-opensearch-secret"
        region: "us-east-1"
        refresh_interval: PT1H
```

È possibile utilizzare un blueprint preconfigurato per creare questa pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

Usa una pipeline OpenSearch di ingestione con Amazon Kinesis Data Streams

Utilizza una pipeline di OpenSearch ingestione con Amazon Kinesis Data Streams per importare i dati dei record di flusso da più flussi in domini e raccolte di Amazon Service. OpenSearch La pipeline OpenSearch Ingestion incorpora l'infrastruttura di inserimento dello streaming per fornire un modo su larga scala e a bassa latenza per importare continuamente i record di streaming da Kinesis.

Argomenti

- Amazon Kinesis Data Streams come fonte
- Amazon Kinesis Data Streams su più account come fonte

Amazon Kinesis Data Streams come fonte

Con la procedura seguente, imparerai come configurare una pipeline di OpenSearch ingestione che utilizza Amazon Kinesis Data Streams come origine dati. Questa sezione descrive i prerequisiti necessari, come la creazione di un dominio di OpenSearch servizio o una collezione OpenSearch Serverless, e illustra i passaggi per configurare il ruolo della pipeline e creare la pipeline.

Prerequisiti

Per configurare la pipeline, sono necessari uno o più Kinesis Data Streams attivi. Questi flussi devono ricevere record o essere pronti a ricevere record da altre fonti. Per ulteriori informazioni, vedere Panoramica sull' OpenSearch ingestione.

Per configurare la tua pipeline

1. Crea un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless

   Per creare un dominio o una raccolta, consulta Guida introduttiva a OpenSearch Ingestion.

   Per creare un ruolo IAM con le autorizzazioni corrette per accedere ai dati di scrittura nella raccolta o nel dominio, consulta Politiche basate sulle risorse.

2. Configura il ruolo della pipeline con le autorizzazioni

   Imposta il ruolo della pipeline che desideri utilizzare nella configurazione della pipeline e aggiungi le seguenti autorizzazioni. Sostituisci *placeholder values* con le informazioni appropriate.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "allowReadFromStream",
            "Effect": "Allow",
            "Action": [
                "kinesis:DescribeStream",
                "kinesis:DescribeStreamConsumer",
                "kinesis:DescribeStreamSummary",
                "kinesis:GetRecords",
                "kinesis:GetShardIterator",
                "kinesis:ListShards",
                "kinesis:ListStreams",
                "kinesis:ListStreamConsumers",
                "kinesis:RegisterStreamConsumer",
                "kinesis:SubscribeToShard"
            ],
            "Resource": [
                "arn:aws:kinesis:us-east-1:111122223333:stream/stream-name"
            ]
        }
    ]
}
```

   Se la crittografia lato server è abilitata negli stream, la seguente AWS KMS politica consente di decrittografare i record. Sostituisci *placeholder values* con le informazioni appropriate.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "allowDecryptionOfCustomManagedKey",
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt",
                "kms:GenerateDataKey"
            ],
            "Resource": "arn:aws:kms:us-east-1:111122223333:key/key-id"
        }
    ]
}
```

Affinché una pipeline possa scrivere dati su un dominio, il dominio deve disporre di una politica di accesso a livello di dominio che consenta al ruolo della pipeline sts_role_arn di accedervi.

L'esempio seguente è una politica di accesso al dominio che consente al ruolo pipeline creato nel passaggio precedente (), di scrivere dati nel dominio. `pipeline-role` `ingestion-domain` Sostituisci *`placeholder values`* con le informazioni appropriate.

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::your-account-id:role/pipeline-role"
      },
      "Action": ["es:DescribeDomain", "es:ESHttp*"],
      "Resource": "arn:aws:es:Regione AWS:account-id:domain/domain-name/*"
    }
  ]
}
```

3. Crea la pipeline

Configura una pipeline OpenSearch di ingestione specificando K come origine. inesis-data-streams È possibile individuare un blueprint già pronto e disponibile nella OpenSearch Ingestion Console per creare una pipeline di questo tipo. (Facoltativo) Per creare la pipeline utilizzando AWS CLI, è possibile utilizzare un blueprint denominato "». **`AWS-KinesisDataStreamsPipeline`** Sostituisci *`placeholder values`* con le informazioni appropriate.

```
version: "2"
kinesis-pipeline:
  source:
    kinesis_data_streams:
      acknowledgments: true
      codec:
        # Based on whether kinesis records are aggregated or not, you could choose
 json, newline or ndjson codec for processing the records.
        # JSON codec supports parsing nested CloudWatch Events into individual log
 entries that will be written as documents into OpenSearch.
        # json:
          # key_name: "logEvents"
          # These keys contain the metadata sent by CloudWatch Subscription Filters
          # in addition to the individual log events:
          # include_keys: [ 'owner', 'logGroup', 'logStream' ]
        newline:
      streams:
        - stream_name: "stream name"
          # Enable this if ingestion should start from the start of the stream.
          # initial_position: "EARLIEST"
          # checkpoint_interval: "PT5M"
          # Compression will always be gzip for CloudWatch, but will vary for other
 sources:
          # compression: "gzip"
        - stream_name: "stream name"
          # Enable this if ingestion should start from the start of the stream.
          # initial_position: "EARLIEST"
          # checkpoint_interval: "PT5M"
          # Compression will always be gzip for CloudWatch, but will vary for other
 sources:
          # compression: "gzip"
```

```
        # buffer_timeout: "1s"
        # records_to_accumulate: 100
        # Change the consumer strategy to "polling". Default consumer strategy will
 use enhanced "fan-out" supported by KDS.
        # consumer_strategy: "polling"
        # if consumer strategy is set to "polling", enable the polling config
 below.
        # polling:
          # max_polling_records: 100
          # idle_time_between_reads: "250ms"
      aws:
        # Provide the Role ARN with access to Amazon Kinesis Data Streams. This
 role should have a trust relationship with osis-pipelines.amazonaws.com
        sts_role_arn: "arn:aws:iam::111122223333:role/Example-Role"
        # Provide the Regione AWS of the Data Stream.
        region: "us-east-1"

  sink:
    - opensearch:
        # Provide an Amazon OpenSearch Serverless domain endpoint
        hosts: [ "https://search-mydomain-1a2a3a4a5a6a7a8a9a0a9a8a7a.us-
east-1.es.amazonaws.com" ]
        index: "index_${getMetadata(\"stream_name\")}"
        # Ensure adding unique document id as a combination of the metadata
 attributes available.
        document_id: "${getMetadata(\"partition_key\")}_
${getMetadata(\"sequence_number\")}_${getMetadata(\"sub_sequence_number\")}"
        aws:
          # Provide a Role ARN with access to the domain. This role should have a
 trust relationship with osis-pipelines.amazonaws.com
          sts_role_arn: "arn:aws:iam::111122223333:role/Example-Role"
          # Provide the Regione AWS of the domain.
          region: "us-east-1"
          # Enable the 'serverless' flag if the sink is an Amazon OpenSearch
 Serverless collection
          serverless: false
          # serverless_options:
            # Specify a name here to create or update network policy for the
 serverless collection
            # network_policy_name: "network-policy-name"
        # Enable the 'distribution_version' setting if the OpenSearch Serverless
 domain is of version Elasticsearch 6.x
        # distribution_version: "es6"
```

```
        # Enable and switch the 'enable_request_compression' flag if the default
 compression setting is changed in the domain. See https://docs.aws.amazon.com/
opensearch-service/latest/developerguide/gzip.html
        # enable_request_compression: true/false
        # Optional: Enable the S3 DLQ to capture any failed requests in an S3
 bucket. Delete this entire block if you don't want a DLQ.
        dlq:
          s3:
            # Provide an S3 bucket
            bucket: "your-dlq-bucket-name"
            # Provide a key path prefix for the failed requests
            # key_path_prefix: "kinesis-pipeline/logs/dlq"
            # Provide the region of the bucket.
            region: "us-east-1"
            # Provide a Role ARN with access to the bucket. This role should have a
 trust relationship with osis-pipelines.amazonaws.com
            sts_role_arn: "arn:aws:iam::111122223333:role/Example-Role"
```

Opzioni di configurazione

Per le opzioni di configurazione di Kinesis, consulta Opzioni di configurazione nella OpenSearchdocumentazione.

Attributi dei metadati disponibili

- stream_name — Nome del Kinesis Data Streams da cui è stato importato il record
- partition_key — Chiave di partizione del record Kinesis Data Streams che viene importato
- sequence_number — Numero di sequenza del record Kinesis Data Streams che viene importato
- sub_sequence_number — Numero di sottosequenza del record Kinesis Data Streams che viene importato

4. (Facoltativo) Configura le unità di calcolo consigliate (OCUs) per la pipeline Kinesis Data Streams

   Una pipeline di origine OpenSearch Kinesis Data Streams può anche essere configurata per importare record di stream da più di un flusso. Per un'acquisizione più rapida, ti consigliamo di aggiungere un'unità di calcolo aggiuntiva per ogni nuovo flusso aggiunto.

Coerenza dei dati

OpenSearch Ingestion supporta il end-to-end riconoscimento per garantire la durabilità dei dati. Quando la pipeline legge i record dei flussi da Kinesis, distribuisce dinamicamente il lavoro di lettura dei record dei flussi in base agli shard associati agli stream. Pipeline effettuerà automaticamente il checkpoint dei flussi quando riceve una conferma dopo aver acquisito tutti i record del dominio o della raccolta. OpenSearch Ciò eviterà l'elaborazione duplicata dei record di stream.

Per creare l'indice in base al nome dello stream, definisci l'indice nella sezione opensearch sink come «index_$ {getMetaData (\" stream_name\»)}».

Amazon Kinesis Data Streams su più account come fonte

Puoi concedere l'accesso a più account con Amazon Kinesis Data Streams in OpenSearch modo che le pipeline di ingestione possano accedere a Kinesis Data Streams in un altro account come origine. Completa i seguenti passaggi per abilitare l'accesso tra più account:

Configurare l'accesso tra più account

1. Imposta la politica delle risorse nell'account che ha il flusso Kinesis

   Sostituisci *placeholder values* con le informazioni appropriate.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "StreamReadStatementID",
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::111122223333:role/Pipeline-Role"
            },
            "Action": [
                "kinesis:DescribeStreamSummary",
                "kinesis:GetRecords",
                "kinesis:GetShardIterator",
                "kinesis:ListShards"
            ],
```

```
            "Resource": "arn:aws:kinesis:us-east-1:444455556666:stream/
stream-name"
        },
        {
            "Sid": "StreamEFOReadStatementID",
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::111122223333:role/Pipeline-Role"
            },
            "Action": [
                "kinesis:DescribeStreamSummary",
                "kinesis:ListShards"
            ],
            "Resource": "arn:aws:kinesis:us-
east-1:444455556666:stream/stream-name/consumer/consumer-name"
        }
    ]
}
```

2. (Facoltativo) Imposta la politica relativa alle risorse destinate ai consumatori e ai consumatori

   Questo passaggio è facoltativo e sarà necessario solo se prevedi di utilizzare la strategia Enhanced Fanout Consumer per leggere i record degli stream. Per ulteriori informazioni, consulta Sviluppare utenti fan-out avanzati con throughput dedicato.

   a. Configura il consumatore

      Per riutilizzare un consumatore esistente, puoi saltare questo passaggio. Per ulteriori informazioni, consulta RegisterStreamConsumerAmazon Kinesis Data Streams API Reference.

      Nel seguente comando CLI di esempio, sostituisci il *placeholder values* con le tue informazioni.

      Example: Esempio di comando CLI

```
aws kinesis register-stream-consumer \
--stream-arn "arn:aws:kinesis:Regione AWS:account-id:stream/stream-name" \
--consumer-name consumer-name
```

b. Imposta la politica delle risorse per i consumatori

Nella seguente dichiarazione, sostituiscili *placeholder values* con le tue informazioni.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "ConsumerEFOReadStatementID",
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::111122223333:role/Pipeline-Role"
            },
            "Action": [
                "kinesis:DescribeStreamConsumer",
                "kinesis:SubscribeToShard"
            ],
            "Resource": "arn:aws:kinesis:us-
east-1:444455556666:stream/stream-1/consumer/consumer-name"
        }
    ]
}
```

3. Configurazione della pipeline

Per l'ingestione tra più account, aggiungi i seguenti attributi sotto `kinesis_data_streams` per ogni stream:

- `stream_arn`- l'arn dello stream appartenente all'account in cui esiste lo stream
- `consumer_arn`- questo è un attributo facoltativo e deve essere specificato se viene scelta la strategia di consumo fanout avanzata predefinita. Specificare l'effettivo valore del consumatore per questo campo. Sostituisci *placeholder values* con le informazioni appropriate.

```
version: "2"
     kinesis-pipeline:
```

```
      source:
        kinesis_data_streams:
          acknowledgments: true
          codec:
            newline:
          streams:
            - stream_arn: "arn:aws:kinesis:region:stream-account-id:stream/stream-
name"
              consumer_arn: "consumer arn"
              # Enable this if ingestion should start from the start of the
 stream.
              # initial_position: "EARLIEST"
              # checkpoint_interval: "PT5M"
            - stream_arn: "arn:aws:kinesis:region:stream-account-id:stream/stream-
name"
              consumer_arn: "consumer arn"
               # initial_position: "EARLIEST"

          # buffer_timeout: "1s"
          # records_to_accumulate: 100
          # Enable the consumer strategy to "polling". Default consumer strategy
 will use enhanced "fan-out" supported by KDS.
          # consumer_strategy: "polling"
          # if consumer strategy is set to "polling", enable the polling config
 below.
          # polling:
            # max_polling_records: 100
            # idle_time_between_reads: "250ms"
          aws:
            # Provide the Role ARN with access to Kinesis. This role should have a
 trust relationship with osis-pipelines.amazonaws.com
            sts_role_arn: "arn:aws:iam::111122223333:role/Example-Role"
            # Provide the Regione AWS of the domain.
            region: "us-east-1"

      sink:
        - opensearch:
            # Provide an OpenSearch Serverless domain endpoint
            hosts: [ "https://search-mydomain-1a2a3a4a5a6a7a8a9a0a9a8a7a.us-
east-1.es.amazonaws.com" ]
            index: "index_${getMetadata(\"stream_name\")}"
            # Mapping for documentid based on partition key, shard sequence number
 and subsequence number metadata attributes
```

```
            document_id: "${getMetadata(\"partition_key\")}_
${getMetadata(\"sequence_number\")}_${getMetadata(\"sub_sequence_number\")}"
            aws:
              # Provide a Role ARN with access to the domain. This role should
 have a trust relationship with osis-pipelines.amazonaws.com
              sts_role_arn: "arn:aws:iam::111122223333:role/Example-Role"
              # Provide the Regione AWS of the domain.
              region: "us-east-1"
              # Enable the 'serverless' flag if the sink is an OpenSearch
 Serverless collection
              serverless: false
                # serverless_options:
                # Specify a name here to create or update network policy for the
 serverless collection
              # network_policy_name: network-policy-name
            # Enable the 'distribution_version' setting if the OpenSearch
 Serverless domain is of version Elasticsearch 6.x
            # distribution_version: "es6"
            # Enable and switch the 'enable_request_compression' flag if
 the default compression setting is changed in the domain. See https://
docs.aws.amazon.com/opensearch-service/latest/developerguide/gzip.html
            # enable_request_compression: true/false
            # Optional: Enable the S3 DLQ to capture any failed requests in an S3
 bucket. Delete this entire block if you don't want a DLQ.
            dlq:
              s3:
                # Provide an Amazon S3 bucket
                bucket: "your-dlq-bucket-name"
                # Provide a key path prefix for the failed requests
                # key_path_prefix: "alb-access-log-pipeline/logs/dlq"
                # Provide the Regione AWS of the bucket.
                region: "us-east-1"
                # Provide a Role ARN with access to the bucket. This role should
 have a trust relationship with osis-pipelines.amazonaws.com
                sts_role_arn: "arn:aws:iam::111122223333:role/Example-Role"
```

4. Ruolo della pipeline OSI Kinesis Data Streams

   a. Politica IAM

      Aggiungi la seguente politica al ruolo della pipeline. Sostituisci *placeholder values* con le informazioni appropriate.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kinesis:DescribeStreamConsumer",
                "kinesis:SubscribeToShard"
            ],
            "Resource": [
            "arn:aws:kinesis:us-east-1:111122223333:stream/my-stream"
            ]
        },
        {
            "Sid": "allowReadFromStream",
            "Effect": "Allow",
            "Action": [
                "kinesis:DescribeStream",
                "kinesis:DescribeStreamSummary",
                "kinesis:GetRecords",
                "kinesis:GetShardIterator",
                "kinesis:ListShards",
                "kinesis:ListStreams",
                "kinesis:ListStreamConsumers",
                "kinesis:RegisterStreamConsumer"
            ],
            "Resource": [
                "arn:aws:kinesis:us-east-1:111122223333:stream/my-stream"
            ]
        }
    ]
}
```

b. Policy di attendibilità

Per importare i dati dall'account di streaming, è necessario stabilire una relazione di fiducia tra il ruolo di importazione della pipeline e l'account di streaming. Aggiungi quanto segue al ruolo pipeline. Sostituisci *placeholder values* con le informazioni appropriate.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
     "Effect": "Allow",
     "Principal": {
       "AWS": "arn:aws:iam::111122223333:root"
      },
     "Action": "sts:AssumeRole"
  }]
}
```

Fasi successive

Dopo aver esportato i dati in una pipeline, puoi interrogarli dal dominio di OpenSearch servizio configurato come sink per la pipeline. Le seguenti risorse possono aiutarti a iniziare:

- Osservabilità
- the section called “Scopri le tracce”
- Osservabilità

Utilizzo di una pipeline OpenSearch di ingestione con AWS Lambda

Utilizza il AWS Lambda processore per arricchire i dati da qualsiasi fonte o destinazione supportata da OpenSearch Ingestion utilizzando codice personalizzato. Con il processore Lambda, puoi applicare le tue trasformazioni o arricchimenti dei dati e quindi restituire gli eventi elaborati alla pipeline per un'ulteriore elaborazione. Questo processore consente l'elaborazione personalizzata dei dati e offre il pieno controllo sul modo in cui i dati vengono manipolati prima che entrino nella pipeline.

> **Note**
>
> Il limite di dimensione del payload per un singolo evento elaborato da un processore Lambda è di 5 MB. Inoltre, il processore Lambda supporta solo risposte in formato array JSON.

Prerequisiti

Prima di creare una pipeline con un processore Lambda, create le seguenti risorse:

- Una AWS Lambda funzione che arricchisce e trasforma i dati di origine. Per istruzioni, consulta Creare la tua prima funzione Lambda.
- Un dominio OpenSearch di servizio o una raccolta OpenSearch Serverless che fungerà da sink della pipeline. Per ulteriori informazioni, consultare the section called “Creazione OpenSearch di domini di servizio” e the section called “Creazione di raccolte”.
- Un ruolo della pipeline che include le autorizzazioni di scrittura nel dominio o nel sink di raccolta. Per ulteriori informazioni, consulta the section called “Ruolo della pipeline”.

  Il ruolo pipeline necessita inoltre di una politica di autorizzazioni allegata che gli consenta di richiamare la funzione Lambda specificata nella configurazione della pipeline. Esempio:

  JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "allowinvokeFunction",
            "Effect": "Allow",
            "Action": [
                "lambda:invokeFunction",
                "lambda:InvokeAsync",
                "lambda:ListFunctions"
            ],
            "Resource": "arn:aws:lambda:us-
east-1:111122223333:function:function-name"

        }
    ]
}
```

Crea una pipeline

Per utilizzarlo AWS Lambda come processore, configura una pipeline di OpenSearch ingestione e specificala come processore. `aws_lambda` È inoltre possibile utilizzare il blueprint di arricchimento AWS Lambda personalizzato per creare la pipeline. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

La pipeline di esempio seguente riceve dati da un'origine HTTP, li arricchisce utilizzando un elaboratore di dati e il AWS Lambda processore e inserisce i dati elaborati in un dominio. OpenSearch

```
version: "2"
lambda-processor-pipeline:
  source:
    http:
      path: "/${pipelineName}/logs"
  processor:
      - date:
        destination: "@timestamp"
        from_time_received: true
    - aws_lambda:
        function_name: "my-lambda-function"

        tags_on_failure: ["lambda_failure"]
        batch:
            key_name: "events"
        aws:
          region: region
  sink:
    - opensearch:
        hosts: [ "https://search-mydomain.us-east-1es.amazonaws.com" ]
        index: "table-index"
        aws:
          region: "region"
          serverless: false
```

La AWS Lambda funzione di esempio seguente trasforma i dati in entrata aggiungendo una nuova coppia chiave-valore (`"transformed": "true"`) a ciascun elemento nell'array di eventi fornito, quindi restituisce la versione modificata.

```
import json
```

```
def lambda_handler(event, context):
    input_array = event.get('events', [])
    output = []
    for input in input_array:
        input["transformed"] = "true";
        output.append(input)

    return output
```

Batching

Le pipeline inviano eventi in batch al processore Lambda e regolano dinamicamente la dimensione del batch per garantire che rimanga al di sotto del limite di 5 MB.

Di seguito è riportato un esempio di batch di pipeline:

```
batch:
    key_name: "events"

input_arrary = event.get('events', [])
```

> **Note**
>
> Quando crei una pipeline, assicurati che l'`key_name`opzione nella configurazione del processore Lambda corrisponda alla chiave dell'evento nel gestore Lambda.

Filtraggio condizionale

Il filtraggio condizionale consente di controllare quando il AWS Lambda processore richiama la funzione Lambda in base a condizioni specifiche nei dati degli eventi. Ciò è particolarmente utile quando si desidera elaborare selettivamente determinati tipi di eventi ignorandone altri.

La seguente configurazione di esempio utilizza il filtro condizionale:

```
processors:
  - aws_lambda:
      function_name: "my-lambda-function"
      aws:
        region: "region"
      lambda_when: "/sourceIp == 10.10.10.10"
```

Migrazione dei dati tra domini e raccolte utilizzando Amazon Ingestion OpenSearch

Puoi utilizzare le pipeline OpenSearch di Ingestion per migrare i dati tra domini Amazon OpenSearch Service o OpenSearch raccolte VPC Serverless. A tale scopo, configuri una pipeline in cui configuri un dominio o una raccolta come origine e un altro dominio o raccolta come sink. Questo consente di migrare efficacemente i dati da un dominio o una raccolta all'altra.

Per migrare i dati, è necessario disporre delle seguenti risorse:

- Un dominio di OpenSearch servizio di origine o una raccolta VPC OpenSearch Serverless. Questo dominio o raccolta contiene i dati che desideri migrare. Se utilizzi un dominio, deve eseguire la OpenSearch versione 1.0 o successiva oppure Elasticsearch versione 7.4 o successiva. Il dominio deve inoltre avere una politica di accesso che conceda le autorizzazioni appropriate al tuo ruolo di pipeline.
- Un dominio o una raccolta VPC separata verso cui migrare i dati. Questo dominio o raccolta fungerà da serbatoio della pipeline.
- Un ruolo della pipeline che OpenSearch Ingestion utilizzerà per leggere e scrivere nella raccolta o nel dominio. Includi l'Amazon Resource Name (ARN) di questo ruolo nella configurazione della pipeline. Per maggiori informazioni, consulta le seguenti risorse:
  - the section called “Concedere alle pipeline l'accesso ai domini”
  - the section called “Concedere alle pipeline l'accesso alle raccolte”

Argomenti

- Limitazioni
- OpenSearch Il servizio come fonte
- Specificazione di più sink di dominio di servizio OpenSearch
- Migrazione dei dati verso una raccolta OpenSearch VPC serverless

Limitazioni

Le seguenti limitazioni si applicano quando si designano domini di OpenSearch servizio o raccolte OpenSearch Serverless come sink:

- Una pipeline non può scrivere su più di un dominio VPC.

- Puoi migrare i dati solo da o verso raccolte OpenSearch Serverless che utilizzano l'accesso VPC. Le raccolte pubbliche non sono supportate.
- Non è possibile specificare una combinazione di VPC e domini pubblici in una configurazione a pipeline singola.
- È possibile avere un massimo di 20 sink non collegati alla pipeline all'interno di una singola configurazione di pipeline.
- È possibile specificare i sink scegliendo tra un massimo di tre tipi diversi Regioni AWS in una configurazione a tubazione singola.
- Una pipeline con più sink potrebbe subire una riduzione della velocità di elaborazione nel tempo se uno dei sink rimane inattivo per troppo tempo o non dispone di una capacità sufficiente per ricevere i dati in entrata.

OpenSearch Il servizio come fonte

Il dominio o la raccolta che specifichi come origine è il luogo da cui vengono migrati i dati.

Creazione di un ruolo di pipeline in IAM

Per creare la pipeline di OpenSearch Ingestion, devi prima creare un ruolo di pipeline per concedere l'accesso in lettura e scrittura tra domini o raccolte. A tal fine, segui questa procedura:

1. Crea una nuova politica di autorizzazioni in IAM da collegare al ruolo della pipeline. Assicurati di consentire le autorizzazioni per leggere dalla fonte e scrivere nel sink. Per ulteriori informazioni sull'impostazione delle autorizzazioni della pipeline IAM per i domini di OpenSearch servizio, consulta e. the section called “Concedere alle pipeline l'accesso ai domini” the section called “Concedere alle pipeline l'accesso alle raccolte”
2. Specificate le seguenti autorizzazioni all'interno del ruolo della pipeline da leggere dalla fonte:

   JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Action":"es:ESHttpGet",
         "Resource":[
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/",
```

```
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/_cat/
indices",
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/_search",
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/_search/
scroll",
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/*/_search"
         ]
      },
      {
         "Effect":"Allow",
         "Action":"es:ESHttpPost",
         "Resource":[
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/*/_search/
point_in_time",
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/*/_search/
scroll"
         ]
      },
      {
         "Effect":"Allow",
         "Action":"es:ESHttpDelete",
         "Resource":[
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/_search/
point_in_time",
            "arn:aws:es:us-east-1:111122223333:domain/domain-name/_search/
scroll"
         ]
      }
   ]
}
```

Creare una pipeline

Dopo aver associato la policy al ruolo della pipeline, utilizza il blueprint di AWSOpenSearchDataMigrationPipelinemigrazione per creare la pipeline. Questo blueprint include una configurazione predefinita per la migrazione dei dati tra OpenSearch domini o raccolte di servizi. Per ulteriori informazioni, consulta the section called “Lavorare con i progetti”.

> **Note**
>
> OpenSearch Ingestion utilizza la versione e la distribuzione del dominio di origine per determinare il meccanismo da utilizzare per la migrazione. Alcune versioni supportano l'opzione. `point_in_time` OpenSearch Serverless utilizza l'`search_after`opzione perché non supporta `point_in_time` `oscroll`.

È possibile che durante il processo di migrazione siano in corso di creazione nuovi indici oppure che i documenti vengano aggiornati durante la migrazione. Per questo motivo, potrebbe essere necessario eseguire una scansione singola o più scansioni dei dati dell'indice di dominio per raccogliere dati nuovi o aggiornati.

Specificate il numero di scansioni da eseguire configurando la configurazione `index_read_count` e `interval` nella pipeline. L'esempio seguente mostra come eseguire scansioni multiple:

```
scheduling:
    interval: "PT2H"
    index_read_count: 3
    start_time: "2023-06-02T22:01:30.00Z"
```

OpenSearch Ingestion utilizza la seguente configurazione per garantire che i dati vengano scritti nello stesso indice e mantengano lo stesso ID del documento:

```
index: "${getMetadata(\"opensearch-index\")}"
document_id: "${getMetadata(\"opensearch-document_id\")}"
```

Specificazione di più sink di dominio di servizio OpenSearch

È possibile specificare più domini di OpenSearch servizio pubblici come destinazioni per i dati. È possibile utilizzare questa funzionalità per eseguire il routing condizionale o replicare i dati in entrata in più domini di servizio. OpenSearch È possibile specificare fino a 10 diversi domini di servizio pubblici OpenSearch come sink.

Nell'esempio seguente, i dati in entrata vengono instradati in modo condizionale a diversi domini di servizio: OpenSearch

```
...
```

```
  route:
    - 2xx_status: "/response >= 200 and /response < 300"
    - 5xx_status: "/response >= 500 and /response < 600"
  sink:
    - opensearch:
        hosts: [ "https://search-response-2xx.region.es.amazonaws.com" ]
        aws:
          region: "us-east-1"
        index: "response-2xx"
        routes:
          - 2xx_status
    - opensearch:
        hosts: [ "https://search-response-5xx.region.es.amazonaws.com" ]
        aws:
          region: "us-east-1"
        index: "response-5xx"
        routes:
          - 5xx_status
```

Migrazione dei dati verso una raccolta OpenSearch VPC serverless

È possibile utilizzare OpenSearch Ingestion per migrare i dati da un dominio di OpenSearch servizio di origine o da una raccolta OpenSearch Serverless a un sink di raccolta VPC. È necessario fornire una politica di accesso alla rete all'interno della configurazione della pipeline. Per ulteriori informazioni sull'inserimento di dati nelle raccolte VPC OpenSearch serverless, vedere. the section called “Tutorial: inserisci dati in una raccolta”

Per migrare i dati in una raccolta VPC

1. Crea una raccolta OpenSearch Serverless. Per istruzioni, consulta the section called “Tutorial: inserisci dati in una raccolta”.
2. Crea una politica di rete per la raccolta che specifichi l'accesso VPC sia all'endpoint di raccolta che all'endpoint Dashboards. Per istruzioni, consulta the section called “Accesso alla rete”.
3. Crea il ruolo pipeline se non ne hai già uno. Per istruzioni, consulta the section called “Ruolo della pipeline”.
4. Crea la pipeline. Per istruzioni, consulta the section called “Lavorare con i progetti”.

Utilizzo di AWS SDKs per interagire con Amazon OpenSearch Ingestion

Questa sezione include un esempio di come utilizzare il AWS SDKs per interagire con Amazon OpenSearch Ingestion. L'esempio di codice dimostra come creare un dominio e una pipeline e quindi inserire dati nella pipeline.

Argomenti

- Python

Python

Lo script di esempio seguente utilizza AWS SDK per Python (Boto3)per creare un ruolo della pipeline IAM, un dominio in cui scrivere i dati e una pipeline attraverso cui importare i dati. Quindi inserisce un file di registro di esempio nella pipeline utilizzando la libreria HTTP. `requests`

Per installare le dipendenze richieste, esegui i seguenti comandi:

```
pip install boto3
pip install botocore
pip install requests
pip install requests-auth-aws-sigv4
```

All'interno dello script, sostituisci tutte le istanze di *`account-id`* con il tuo ID. Account AWS

```
import boto3
import botocore
from botocore.config import Config
import requests
from requests_auth_aws_sigv4 import AWSSigV4
import time

# Build the client using the default credential configuration.
# You can use the CLI and run 'aws configure' to set access key, secret
# key, and default region.

opensearch = boto3.client('opensearch', config=my_config)
iam = boto3.client('iam', config=my_config)
osis = boto3.client('osis', config=my_config)
```

```
domainName = 'test-domain'  # The name of the domain
pipelineName = 'test-pipeline' # The name of the pipeline

def createPipelineRole(iam, domainName):
    """Creates the pipeline role"""
    response = iam.create_policy(
        PolicyName='pipeline-policy',
        PolicyDocument=f'{{\"Version\":\"2012-10-17\",\"Statement\":[{{\"Effect\":
\"Allow\",\"Action\":\"es:DescribeDomain\",\"Resource\":\"arn:aws:es:us-east-1:account-
id:domain\/{domainName}\"}},{{\"Effect\":\"Allow\",\"Action\":\"es:ESHttp*\",\"Resource
\":\"arn:aws:es:us-east-1:account-id:domain\/{domainName}\/*\"}}]}}'
    )
    policyarn = response['Policy']['Arn']

    response = iam.create_role(
        RoleName='PipelineRole',
        AssumeRolePolicyDocument='{\"Version\":\"2012-10-17\",\"Statement\":[{\"Effect
\":\"Allow\",\"Principal\":{\"Service\":\"osis-pipelines.amazonaws.com\"},\"Action\":
\"sts:AssumeRole\"}]}'
    )
    rolename=response['Role']['RoleName']

    response = iam.attach_role_policy(
        RoleName=rolename,
        PolicyArn=policyarn
    )

    print('Creating pipeline role...')
    time.sleep(10)
    print('Role created: ' + rolename)

def createDomain(opensearch, domainName):
    """Creates a domain to ingest data into"""
    response = opensearch.create_domain(
        DomainName=domainName,
        EngineVersion='OpenSearch_2.3',
        ClusterConfig={
            'InstanceType': 't2.small.search',
            'InstanceCount': 5,
            'DedicatedMasterEnabled': True,
            'DedicatedMasterType': 't2.small.search',
            'DedicatedMasterCount': 3
        },
        # Many instance types require EBS storage.
```

```
        EBSOptions={
            'EBSEnabled': True,
            'VolumeType': 'gp2',
            'VolumeSize': 10
        },
        AccessPolicies=f'{{\"Version\":\"2012-10-17\",\"Statement\":[{{\"Effect
\":\"Allow\",\"Principal\":{{\"AWS\":\"arn:aws:iam::account-id:role\/PipelineRole
\"}},\"Action\":\"es:*\",\"Resource\":\"arn:aws:es:us-east-1:account-id:domain\/
{domainName}\/*\"}}]}}',
        NodeToNodeEncryptionOptions={
            'Enabled': True
        }
    )
    return(response)

def waitForDomainProcessing(opensearch, domainName):
    """Waits for the domain to be active"""
    try:
        response = opensearch.describe_domain(
            DomainName=domainName
        )
        # Every 30 seconds, check whether the domain is processing.
        while 'Endpoint' not in response['DomainStatus']:
            print('Creating domain...')
            time.sleep(60)
            response = opensearch.describe_domain(
                DomainName=domainName)

        # Once we exit the loop, the domain is ready for ingestion.
        endpoint = response['DomainStatus']['Endpoint']
        print('Domain endpoint ready to receive data: ' + endpoint)
        createPipeline(osis, endpoint)

    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ResourceNotFoundException':
            print('Domain not found.')
        else:
            raise error

def createPipeline(osis, endpoint):
    """Creates a pipeline using the domain and pipeline role"""
    try:
        definition = f'version: \"2\"\nlog-pipeline:\n  source:\n    http:\n      path:
 \"/${{pipelineName}}/logs\"\n  processor:\n    - date:\n        from_time_received:
```

```
 true\n        destination: \"@timestamp\"\n  sink:\n    - opensearch:\n        hosts:
 [ \"https://{endpoint}\" ]\n        index: \"application_logs\"\n        aws:\n
   region: \"us-east-1\"'
        response = osis.create_pipeline(
            PipelineName=pipelineName,
            MinUnits=4,
            MaxUnits=9,
            PipelineConfigurationBody=definition,
            PipelineRoleArn="arn:aws:iam::account-id:role/PipelineRole"
        )

        response = osis.get_pipeline(
                PipelineName=pipelineName
        )

        # Every 30 seconds, check whether the pipeline is active.
        while response['Pipeline']['Status'] == 'CREATING':
            print('Creating pipeline...')
            time.sleep(30)
            response = osis.get_pipeline(
                PipelineName=pipelineName)

        # Once we exit the loop, the pipeline is ready for ingestion.
        ingestionEndpoint = response['Pipeline']['IngestEndpointUrls'][0]
        print('Pipeline ready to ingest data at endpoint: ' + ingestionEndpoint)
        ingestData(ingestionEndpoint)

    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ResourceAlreadyExistsException':
            print('Pipeline already exists.')
            response = osis.get_pipeline(
                PipelineName=pipelineName
            )
            ingestionEndpoint = response['Pipeline']['IngestEndpointUrls'][0]
            ingestData(ingestionEndpoint)
        else:
            raise error


def ingestData(ingestionEndpoint):
    """Ingests a sample log file into the pipeline"""
    endpoint = 'https://' + ingestionEndpoint
    r = requests.request('POST', f'{endpoint}/log-pipeline/logs',
```

```
 data='[{"time":"2014-08-11T11:40:13+00:00","remote_addr":"122.226.223.69","status":"404","requ
 http://www.k2proxy.com//hello.html HTTP/1.1","http_user_agent":"Mozilla/4.0
 (compatible; WOW64; SLCC2;)"}]',
    auth=AWSSigV4('osis'))
    print('Ingesting sample log file into pipeline')
    print('Response: ' + r.text)

def main():
    createPipelineRole(iam, domainName)
    createDomain(opensearch, domainName)
    waitForDomainProcessing(opensearch, domainName)

if __name__ == "__main__":
    main()
```

Sicurezza in Amazon OpenSearch Ingestion

La sicurezza del cloud AWS è la massima priorità. In qualità di AWS cliente, puoi beneficiare di un data center e di un'architettura di rete progettati per soddisfare i requisiti delle organizzazioni più sensibili alla sicurezza.

La sicurezza è una responsabilità condivisa tra AWS te e te. Il modello di responsabilità condivisa descrive questo come sicurezza del cloud e sicurezza nel cloud:

- Sicurezza del cloud: AWS è responsabile della protezione dell'infrastruttura che gestisce AWS i servizi nel AWS cloud. AWS ti fornisce anche servizi che puoi utilizzare in modo sicuro. I revisori di terze parti testano e verificano regolarmente l'efficacia della sicurezza come parte dei programmi di conformitàAWS.
- Sicurezza nel cloud: la tua responsabilità è determinata dal AWS servizio che utilizzi. L'utente è anche responsabile di altri fattori, tra cui la riservatezza dei dati, i requisiti della propria azienda e le leggi e normative vigenti.

Questa documentazione aiuta a capire come applicare il modello di responsabilità condivisa quando si utilizza OpenSearch Ingestion. I seguenti argomenti mostrano come configurare OpenSearch Ingestion per soddisfare gli obiettivi di sicurezza e conformità. Imparerai anche come utilizzare altri AWS servizi che ti aiutano a monitorare e proteggere le tue risorse di OpenSearch Ingestion.

Argomenti

- Configurazione dell'accesso VPC per le pipeline di Amazon Ingestion OpenSearch
- Configurazione delle pipeline OpenSearch di ingestione per l'ingestione tra account
- Identity and Access Management per Amazon OpenSearch Ingestion
- Registrazione delle chiamate API Amazon OpenSearch Ingestion tramite AWS CloudTrail
- Amazon OpenSearch Ingestion e API degli endpoint di interfaccia ()AWS PrivateLink

Configurazione dell'accesso VPC per le pipeline di Amazon Ingestion OpenSearch

Puoi accedere alle tue pipeline di Amazon OpenSearch Ingestion utilizzando un endpoint VPC di interfaccia. Un VPC è una rete virtuale dedicata a te. Account AWSÈ logicamente isolato dalle altre reti virtuali nel AWS cloud. L'accesso a una pipeline tramite un endpoint VPC consente una comunicazione sicura OpenSearch tra Ingestion e altri servizi all'interno del VPC senza la necessità di un gateway Internet, un dispositivo NAT o una connessione VPN. Tutto il traffico rimane sicuro all'interno del Cloud. AWS

OpenSearch Ingestion stabilisce questa connessione privata creando un endpoint di interfaccia, alimentato da. AWS PrivateLink Creiamo un'interfaccia di rete endpoint in ogni sottorete specificata durante la creazione della pipeline. Si tratta di interfacce di rete gestite dai richiedenti che fungono da punto di ingresso per il traffico destinato alla pipeline di ingestione. OpenSearch Puoi anche scegliere di creare e gestire tu stesso gli endpoint dell'interfaccia.

L'utilizzo di un VPC consente di imporre il flusso di dati attraverso le pipeline di OpenSearch ingestione entro i confini del VPC, anziché sulla rete Internet pubblica. Le pipeline che non si trovano all'interno di un VPC inviano e ricevono dati su endpoint pubblici e su Internet.

Una pipeline con accesso VPC può scrivere su domini di servizio pubblici o OpenSearch VPC e su raccolte pubbliche o VPC Serverless. OpenSearch

Argomenti

- Considerazioni
- Limitazioni
- Prerequisiti
- Configurazione dell'accesso VPC per una pipeline
- Endpoint VPC autogestiti
- Ruolo collegato ai servizi per l'accesso VPC

Considerazioni

Considerate quanto segue quando configurate l'accesso VPC per una pipeline.

- Non è necessario che una pipeline si trovi nello stesso VPC del relativo sink. Inoltre, non è necessario stabilire una connessione tra i due. VPCs OpenSearch Ingestion si occupa di collegarli per te.
- Puoi specificare un solo VPC per la tua pipeline.
- A differenza delle pipeline pubbliche, una pipeline VPC deve trovarsi nello Regione AWS stesso dominio o sink di raccolta su cui sta scrivendo. Puoi configurare una sorgente S3 per la pipeline per scrivere su più regioni.
- Puoi scegliere di implementare una pipeline in una, due o tre sottoreti del tuo VPC. Le sottoreti sono distribuite nelle stesse zone di disponibilità in cui sono distribuite le Ingestion Compute OpenSearch Unit (). OCUs
- Se distribuisci una pipeline solo in una sottorete e la zona di disponibilità non funziona, non sarai in grado di importare dati. Per garantire un'elevata disponibilità, consigliamo di configurare le pipeline con due o tre sottoreti.
- La specificazione di un gruppo di sicurezza è facoltativa. Se non fornisci un gruppo di sicurezza, OpenSearch Ingestion utilizza il gruppo di sicurezza predefinito specificato nel VPC.

Limitazioni

Le pipeline con accesso VPC presentano le seguenti limitazioni.

- Non è possibile modificare la configurazione di rete di una pipeline dopo averla creata. Se avvii una pipeline all'interno di un VPC, non puoi modificarla successivamente in un endpoint pubblico e viceversa.
- Puoi avviare la tua pipeline con un endpoint VPC di interfaccia o un endpoint pubblico, ma non puoi fare entrambe le cose. È necessario scegliere l'uno o l'altro quando si crea una pipeline.
- Dopo aver effettuato il provisioning di una pipeline con accesso VPC, non è possibile spostarla su un altro VPC e non è possibile modificarne le sottoreti o le impostazioni del gruppo di sicurezza.
- Se la pipeline scrive su un dominio o un sink di raccolta che utilizza l'accesso VPC, non puoi tornare indietro in un secondo momento e modificare il sink (VPC o pubblico) dopo la creazione della pipeline. È necessario eliminare e ricreare la pipeline con un nuovo sink. Puoi comunque passare da un lavandino pubblico a un lavandino con accesso VPC.
- Non puoi fornire l'accesso di importazione tra account diversi alle pipeline VPC.

Prerequisiti

Prima di effettuare il provisioning di una pipeline con accesso VPC, devi fare quanto segue:

- Crea un VPC

  Per creare il tuo VPC, puoi utilizzare la console Amazon VPC, la AWS CLI o uno dei. AWS SDKs Per ulteriori informazioni, consulta Working with VPCs nella Amazon VPC User Guide. Se hai già un VPC, questa fase può essere ignorata.
- Prenota indirizzi IP

  OpenSearch L'ingestione inserisce un'interfaccia elastica di rete in ogni sottorete specificata durante la creazione della pipeline. Ogni interfaccia di rete è associata a un indirizzo IP. È necessario riservare un indirizzo IP per sottorete per le interfacce di rete.

Configurazione dell'accesso VPC per una pipeline

È possibile abilitare l'accesso VPC per una pipeline all'interno della console di OpenSearch servizio o utilizzando il. AWS CLI

Console

L'accesso al VPC viene configurato durante la creazione della pipeline. In Opzioni di rete di origine, scegli Accesso VPC e configura le seguenti impostazioni:

Impostazione	Description
Gestione degli endpoint	Scegliete se volete creare voi stessi gli endpoint VPC o lasciateli creare da OpenSearch Ingestion per voi.
VPC	Scegli l'ID per il cloud privato virtuale (VPC) da utilizzare. Il VPC e la pipeline devono trovarsi nello stesso ambiente. Regione AWS
Sottoreti	Scegli una o più sottoreti. OpenSearch Il servizio inserisce un endpoint VPC e interfacce di rete elastiche nelle sottoreti.
Gruppi di sicurezza	Scegli uno o più gruppi di sicurezza VPC che consentano all'applicazione richiesta di raggiungere la pipeline di OpenSearch ingestione sulle porte (80 o 443) e sui protocolli (HTTP o HTTPs) esposti dalla pipeline.

Impostazione	Description
Opzioni di collegamento VPC	Se la tua fonte richiede una comunicazione tra VPC, come Amazon DocumentDB, OpenSearch autogestita o Confluent Kafka OpenSearch , Ingestion crea Elastic Network Interfaces ENIs () nelle sottoreti specificate per connettersi a queste sorgenti. OpenSearch Ingestion utilizza in ogni zona di disponibilità per raggiungere le fonti specificate. ENIs L'opzione Allega a VPC collega il VPC del piano OpenSearch dati di inserimento al VPC specificato. Seleziona una prenotazione CIDR per il VPC gestito per distribuire l'interfaccia di rete.

CLI

Per configurare l'accesso al VPC utilizzando AWS CLI, specificare il `--vpc-options` parametro:

```
aws osis create-pipeline \
  --pipeline-name vpc-pipeline \
  --min-units 4 \
  --max-units 10 \
  --vpc-options
 SecurityGroupIds={sg-12345678,sg-9012345},SubnetIds=subnet-1212234567834asdf \
  --pipeline-configuration-body "file://pipeline-config.yaml"
```

Endpoint VPC autogestiti

Quando crei una pipeline, puoi utilizzare la gestione degli endpoint per creare una pipeline con endpoint autogestiti o endpoint gestiti dal servizio. La gestione degli endpoint è facoltativa e per impostazione predefinita utilizza gli endpoint gestiti da Ingestion. OpenSearch

Per creare una pipeline con un endpoint VPC autogestito in Console di gestione AWS, vedi Creazione di pipeline con la console di servizio. OpenSearch Per creare una pipeline con un endpoint VPC autogestito in AWS CLI, puoi utilizzare il `--vpc-options` parametro nel comando create-pipeline:

```
--vpc-options SubnetIds=subnet-abcdef01234567890,VpcEndpointManagement=CUSTOMER
```

Puoi creare tu stesso un endpoint per la tua pipeline quando specifichi il tuo servizio endpoint. Per trovare il tuo servizio endpoint, usa il comando get-pipeline, che restituisce una risposta simile alla seguente:

```
"vpcEndpointService" : "com.amazonaws.osis.us-east-1.pipeline-
id-1234567890abcdef1234567890",
"vpcEndpoints" : [
  {
    "vpcId" : "vpc-1234567890abcdef0",
    "vpcOptions" : {
      "subnetIds" : [ "subnet-abcdef01234567890", "subnet-021345abcdef6789" ],
      "vpcEndpointManagement" : "CUSTOMER"
    }
  }
```

Usa `vpcEndpointService` from the response per creare un endpoint VPC con o. Console di gestione AWS AWS CLI

Se utilizzi endpoint VPC autogestiti, devi abilitare gli attributi DNS `enableDnsSupport` e `enableDnsHostnames` nel tuo VPC. Tieni presente che se disponi di una pipeline con un endpoint autogestito che interrompi e riavvii, devi ricreare l'endpoint VPC nel tuo account.

Ruolo collegato ai servizi per l'accesso VPC

Un ruolo collegato ai servizi è un tipo specifico di ruolo IAM che delega le autorizzazioni a un servizio così che possa creare e gestire le risorse per conto dell'utente. Se scegli un endpoint VPC gestito dal servizio, OpenSearch Ingestion richiede un ruolo collegato al servizio chiamato AWSServiceRoleForAmazonOpenSearchIngestionServiceper accedere al tuo VPC, creare l'endpoint della pipeline e posizionare le interfacce di rete in una sottorete del VPC.

Se scegli un endpoint VPC autogestito OpenSearch , Ingestion richiede un ruolo collegato al servizio chiamato. AWSServiceRoleForOpensearchIngestionSelfManagedVpce Per ulteriori informazioni su questi ruoli, le relative autorizzazioni e su come eliminarli, consulta. the section called “Ruolo di creazione della pipeline”

OpenSearch Ingestion crea automaticamente il ruolo quando si crea una pipeline di ingestione. Affinché questa creazione automatica abbia esito positivo, l'utente che crea la prima pipeline in un account deve disporre delle autorizzazioni per l'azione. `iam:CreateServiceLinkedRole` Per ulteriori informazioni, consultare Autorizzazioni del ruolo collegato ai servizi nella Guida per l'utente di

IAM. Puoi visualizzare il ruolo nella console AWS Identity and Access Management (IAM) dopo la sua creazione.

Configurazione delle pipeline OpenSearch di ingestione per l'ingestione tra account

Per sorgenti basate su push come HTTP e OTel Amazon OpenSearch Ingestion ti consente di condividere le pipeline Account AWS da un cloud privato virtuale (VPC) a un endpoint di pipeline in un VPC separato. I team che condividono le analisi con altri team della propria organizzazione utilizzano questa funzionalità per un mezzo più semplificato, ad esempio, per condividere l'analisi dei log.

Questa sezione utilizza la seguente terminologia:

- Proprietario della pipeline: l'account che possiede e gestisce la pipeline di ingestione. OpenSearch Solo un account può possedere una pipeline.
- Account di connessione: un account che si connette e utilizza una pipeline condivisa. È possibile connettere più account alla stessa pipeline.

VPCs Per configurare la condivisione tra le pipeline di OpenSearch Ingestion Account AWS, completa le seguenti attività, come descritto di seguito:

- (Proprietario della pipeline) Concedi agli account di connessione l'accesso a una pipeline
- (Account di collegamento) Crea una connessione endpoint della pipeline per ogni VPC connesso

Prima di iniziare

Prima di configurare VPCs la condivisione tra le pipeline di OpenSearch Ingestion Account AWS, completate le seguenti attività:

Operazione	Informazioni
Crea una o più pipeline di ingestione OpenSearch	Imposta il numero minimo di unità di OpenSearch calcolo (OSUs) su 2 o superiore . Per ulteriori informazioni, consulta Creazione di pipeline Amazon OpenSearch Ingestion . Per informazioni sull'aggiornamento di una

Operazione	Informazioni
	pipeline, vedere. Aggiornamento delle pipeline di Amazon OpenSearch Ingestion
Creane uno o più VPCs per Ingestion OpenSearch	Per abilitare la condivisione della pipeline tra account, qualsiasi VPC coinvolto per la pipeline e gli endpoint della pipeline deve essere configurato con i seguenti valori DNS: • `enableDnsSupport=true` • `enableDnsHostnames=true` Per ulteriori informazioni, consulta Attributi DNS per il VPC nella Guida per l'utente di Amazon VPC.

Concedi agli account di connessione l'accesso a una pipeline

Le procedure in questa sezione descrivono come utilizzare la console di OpenSearch servizio e configurare l'accesso AWS CLI alla pipeline tra account tramite la creazione di una politica delle risorse. Una politica delle risorse consente al proprietario di una pipeline di specificare altri account che possono accedere a una pipeline. Una volta create, le politiche di pipeline esistono finché esiste la pipeline o fino a quando la policy non viene eliminata.

> **Note**
>
> Le politiche relative alle risorse non sostituiscono l'autorizzazione standard di OpenSearch Ingestion utilizzando le autorizzazioni IAM. Le policy relative alle risorse sono un meccanismo di autorizzazione aggiuntivo per consentire l'accesso alla pipeline tra più account.

Argomenti

- Concedi agli account connessi l'accesso a una pipeline (console)
- Concedi agli account connessi l'accesso a una pipeline (CLI)

Concedi agli account connessi l'accesso a una pipeline (console)

Utilizza la seguente procedura per concedere agli account connessi l'accesso a una pipeline utilizzando la console di Amazon OpenSearch Service.

Per creare una connessione endpoint alla pipeline

1. Nella console di Amazon OpenSearch Service, nel riquadro di navigazione, espandi Ingestion, quindi seleziona Pipelines.
2. Nella sezione Pipelines, scegli il nome di una pipeline a cui desideri concedere l'accesso per un account di connessione.
3. Scegli la scheda Endpoint VPC.
4. Nella sezione Responsabili autorizzati, scegli Autorizza account.
5. Nel campo Account AWS ID, inserisci l'ID dell'account numerico a 12 cifre, quindi seleziona Autorizza.

Concedi agli account connessi l'accesso a una pipeline (CLI)

Utilizzare la procedura seguente per concedere agli account di connessione l'accesso a una pipeline utilizzando. AWS CLI

Per concedere agli account connessi l'accesso alla pipeline

1. Effettuare l'aggiornamento alla versione più recente di AWS CLI (versione 2.0). Per ulteriori informazioni, consulta Installazione o aggiornamento alla versione più recente della AWS CLI.
2. Apri la CLI nell'account e Regione AWS con la pipeline che desideri condividere.
3. Esegui il comando seguente per creare una politica delle risorse per la pipeline. Questa politica fornisce l'`osis:CreatePipelineEndpoint`autorizzazione sulla pipeline. La politica include un parametro in cui è possibile Account AWS IDs elencare i dati consentiti.

> **Note**
>
> Nel comando seguente, è necessario utilizzare la forma abbreviata dell'ID dell'account fornendo solo l'ID dell'account a dodici cifre. L'utilizzo di un ARN non funzionerà. È inoltre necessario fornire l'Amazon Resource Name (ARN) della pipeline nel parametro CLI per `resource-arn` e nella policy JSON sotto riportata, come illustrato. `Resource`

```
aws --region region osis put-resource-policy \
  --resource-arn arn:aws:osis:region:pipeline-owner-account-ID:pipeline/pipeline-name
  --policy 'IAM-policy'
```

Utilizza una politica come la seguente per *IAM-policy*

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
  {
  "Sid": "AllowAccess",
  "Effect": "Allow",
  "Principal": {
  "AWS": [
  "111122223333",
  "444455556666"
  ]
  },
  "Action":
  "osis:CreatePipelineEndpoint",
  "Resource": "arn:aws:osis:us-east-1:123456789012:pipeline/pipeline-name"
  }
  ]
 }
```

Crea una connessione endpoint della pipeline per ogni VPC connesso

Dopo che il proprietario della pipeline ha concesso l'accesso a una pipeline nel proprio VPC utilizzando la procedura precedente, un utente dell'account di connessione crea un endpoint della pipeline nel proprio VPC. Questa sezione include le procedure per la creazione di endpoint utilizzando la console di servizio e il. OpenSearch AWS CLI Quando si crea un endpoint, OpenSearch Ingestion esegue le seguenti azioni:

- Crea il ruolo AWSServiceRoleForAmazonOpenSearchIngestionServicecollegato al servizio nel tuo account, se non esiste già. Questo ruolo fornisce all'utente dell'account di connessione l'autorizzazione a richiamare l'azione CreatePipelineEndpointAPI.
- Crea l'endpoint della pipeline.
- Configura l'endpoint della pipeline per importare i dati dalla pipeline condivisa nel VPC del proprietario della pipeline.

Argomenti

- Creazione di una connessione all'endpoint della pipeline (console)
- Creazione di una connessione endpoint della pipeline (CLI)

Creazione di una connessione all'endpoint della pipeline (console)

Utilizzare la procedura seguente per creare una connessione endpoint della pipeline utilizzando la console di servizio. OpenSearch

Per creare una connessione endpoint alla pipeline

1. Nella console di Amazon OpenSearch Service, nel riquadro di navigazione, espandi Ingestion, quindi seleziona Endpoint VPC.
2. Nella pagina degli endpoint VPC, scegli Crea.
3. Per Pipeline location, scegli un'opzione. Se scegli Conto corrente, scegli la pipeline dall'elenco. Se scegli Cross-account, specifica l'ARN della pipeline nel campo. Il proprietario della pipeline deve aver concesso l'accesso alla pipeline, come descritto in. Concedi agli account di connessione l'accesso a una pipeline
4. Nella sezione Impostazioni VPC, per VPC, scegli un VPC dall'elenco.
5. Per Subnets (Sottoreti), scegliere una sottorete.
6. Per i gruppi di sicurezza, scegli un gruppo.
7. Seleziona Crea endpoint.

Attendi lo stato dell'endpoint che hai creato a cui passare`ACTIVE`. Una volta completata la pipeline`ACTIVE`, vedrai un nuovo campo denominato. `ingestEndpointUrl` Usa questo endpoint per accedere alla pipeline e importare dati utilizzando un client come. FluentBit Per ulteriori informazioni sull'utilizzo per FluentBit importare dati, consulta. Utilizzo di una pipeline di OpenSearch ingestione con Fluent Bit

> **Note**
>
> `ingestEndpointUrl`È lo stesso URL per tutti gli account collegati.

Creazione di una connessione endpoint della pipeline (CLI)

Utilizzare la procedura seguente per creare una connessione all'endpoint della pipeline utilizzando. AWS CLI

Per creare una connessione all'endpoint della pipeline

1. Se non l'hai già fatto, esegui l'aggiornamento alla versione più recente di AWS CLI (versione 2.0). Per ulteriori informazioni, consulta Installazione o aggiornamento alla versione più recente della AWS CLI.
2. Apri la CLI nell'account di connessione nella pipeline Regione AWS condivisa.
3. Esegui il comando seguente per creare un endpoint della pipeline.

   > **Note**
   >
   > È necessario fornire almeno una sottorete e un gruppo di sicurezza per il VPC dell'account di connessione. Il gruppo di sicurezza deve includere la porta 443 e supportare i client nella connessione del VPC dell'account.

   ```
   aws osis --region region create-pipeline-endpoint \
     --pipeline-arn arn:aws:osis:region:connecting-account-ID:pipeline/shared-pipeline-name
     --vpc-options SecurityGroupIds={sg-security-group-ID-1,sg-security-group-ID-2},SubnetIds=subnet-subnet-ID
   ```

4. Esegui il comando seguente per elencare gli endpoint nella regione specificata nel comando precedente:

   ```
   aws osis --region region list-pipeline-endpoints
   ```

Attendi la transizione allo stato dell'endpoint che hai creato. `ACTIVE` Una volta completata la pipeline`ACTIVE`, vedrai un nuovo campo denominato. `ingestEndpointUrl` Usa questo endpoint

per accedere alla pipeline e importare dati utilizzando un client come. FluentBit Per ulteriori informazioni sull'utilizzo per FluentBit importare dati, consulta. Utilizzo di una pipeline di OpenSearch ingestione con Fluent Bit

> **Note**
>
> `ingestEndpointUrl`È lo stesso URL per tutti gli account collegati.

Rimozione degli endpoint della pipeline

Se non desideri più fornire l'accesso a una pipeline condivisa, puoi rimuovere un endpoint della pipeline utilizzando uno dei seguenti metodi:

- Eliminare l'endpoint della pipeline (account di connessione).
- Revoca l'endpoint della pipeline (proprietario della pipeline).

Utilizzate la seguente procedura per eliminare un endpoint della pipeline in un account di connessione.

Per eliminare un endpoint della pipeline (account di connessione)

1. Apri la CLI nell'account di connessione nella pipeline Regione AWS condivisa.
2. Esegui il comando seguente per elencare gli endpoint della pipeline nella regione:

   ```
   aws osis --region region list-pipeline-endpoints
   ```

   Prendi nota dell'ID della pipeline che desideri eliminare.
3. Esegui il comando seguente per eliminare l'endpoint della pipeline:

   ```
   aws osis --region region delete-pipeline-endpoint \
     --endpoint-id 'ID'
   ```

In qualità di proprietario della pipeline condivisa, utilizzate la procedura seguente per revocare un endpoint della pipeline.

Per revocare un endpoint della pipeline (proprietario della pipeline)

1. Apri la CLI nell'account di connessione nella pipeline Regione AWS condivisa.
2. Esegui il comando seguente per elencare le connessioni degli endpoint della pipeline nella regione:

   ```
   aws osis --region region list-pipeline-endpoint-connections
   ```

   Prendi nota dell'ID della pipeline che desideri eliminare.
3. Esegui il comando seguente per eliminare l'endpoint della pipeline:

   ```
   aws osis --region region revoke-pipeline-endpoint-connections \
     --pipeline-arn pipeline-arn --endpoint-ids ID
   ```

   Il comando supporta la specificazione di un solo ID di endpoint.

Identity and Access Management per Amazon OpenSearch Ingestion

AWS Identity and Access Management (IAM) è un software Servizio AWS che aiuta un amministratore a controllare in modo sicuro l'accesso alle risorse. AWS Gli amministratori IAM controllano chi può essere autenticato (effettuato l'accesso) e autorizzato (disporre delle autorizzazioni) a utilizzare le risorse di Ingestion. OpenSearch IAM è uno strumento Servizio AWS che puoi utilizzare senza costi aggiuntivi.

Argomenti

- Politiche basate sull'identità per Ingestion OpenSearch
- OpenSearch Azioni politiche per Ingestion
- OpenSearch Risorse politiche per Ingestion
- Chiavi delle condizioni delle politiche per Amazon OpenSearch Ingestion
- ABAC con Ingestion OpenSearch
- Utilizzo di credenziali temporanee con Ingestion OpenSearch
- Ruoli collegati ai servizi per Ingestion OpenSearch
- Esempi di policy basate sull'identità per Ingestion OpenSearch

Politiche basate sull'identità per Ingestion OpenSearch

Supporta le policy basate sull'identità: sì

Le policy basate sull'identità sono documenti di policy di autorizzazione JSON che è possibile allegare a un'identità (utente, gruppo di utenti o ruolo IAM). Tali policy definiscono le operazioni che utenti e ruoli possono eseguire, su quali risorse e in quali condizioni. Per informazioni su come creare una policy basata su identità, consulta Definizione di autorizzazioni personalizzate IAM con policy gestite dal cliente nella Guida per l'utente di IAM.

Con le policy basate sull'identità di IAM, è possibile specificare quali operazioni e risorse sono consentite o respinte, nonché le condizioni in base alle quali le operazioni sono consentite o respinte. Per informazioni su tutti gli elementi utilizzabili in una policy JSON, consulta Guida di riferimento agli elementi delle policy JSON IAM nella Guida per l'utente IAM.

Esempi di policy basate sull'identità per Ingestion OpenSearch

Per visualizzare esempi di politiche basate sull'identità di OpenSearch Ingestion, vedere. the section called "Esempi di policy basate su identità"

OpenSearch Azioni politiche per Ingestion

Supporta le operazioni di policy: si

L'elemento `Action` di una policy JSON descrive le operazioni che è possibile utilizzare per consentire o negare l'accesso in una policy. Le azioni politiche in genere hanno lo stesso nome dell'operazione AWS API associata. Ci sono alcune eccezioni, ad esempio le operazioni di sola autorizzazione che non hanno un'operazione API corrispondente. Esistono anche alcune operazioni che richiedono più operazioni in una policy. Queste operazioni aggiuntive sono denominate operazioni dipendenti.

Includere le operazioni in una policy per concedere le autorizzazioni a eseguire l'operazione associata.

Le azioni politiche in OpenSearch Ingestion utilizzano il seguente prefisso prima dell'azione:

```
osis
```

Per specificare più operazioni in una sola istruzione, occorre separarle con la virgola.

```
"Action": [
```

```
      "osis:action1",
      "osis:action2"
         ]
```

Puoi specificare più operazioni tramite caratteri jolly (*). Ad esempio, per specificare tutte le azioni che iniziano con la parola `List`, includi la seguente azione:

```
"Action": "osis:List*"
```

Per visualizzare esempi di politiche basate sull'identità di OpenSearch Ingestion, vedere. Esempi di policy basate sull'identità per Serverless OpenSearch

OpenSearch Risorse politiche per Ingestion

Supporta le risorse relative alle policy: sì

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. In altre parole, quale entità principale può eseguire operazioni su quali risorse e in quali condizioni.

L'elemento JSON `Resource` della policy specifica l'oggetto o gli oggetti ai quali si applica l'operazione. Come best practice, specifica una risorsa utilizzando il suo nome della risorsa Amazon (ARN). Per le azioni che non supportano le autorizzazioni a livello di risorsa, si utilizza un carattere jolly (*) per indicare che l'istruzione si applica a tutte le risorse.

```
"Resource": "*"
```

Chiavi delle condizioni delle politiche per Amazon OpenSearch Ingestion

Supporta le chiavi di condizione delle policy specifiche del servizio: No

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. In altre parole, quale entità principale può eseguire operazioni su quali risorse e in quali condizioni.

L'elemento `Condition` specifica quando le istruzioni vengono eseguite in base a criteri definiti. È possibile compilare espressioni condizionali che utilizzano operatori di condizione, ad esempio uguale a o minore di, per soddisfare la condizione nella policy con i valori nella richiesta. Per visualizzare tutte le chiavi di condizione AWS globali, consulta le chiavi di contesto delle condizioni AWS globali nella Guida per l'utente IAM.

Per visualizzare un elenco delle chiavi delle condizioni di OpenSearch ingestione, consulta Condition keys for Amazon OpenSearch Ingestion nel Service Authorization Reference. Per sapere con

quali azioni e risorse puoi utilizzare una chiave di condizione, consulta Azioni definite da Amazon OpenSearch Ingestion.

ABAC con Ingestion OpenSearch

Supporta ABAC (tag nelle policy): sì

Il controllo degli accessi basato su attributi (ABAC) è una strategia di autorizzazione che definisce le autorizzazioni in base ad attributi chiamati tag. Puoi allegare tag a entità e AWS risorse IAM, quindi progettare politiche ABAC per consentire le operazioni quando il tag del principale corrisponde al tag sulla risorsa.

Per controllare l'accesso basato su tag, fornire informazioni sui tag nell'elemento condizione di una policy utilizzando le chiavi di condizione `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` o `aws:TagKeys`.

Se un servizio supporta tutte e tre le chiavi di condizione per ogni tipo di risorsa, il valore per il servizio è Sì. Se un servizio supporta tutte e tre le chiavi di condizione solo per alcuni tipi di risorsa, allora il valore sarà Parziale.

Per maggiori informazioni su ABAC, consulta Definizione delle autorizzazioni con autorizzazione ABAC nella Guida per l'utente di IAM. Per visualizzare un tutorial con i passaggi per l'impostazione di ABAC, consulta Utilizzo del controllo degli accessi basato su attributi (ABAC) nella Guida per l'utente di IAM.

Per ulteriori informazioni sull'etichettatura delle risorse di OpenSearch Ingestion, consulta. the section called "Etichettatura delle pipeline"

Utilizzo di credenziali temporanee con Ingestion OpenSearch

Supporta le credenziali temporanee: sì

Le credenziali temporanee forniscono l'accesso a breve termine alle AWS risorse e vengono create automaticamente quando si utilizza la federazione o si cambia ruolo. AWS consiglia di generare dinamicamente credenziali temporanee anziché utilizzare chiavi di accesso a lungo termine. Per ulteriori informazioni, consulta Credenziali di sicurezza temporanee in IAM e Servizi AWS compatibili con IAM nella Guida per l'utente IAM.

Ruoli collegati ai servizi per Ingestion OpenSearch

Supporta i ruoli collegati ai servizi: sì

Un ruolo collegato al servizio è un tipo di ruolo di servizio collegato a un. Servizio AWS Il servizio può assumere il ruolo per eseguire un'operazione per tuo conto. I ruoli collegati al servizio vengono visualizzati nel tuo account Account AWS e sono di proprietà del servizio. Un amministratore IAM può visualizzare le autorizzazioni per i ruoli collegati al servizio, ma non modificarle.

OpenSearch Ingestion utilizza un ruolo collegato al servizio chiamato. `AWSServiceRoleForAmazonOpenSearchIngestionService` Il ruolo collegato ai servizi denominato `AWSServiceRoleForOpensearchIngestionSelfManagedVpce` è disponibile anche per le pipeline con endpoint VPC autogestiti. Per informazioni dettagliate sulla creazione e la gestione dei ruoli collegati al servizio Ingestion, consulta OpenSearch . the section called "Ruolo di creazione della pipeline"

Esempi di policy basate sull'identità per Ingestion OpenSearch

Per impostazione predefinita, gli utenti e i ruoli non dispongono dell'autorizzazione per creare o modificare le risorse di Ingestion. OpenSearch Per concedere agli utenti l'autorizzazione a eseguire azioni sulle risorse di cui hanno bisogno, un amministratore IAM può creare policy IAM.

Per informazioni su come creare una policy basata su identità IAM utilizzando questi documenti di policy JSON di esempio, consulta Creazione di policy IAM (console) nella Guida per l'utente di IAM.

Per dettagli sulle azioni e sui tipi di risorse definiti da Amazon OpenSearch Ingestion, incluso il formato di ARNs per ogni tipo di risorsa, consulta Azioni, risorse e chiavi di condizione per Amazon OpenSearch Ingestion nel Service Authorization Reference.

Argomenti

- Best practice per le policy
- Utilizzo di OpenSearch Ingestion nella console
- OpenSearch Amministrazione delle pipeline di Ingestion
- Inserimento di dati in una pipeline di ingestione OpenSearch

Best practice per le policy

Le policy basate su identità sono molto efficaci. Determinano se qualcuno può creare, accedere o eliminare le risorse di OpenSearch Ingestion nel tuo account. Queste azioni possono comportare costi aggiuntivi per l' Account AWS. Quando si creano o modificano policy basate sull'identità, seguire queste linee guida e raccomandazioni:

Le politiche basate sull'identità determinano se qualcuno può creare, accedere o eliminare le risorse di OpenSearch Ingestion nel tuo account. Queste azioni possono comportare costi aggiuntivi per l'Account AWS. Quando si creano o modificano policy basate sull'identità, seguire queste linee guida e raccomandazioni:

- Inizia con le policy AWS gestite e passa alle autorizzazioni con privilegi minimi: per iniziare a concedere autorizzazioni a utenti e carichi di lavoro, utilizza le politiche gestite che concedono le autorizzazioni per molti casi d'uso comuni.AWS Sono disponibili nel tuo. Account AWS Ti consigliamo di ridurre ulteriormente le autorizzazioni definendo politiche gestite dai AWS clienti specifiche per i tuoi casi d'uso. Per maggiori informazioni, consulta Policy gestite da AWS o Policy gestite da AWS per le funzioni dei processi nella Guida per l'utente di IAM.
- Applicazione delle autorizzazioni con privilegio minimo - Quando si impostano le autorizzazioni con le policy IAM, concedere solo le autorizzazioni richieste per eseguire un'attività. È possibile farlo definendo le azioni che possono essere intraprese su risorse specifiche in condizioni specifiche, note anche come autorizzazioni con privilegio minimo. Per maggiori informazioni sull'utilizzo di IAM per applicare le autorizzazioni, consulta Policy e autorizzazioni in IAM nella Guida per l'utente di IAM.
- Condizioni d'uso nelle policy IAM per limitare ulteriormente l'accesso - Per limitare l'accesso ad azioni e risorse è possibile aggiungere una condizione alle policy. Ad esempio, è possibile scrivere una condizione di policy per specificare che tutte le richieste devono essere inviate utilizzando SSL. Puoi anche utilizzare le condizioni per concedere l'accesso alle azioni del servizio se vengono utilizzate tramite uno specifico Servizio AWS, ad esempio CloudFormation. Per maggiori informazioni, consultare la sezione Elementi delle policy JSON di IAM: condizione nella Guida per l'utente di IAM.
- Utilizzo dello strumento di analisi degli accessi IAM per convalidare le policy IAM e garantire autorizzazioni sicure e funzionali - Lo strumento di analisi degli accessi IAM convalida le policy nuove ed esistenti in modo che aderiscano al linguaggio (JSON) della policy IAM e alle best practice di IAM. Lo strumento di analisi degli accessi IAM offre oltre 100 controlli delle policy e consigli utili per creare policy sicure e funzionali. Per maggiori informazioni, consultare Convalida delle policy per il Sistema di analisi degli accessi IAM nella Guida per l'utente di IAM.
- Richiedi l'autenticazione a più fattori (MFA): se hai uno scenario che richiede utenti IAM o un utente root nel Account AWS tuo, attiva l'MFA per una maggiore sicurezza. Per richiedere la MFA quando vengono chiamate le operazioni API, aggiungere le condizioni MFA alle policy. Per maggiori informazioni, consultare Protezione dell'accesso API con MFA nella Guida per l'utente di IAM.

Per maggiori informazioni sulle best practice in IAM, consulta Best practice di sicurezza in IAM nella Guida per l'utente di IAM.

Utilizzo di OpenSearch Ingestion nella console

Per accedere a OpenSearch Ingestion dalla console di OpenSearch servizio, è necessario disporre di un set minimo di autorizzazioni. Queste autorizzazioni devono consentire all'utente di elencare e visualizzare i dettagli sulle risorse di OpenSearch Ingestion presenti nell'account. AWS Se crei una policy basata su identità più restrittiva rispetto alle autorizzazioni minime richieste, la console non funzionerà nel modo previsto per le entità (come i ruoli IAM) associate a tale policy.

Non è necessario consentire autorizzazioni minime per la console per gli utenti che effettuano chiamate solo verso o l' AWS CLI API. AWS Al contrario, è possibile accedere solo alle operazioni che soddisfano l'operazione API che stai cercando di eseguire.

La seguente politica consente a un utente di accedere a OpenSearch Ingestion dalla console di servizio: OpenSearch

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Resource": "*",
            "Effect": "Allow",
            "Action": [
                "osis:ListPipelines",
                "osis:GetPipeline",
                "osis:ListPipelineBlueprints",
                "osis:GetPipelineBlueprint",
                "osis:GetPipelineChangeProgress"
            ]
        }
    ]
}
```

In alternativa, è possibile utilizzare la policy the section called "AmazonOpenSearchIngestionReadOnlyAccess" AWS gestita, che concede l'accesso in sola lettura a tutte le risorse di Ingestion per un. OpenSearch Account AWS

OpenSearch Amministrazione delle pipeline di Ingestion

Questa policy è un esempio di policy di «pipeline admin» che consente a un utente di gestire e amministrare le pipeline di Amazon OpenSearch Ingestion. L'utente può creare, visualizzare ed eliminare le pipeline.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Resource": "arn:aws:osis:us-east-1:111122223333:pipeline/*",
            "Action": [
                "osis:CreatePipeline",
                "osis:DeletePipeline",
                "osis:UpdatePipeline",
                "osis:ValidatePipeline",
                "osis:StartPipeline",
                "osis:StopPipeline"
            ],
            "Effect": "Allow"
        },
        {
            "Resource": "*",
            "Action": [
                "osis:ListPipelines",
                "osis:GetPipeline",
                "osis:ListPipelineBlueprints",
                "osis:GetPipelineBlueprint",
                "osis:GetPipelineChangeProgress"
            ],
            "Effect": "Allow"
        }
    ]
}
```

Inserimento di dati in una pipeline di ingestione OpenSearch

Questa policy di esempio consente a un utente o a un'altra entità di inserire dati in una pipeline di Amazon OpenSearch Ingestion nel proprio account. L'utente non può modificare le pipeline.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Resource": "arn:aws:osis:us-east-1:123456789012:pipeline/*",
            "Action": [
                "osis:Ingest"
            ],
            "Effect": "Allow"
        }
    ]
}
```

Registrazione delle chiamate API Amazon OpenSearch Ingestion tramite AWS CloudTrail

Amazon OpenSearch Ingestion è integrato con AWS CloudTrail, un servizio che fornisce un registro delle azioni intraprese da un utente, ruolo o AWS servizio in OpenSearch Ingestion.

CloudTrail acquisisce tutte le chiamate API per Ingestion come eventi. OpenSearch Le chiamate acquisite includono chiamate dalla sezione OpenSearch Ingestion della console di OpenSearch servizio e chiamate di codice alle operazioni dell'API Ingestion. OpenSearch

Se crei un trail, puoi abilitare la distribuzione continua di CloudTrail eventi a un bucket Amazon S3, inclusi gli eventi per Ingestion. OpenSearch Se non configuri un percorso, puoi comunque visualizzare gli eventi più recenti nella CloudTrail console nella cronologia degli eventi.

Utilizzando le informazioni raccolte da CloudTrail, è possibile determinare la richiesta effettuata a OpenSearch Ingestion, l'indirizzo IP da cui è stata effettuata la richiesta, chi ha effettuato la richiesta, quando è stata effettuata e ulteriori dettagli.

Per ulteriori informazioni CloudTrail, consulta la Guida per l'AWS CloudTrail utente.

OpenSearch Informazioni sull'ingestione in CloudTrail

CloudTrail è abilitato sul tuo Account AWS quando crei l'account. Quando si verifica un'attività in OpenSearch Ingestion, tale attività viene registrata in un CloudTrail evento insieme ad altri eventi di AWS servizio nella cronologia degli eventi. Puoi visualizzare, cercare e scaricare gli eventi recenti

nell' Account AWS. Per ulteriori informazioni, vedere Visualizzazione degli eventi con CloudTrail la cronologia degli eventi.

Per una registrazione continua degli eventi in tuo Account AWS, inclusi gli eventi per OpenSearch Ingestion, crea un percorso. Un trail consente di CloudTrail inviare file di log a un bucket Amazon S3. Per impostazione predefinita, quando si crea un percorso nella console, questo sarà valido in tutte le Regioni AWS.

Il trail registra gli eventi di tutte le regioni della AWS partizione e consegna i file di log al bucket Amazon S3 specificato. Inoltre, puoi configurare altri AWS servizi per analizzare ulteriormente e agire in base ai dati sugli eventi raccolti nei log. CloudTrail Per ulteriori informazioni, consulta gli argomenti seguenti:

- Panoramica della creazione di un percorso
- CloudTrail servizi e integrazioni supportati
- Configurazione delle notifiche Amazon SNS per CloudTrail
- Ricezione di file di CloudTrail registro da più regioni e ricezione di file di CloudTrail registro da più account

Tutte le azioni OpenSearch di ingestione vengono registrate CloudTrail e documentate nel riferimento all'API di Ingestion. OpenSearch Ad esempio, le chiamate alle operazioni `CreateCollection`, `ListCollections` e `DeleteCollection` generano voci nei file di log CloudTrail .

Ogni evento o voce di log contiene informazioni sull'utente che ha generato la richiesta. Le informazioni di identità consentono di stabilire:

- Se la richiesta è stata effettuata con credenziali utente root o AWS Identity and Access Management (IAM).
- Se la richiesta è stata effettuata con le credenziali di sicurezza temporanee per un ruolo o un utente federato.
- Se la richiesta è stata effettuata da un altro AWS servizio.

Per ulteriori informazioni, consulta Elemento CloudTrail userIdentity.

Comprensione delle OpenSearch voci dei file di registro di ingestione

Un trail è una configurazione che consente la distribuzione di eventi come file di log in un bucket Amazon S3 specificato dall'utente. CloudTrail i file di registro contengono una o più voci di registro.

Un evento rappresenta una singola richiesta da un'origine. Include informazioni sull'azione richiesta, la data e l'ora dell'azione, i parametri della richiesta e così via. CloudTrail i file di registro non sono una traccia ordinata dello stack delle chiamate API pubbliche, quindi non vengono visualizzati in un ordine specifico.

L'esempio seguente mostra una voce di CloudTrail registro che illustra l'`DeletePipeline`azione.


```
{
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "AIDACKCEVSQ6C2EXAMPLE",
        "arn":"arn:aws:iam::123456789012:user/test-user",
        "accountId": "123456789012",
        "accessKeyId": "access-key",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AIDACKCEVSQ6C2EXAMPLE",
                "arn": "arn:aws:iam::123456789012:role/Admin",
                "accountId": "123456789012",
                "userName": "Admin"
            },
            "webIdFederationData": {},
            "attributes": {
                "creationDate": "2023-04-21T16:48:33Z",
                "mfaAuthenticated": "false"
            }
        }
    },
    "eventTime": "2023-04-21T16:49:22Z",
    "eventSource": "osis.amazonaws.com",
    "eventName": "UpdatePipeline",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "123.456.789.012",
    "userAgent": "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36
 (KHTML, like Gecko) Chrome/112.0.0.0 Safari/537.36",
    "requestParameters": {
        "pipelineName": "my-pipeline",
        "pipelineConfigurationBody": "version: \"2\"\nlog-pipeline:\n  source:\n
 http:\n        path: \"/test/logs\"\n  processor:\n    - grok:\n        match:\n
        log: [ '%{COMMONAPACHELOG}' ]\n    - date:\n        from_time_received: true
\n        destination: \"@timestamp\"\n  sink:\n    - opensearch:\n        hosts:
```

```
 [ \"https://search-b5zd22mwxhggheqpj5ftslgyle.us-west-2.es.amazonaws.com\" ]\n
  index: \"apache_logs2\"\n        aws_sts_role_arn: \"arn:aws:iam::709387180454:role/
canary-bootstrap-OsisRole-J1BARLD26QKN\"\n        aws_region: \"us-west-2\"\n
 aws_sigv4: true\n"
    },
    "responseElements": {
        "pipeline": {
            "pipelineName": "my-pipeline",sourceIPAddress
            "pipelineArn": "arn:aws:osis:us-west-2:123456789012:pipeline/my-pipeline",
            "minUnits": 1,
            "maxUnits": 1,
            "status": "UPDATING",
            "statusReason": {
                "description": "An update was triggered for the pipeline. It is still
 available to ingest data."
            },
            "pipelineConfigurationBody": "version: \"2\"\nlog-pipeline:\n  source:\n
    http:\n        path: \"/test/logs\"\n  processor:\n    - grok:\n        match:
\n          log: [ '%{COMMONAPACHELOG}' ]\n    - date:\n        from_time_received:
 true\n        destination: \"@timestamp\"\n  sink:\n    - opensearch:\n        hosts:
 [ \"https://search-b5zd22mwxhggheqpj5ftslgyle.us-west-2.es.amazonaws.com\" ]\n
  index: \"apache_logs2\"\n        aws_sts_role_arn: \"arn:aws:iam::709387180454:role/
canary-bootstrap-OsisRole-J1BARLD26QKN\"\n        aws_region: \"us-west-2\"\n
 aws_sigv4: true\n",
            "createdAt": "Mar 29, 2023 1:03:44 PM",
            "lastUpdatedAt": "Apr 21, 2023 9:49:21 AM",
            "ingestEndpointUrls": [
                "my-pipeline-tu33ldsgdltgv7x7tjqiudvf7m.us-west-2.osis.amazonaws.com"
            ]
        }
    },
    "requestID": "12345678-1234-1234-1234-987654321098",
    "eventID": "12345678-1234-1234-1234-987654321098",
    "readOnly": false,
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "709387180454",
    "eventCategory": "Management",
    "tlsDetails": {
        "tlsVersion": "TLSv1.2",
        "cipherSuite": "ECDHE-RSA-AES128-GCM-SHA256",
        "clientProvidedHostHeader": "osis.us-west-2.amazonaws.com"
    },
    "sessionCredentialFromConsole": "true"
```

```
}
```

Amazon OpenSearch Ingestion e API degli endpoint di interfaccia ()AWS PrivateLink

Puoi stabilire una connessione privata tra il tuo VPC e gli endpoint dell'API OpenSearch Ingestion creando un endpoint VPC di interfaccia. Endpoint di interfaccia con tecnologia AWS PrivateLink.

AWS PrivateLink consente di accedere in modo privato alle operazioni dell'API OpenSearch Ingestion senza un gateway Internet, un dispositivo NAT, una connessione VPN o una connessione. Direct Connect Le risorse nel tuo VPC non necessitano di indirizzi IP pubblici per comunicare con gli endpoint dell'API OpenSearch Ingestion per creare, modificare o eliminare pipeline. Il traffico tra il tuo VPC e OpenSearch Ingestion non esce dalla rete Amazon.

> **Note**
>
> Questo argomento tratta gli endpoint VPC per l'accesso all'API OpenSearch Ingestion, che consente di gestire le pipeline (creazione, aggiornamento, eliminazione) dall'interno del VPC. Ciò è diverso dalla configurazione dell'accesso VPC per le pipeline stesse, che controlla il modo in cui i dati vengono importati nelle pipeline da fonti all'interno del tuo VPC. Per informazioni sulla configurazione dell'accesso VPC per le pipeline, vedere. the section called “Configurazione dell'accesso VPC per le pipeline”

Ogni endpoint dell'interfaccia è rappresentato da una o più interfacce di rete elastiche nelle sottoreti. Per maggiori informazioni sulle interfacce di rete elastiche, consulta le interfacce di rete elastiche nella Guida dell'utente di Amazon EC2.

Per ulteriori informazioni sugli endpoint dei VPC, consulta Endpoint VPC di interfaccia (AWS PrivateLink) nella Guida per l'utente di Amazon VPC. Per ulteriori informazioni sulle operazioni dell'API OpenSearch Ingestion, consulta il riferimento all'API Ingestion. OpenSearch

Considerazioni sugli endpoint VPC

Prima di configurare un endpoint VPC di interfaccia per gli endpoint dell'API OpenSearch Ingestion, assicurati di esaminare le proprietà e le limitazioni degli endpoint dell'interfaccia nella Amazon VPC User Guide.

Tutte le operazioni OpenSearch dell'API Ingestion relative alla gestione delle risorse di OpenSearch Ingestion sono disponibili tramite il tuo VPC utilizzando. AWS PrivateLink

Le policy degli endpoint VPC sono supportate per gli endpoint dell'API OpenSearch Ingestion. Per impostazione predefinita, l'accesso completo alle operazioni dell'API OpenSearch Ingestion è consentito tramite l'endpoint. Per ulteriori informazioni, consulta Controllo degli accessi ai servizi con endpoint VPC nella Guida per l'utente di Amazon VPC.

Disponibilità

OpenSearch L'API Ingestion attualmente supporta gli endpoint VPC in tutte le regioni di ingestione. OpenSearch

Al momento, gli endpoint FIPS non sono supportati.

Creazione di un endpoint VPC di interfaccia per l'API Ingestion OpenSearch

Puoi creare un endpoint VPC per l'API OpenSearch Ingestion utilizzando la console Amazon VPC o (). AWS Command Line Interface AWS CLI Per ulteriori informazioni, consultare Creazione di un endpoint di interfaccia nella Guida per l'utente di Amazon VPC.

Crea un endpoint VPC per l'API OpenSearch Ingestion utilizzando il nome del servizio. `com.amazonaws.region.osis`

Se abiliti il DNS privato per l'endpoint, puoi effettuare richieste API a OpenSearch Ingestion con l'endpoint VPC utilizzando il nome DNS predefinito per la regione, ad esempio. AWS `osis.us-east-1.amazonaws.com`

Per ulteriori informazioni, consultare Accesso a un servizio tramite un endpoint di interfaccia in Guida per l'utente di Amazon VPC.

Creazione di una policy di endpoint VPC per l'API Ingestion OpenSearch

Puoi allegare una policy per gli endpoint all'endpoint VPC che controlla l'accesso all' OpenSearch API Ingestion. La policy specifica le informazioni riportate di seguito:

- Il principale che può eseguire operazioni.
- Le azioni che possono essere eseguite.
- Le risorse sui cui si possono eseguire operazioni.

Per ulteriori informazioni, consulta Controllo degli accessi ai servizi con endpoint VPC in Guida per l'utente di Amazon VPC.

Esempio: policy degli endpoint VPC per OpenSearch le azioni dell'API Ingestion

Di seguito è riportato un esempio di policy sugli endpoint per l'API Ingestion. OpenSearch Se collegata a un endpoint, questa policy consente l'accesso alle azioni API di OpenSearch Ingestion elencate per tutti i principali su tutte le risorse.

```
{
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "osis:CreatePipeline",
            "osis:UpdatePipeline",
            "osis:DeletePipeline"
         ],
         "Resource":"*"
      }
   ]
}
```

Esempio: policy degli endpoint VPC che nega tutti gli accessi da un account specificato AWS

La seguente politica degli endpoint VPC nega all' AWS account `123456789012` tutti gli accessi alle risorse che utilizzano l'endpoint. La policy consente tutte le operazioni da altri account.

```
{
  "Statement": [
    {
      "Action": "*",
      "Effect": "Allow",
      "Resource": "*",
      "Principal": "*"
    },
    {
      "Action": "*",
      "Effect": "Deny",
      "Resource": "*",
      "Principal": { "AWS": [ "123456789012" ] }
```

```
      }
    ]
}
```

Etichettatura delle pipeline di Amazon OpenSearch Ingestion

I tag consentono di assegnare informazioni arbitrarie a una pipeline di Amazon OpenSearch Ingestion in modo da poter classificare e filtrare tali informazioni. Un tag è un'etichetta di metadati che assegni o che assegni a una risorsa. AWS AWS Ciascun tag è formato da una chiave e da un valore, Per i tag assegnati da te, puoi definire la chiave e il valore. Ad esempio, potresti definire la chiave come `stagee` il valore di una risorsa come `test`.

I tag consentono di:

- Identifica e organizza le tue risorse. AWS Molti AWS servizi supportano l'etichettatura, quindi puoi assegnare lo stesso tag a risorse di servizi diversi per indicare che le risorse sono correlate. Ad esempio, puoi assegnare lo stesso tag a una pipeline di OpenSearch Ingestion che assegni a un dominio Amazon Service. OpenSearch
- Tieni traccia dei costi. AWS Attivi questi tag sulla Gestione dei costi e fatturazione AWS dashboard. AWS utilizza i tag per classificare i costi e fornirti un rapporto mensile sull'allocazione dei costi. Per ulteriori informazioni, consulta la pagina sull'utilizzo dei tag per l'allocazione dei costi nella Guida per l'utente di AWS Billing.
- Limita l'accesso alle pipeline utilizzando il controllo degli accessi basato sugli attributi. Per ulteriori informazioni, consulta Controllo dell'accesso in base alle chiavi di tag nella Guida per l'utente IAM.

In OpenSearch Ingestion, la risorsa principale è una pipeline. Puoi utilizzare la console di OpenSearch servizio, la AWS CLI, OpenSearch Ingestion o AWS SDKs per aggiungere APIs, gestire e rimuovere tag da una pipeline.

Argomenti

- Autorizzazioni richieste
- Utilizzo dei tag (console)
- Utilizzo dei tag (AWS CLI)

Autorizzazioni richieste

OpenSearch Ingestion utilizza le seguenti autorizzazioni AWS Identity and Access Management Access Analyzer (IAM) per etichettare le pipeline:

- `osis:TagResource`
- `osis:ListTagsForResource`
- `osis:UntagResource`

Per ulteriori informazioni su ciascuna autorizzazione, consulta Azioni, risorse e chiavi di condizione per OpenSearch Ingestion nel Service Authorization Reference.

Utilizzo dei tag (console)

La console è il modo più semplice per etichettare una pipeline.

Per creare un tag

1. Accedi alla console di Amazon OpenSearch Service su https://console.aws.amazon.com/aos/osis/home. Ti troverai nella pagina Pipelines.
2. Seleziona la pipeline a cui desideri aggiungere i tag e vai alla scheda Tag.
3. Scegli Gestisci, quindi seleziona Aggiungi nuovo tag.
4. Inserire una chiave di tag e un valore facoltativo.
5. Scegli Save (Salva).

Per eliminare un tag, esegui la stessa procedura e scegli Rimuovi nella pagina Gestisci tag.

Per ulteriori informazioni sull'utilizzo della console per il funzionamento con i tag, consulta Editor di tag nella Guida alle operazioni di base della Console di gestione AWS .

Utilizzo dei tag (AWS CLI)

Per etichettare una pipeline utilizzando il AWS CLI, invia una `TagResource` richiesta:

```
aws osis tag-resource
  --arn arn:aws:osis:us-east-1:123456789012:pipeline/my-pipeline
  --tags Key=service,Value=osis Key=source,Value=otel
```

Rimuovi i tag da una pipeline usando il `UntagResource` comando:

```
aws osis untag-resource
  --arn arn:aws:osis:us-east-1:123456789012:pipeline/my-pipeline
  --tag-keys service
```

Visualizza i tag esistenti per una pipeline con il `ListTagsForResource` comando:

```
aws osis list-tags-for-resource
  --arn arn:aws:osis:us-east-1:123456789012:pipeline/my-pipeline
```

Registrazione e monitoraggio di Amazon OpenSearch Ingestion con Amazon CloudWatch

Amazon OpenSearch Ingestion pubblica metriche e log su Amazon. CloudWatch

Argomenti

- Monitoraggio dei log della pipeline
- Monitoraggio delle metriche della pipeline

Monitoraggio dei log della pipeline

Puoi abilitare la registrazione per le pipeline di Amazon OpenSearch Ingestion per esporre i messaggi di errore e di avviso generati durante le operazioni di pipeline e l'attività di ingestione. OpenSearch Ingestion pubblica tutti i log su Amazon Logs. CloudWatch CloudWatch I log possono monitorare le informazioni nei file di registro e avvisarti quando vengono raggiunte determinate soglie. Puoi inoltre archiviare i dati del log in storage estremamente durevole. Per ulteriori informazioni, consulta la Amazon CloudWatch Logs User Guide.

I log di OpenSearch Ingestion potrebbero indicare un'elaborazione non riuscita delle richieste, errori di autenticazione dall'origine al sink e altri avvisi che possono essere utili per la risoluzione dei problemi. Per i suoi registri, OpenSearch Ingestion utilizza i livelli di registro di,, e. `INFO` `WARN` `ERROR` `FATAL` Si consiglia di abilitare la pubblicazione dei log per tutte le pipeline.

Autorizzazioni richieste

Per consentire a OpenSearch Ingestion di inviare i log ai CloudWatch registri, è necessario accedere come utente con determinate autorizzazioni IAM.

Sono necessarie le seguenti autorizzazioni per CloudWatch i log per creare e aggiornare le risorse di consegna dei log:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Resource": "*",
            "Action": [
                "logs:CreateLogDelivery",
                "logs:PutResourcePolicy",
                "logs:UpdateLogDelivery",
                "logs:DeleteLogDelivery",
                "logs:DescribeResourcePolicies",
                "logs:GetLogDelivery",
                "logs:ListLogDeliveries"
            ]
        }
    ]
}
```

Attivazione della pubblicazione dei log

È possibile abilitare la pubblicazione dei log su pipeline esistenti o durante la creazione di una pipeline. Per i passaggi per abilitare la pubblicazione dei log durante la creazione della pipeline, consulta. the section called “Creazione di pipeline”

Console

Per abilitare la pubblicazione dei log su una pipeline esistente

1. Accedi alla console di Amazon OpenSearch Service su https://console.aws.amazon.com/aos/osis/home. Ti troverai nella pagina Pipelines.
2. Apri la pipeline in cui desideri abilitare i log, quindi scegli Azioni, Modifica le opzioni di pubblicazione dei log.
3. Abilita la pubblicazione nei registri CloudWatch .

4. Crea un nuovo gruppo di log o selezionane uno esistente. Si consiglia di formattare il nome come percorso, ad esempio `/aws/vendedlogs/OpenSearchIngestion/pipeline-name/audit-logs`. Questo formato semplifica l'applicazione di una politica di CloudWatch accesso che concede le autorizzazioni a tutti i gruppi di log in un percorso specifico, ad esempio. `/aws/vendedlogs/OpenSearchIngestion`

   > **Important**
   >
   > È necessario includere il prefisso `vendedlogs` nel nome del gruppo di log, altrimenti la creazione non riesce.

5. Scegli Save (Salva).

CLI

Per abilitare la pubblicazione dei log utilizzando il AWS CLI, invia la seguente richiesta:

```
aws osis update-pipeline \
  --pipeline-name my-pipeline \
  --log-publishing-options  IsLoggingEnabled=true,CloudWatchLogDestination={LogGroup="/
aws/vendedlogs/OpenSearchIngestion/pipeline-name"}
```

Monitoraggio delle metriche della pipeline

Puoi monitorare le pipeline di Amazon OpenSearch Ingestion utilizzando Amazon CloudWatch, che raccoglie dati grezzi e li elabora in metriche leggibili quasi in tempo reale. Queste statistiche vengono conservate per un periodo di 15 mesi, per permettere l'accesso alle informazioni storiche e offrire una prospettiva migliore sulle prestazioni del servizio o dell'applicazione web. È anche possibile impostare allarmi che controllano determinate soglie e inviare notifiche o intraprendere azioni quando queste soglie vengono raggiunte. Per ulteriori informazioni, consulta la Amazon CloudWatch User Guide.

La console OpenSearch Ingestion mostra una serie di grafici basati sui dati grezzi della scheda Performance per ogni pipeline. CloudWatch

OpenSearch Ingestion riporta le metriche della maggior parte dei plugin supportati. Se alcuni plugin non dispongono di una tabella propria qui sotto, significa che non riportano alcuna metrica specifica del plug-in. Le metriche della pipeline vengono pubblicate nel namespace. `AWS/OSIS`

Argomenti

- Metriche comuni
- Metriche del buffer
- Metriche Signature V4
- Metriche del buffer di blocco limitato
- Metriche del codice sorgente di Otel Trace
- Otel metrics (metriche di origine)
- Metriche Http
- Metriche S3
- Parametro Aggregate
- Metriche relative alla data
- parametri Lambda
- Metriche Grok
- Metriche grezze di Otel Trace
- Metriche del gruppo Otel trace
- Metriche statiche della mappa dei servizi
- OpenSearch metriche
- Metriche di sistema e misurazione

Metriche comuni

Le seguenti metriche sono comuni a tutti i processori e i sink.

Ogni metrica è preceduta dal nome della sub-pipeline e dal nome del plugin, nel formato < >< >< >. *sub_pipeline_name plugin metric_name* Ad esempio, il nome completo della `recordsIn.count` metrica per una sottopipeline denominata e il processore di data sarebbero. `my-pipeline` `my-pipeline.date.recordsIn.count`

Suffisso metrico	Description
`recordsIn.count`	L'ingresso dei record in un componente della pipeline. Questa metrica si applica ai processori e ai sink. Statistiche rilevanti: Sum (Somma)

Suffisso metrico	Description
	Dimensione: `PipelineName`
`recordsOut.count`	L'uscita dei record da un componente della pipeline. Questa metrica si applica ai processori e alle sorgenti. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`timeElapsed.count`	Un conteggio dei punti dati registrati durante l'esecuzione di un componente della pipeline. Questa metrica si applica ai processori e ai sink. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`timeElapsed.sum`	Il tempo totale trascorso durante l'esecuzione di un componente della pipeline. Questa metrica si applica ai processori e ai sink, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`timeElapsed.max`	Il tempo massimo trascorso durante l'esecuzione di un componente della pipeline. Questa metrica si applica ai processori e ai sink, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`

Metriche del buffer

Le seguenti metriche si applicano al buffer di blocco Bounded predefinito che OpenSearch Ingestion configura automaticamente per tutte le pipeline.

Ogni metrica è preceduta dal nome della subpipeline e dal nome del buffer, nel formato < >< >< >. *sub_pipeline_name buffer_name metric_name* Ad esempio, il nome completo della `recordsWritten.count` metrica per una sotto-pipeline denominata sarebbe. `my-pipeline` `my-pipeline.BlockingBuffer.recordsWritten.count`

Suffisso metrico	Description
`recordsWritten.count`	Il numero di record scritti in un buffer. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`recordsRead.count`	Il numero di record letti da un buffer. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`recordsInFlight.value`	Il numero di record non controllati letti da un buffer. Statistiche pertinenti: media Dimensione: `PipelineName`
`recordsInBuffer.value`	Il numero di record attualmente presenti in un buffer. Statistiche pertinenti: media Dimensione: `PipelineName`
`recordsProcessed.count`	Il numero di record letti da un buffer ed elaborati da una pipeline. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`recordsWriteFailed.count`	Il numero di record che la pipeline non è riuscita a scrivere nel sink. Statistiche rilevanti: Sum (Somma)

Suffisso metrico	Description
	Dimensione: `PipelineName`
`writeTimeElapsed.count`	Un conteggio dei punti dati registrati durante la scrittura su un buffer. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`writeTimeElapsed.sum`	Il tempo totale trascorso durante la scrittura su un buffer, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`writeTimeElapsed.max`	Il tempo massimo trascorso durante la scrittura su un buffer, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`writeTimeouts.count`	Il conteggio dei timeout di scrittura in un buffer. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`readTimeElapsed.count`	Un conteggio dei punti dati registrati durante la lettura da un buffer. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`readTimeElapsed.sum`	Il tempo totale trascorso durante la lettura da un buffer, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`readTimeElapsed.max`	Il tempo massimo trascorso durante la lettura da un buffer, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`checkpointTimeElapsed.count`	Un conteggio dei punti dati registrati durante il checkpoint. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`checkpointTimeElapsed.sum`	Il tempo totale trascorso durante il checkpoint, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`checkpointTimeElapsed.max`	Il tempo massimo trascorso durante il checkpoint, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`

Metriche Signature V4

Le seguenti metriche si applicano all'endpoint di ingestione per una pipeline e sono associate ai plugin di origine (, e). `http` `otel_trace` `otel_metrics` Tutte le richieste all'endpoint di ingestione devono essere firmate utilizzando la versione 4 di Signature. Queste metriche possono aiutarti a

identificare i problemi di autorizzazione durante la connessione alla tua pipeline o a confermare che l'autenticazione è avvenuta correttamente.

Ogni metrica è preceduta dal nome della sottopipeline e. `osis_sigv4_auth` Ad esempio, *`sub_pipeline_name`*`.osis_sigv4_auth.httpAuthSuccess.count`.

Suffisso metrico	Description
`httpAuthSuccess.count`	Il numero di richieste Signature V4 riuscite alla pipeline. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`httpAuthFailure.count`	Il numero di richieste Signature V4 non riuscite alla pipeline. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`httpAuthServerError.count`	Il numero di richieste Signature V4 alla pipeline che hanno restituito errori del server. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Metriche del buffer di blocco limitato

Le seguenti metriche si applicano al buffer di blocco limitato. Ogni metrica è preceduta dal nome della sottopipeline e. `BlockingBuffer` Ad esempio, *`sub_pipeline_name`*`.BlockingBuffer.bufferUsage.value`.

Suffisso metrico	Description
`bufferUsage.value`	Percentuale di utilizzo di in `buffer_size` base al numero di record nel buffer. `buffer_size` rappresenta

Suffisso metrico	Description
	il numero massimo di record scritti nel buffer e i record in corso che non sono stati controllati. Statistiche pertinenti: media Dimensione: `PipelineName`

Metriche del codice sorgente di Otel Trace

Le seguenti metriche si applicano alla sorgente di OTel traccia. Ogni metrica è preceduta dal nome della sottopipeline e. `otel_trace_source` Ad esempio, *`sub_pipeline_name`*`.otel_trace_source.requestTimeouts.count`.

Suffisso metrico	Description
`requestTimeouts.count`	Il numero di richieste scadute. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestsReceived.count`	Il numero di richieste ricevute dal plugin. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`successRequests.count`	Il numero di richieste che sono state elaborate con successo dal plugin. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`badRequests.count`	Il numero di richieste con un formato non valido che sono state elaborate dal plugin. Statistiche rilevanti: Sum (Somma)

Suffisso metrico	Description
	Dimensione: `PipelineName`
`requestsTooLarge.count`	Il numero di richieste il cui numero di intervalli nel contenuto è superiore alla capacità del buffer. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`internalServerError.count`	Il numero di richieste elaborate dal plugin con un tipo di eccezione personalizzato. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestProcessDuration.count`	Un conteggio dei punti dati registrati durante l'elaborazione delle richieste dal plugin. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestProcessDuration.sum`	La latenza totale delle richieste elaborate dal plugin, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestProcessDuration.max`	La latenza massima delle richieste elaborate dal plugin, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`

Suffisso metrico	Description
`payloadSize.count`	Un conteggio della distribuzione delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`payloadSize.sum`	La distribuzione totale delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`payloadSize.max`	La distribuzione massima delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`

Otel metrics (metriche di origine)

Le seguenti metriche si applicano alla fonte delle metriche. OTel Ogni metrica è preceduta dal nome della pipeline secondaria e. `otel_metrics_source` Ad esempio, *`sub_pipeline_name`*`.otel_metrics_source.requestTimeouts.count`.

Suffisso metrico	Description
`requestTimeouts.count`	Il numero totale di richieste al plugin scadute. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestsReceived.count`	Il numero totale di richieste ricevute dal plugin. Statistiche rilevanti: Sum (Somma)

Suffisso metrico	Description
	Dimensione: `PipelineName`
`successRequests.count`	Il numero di richieste elaborate con successo (200 codice di stato della risposta) dal plugin. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestProcessDuration.count`	Un conteggio della latenza delle richieste elaborate dal plugin, in secondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestProcessDuration.sum`	La latenza totale delle richieste elaborate dal plugin, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestProcessDuration.max`	La latenza massima delle richieste elaborate dal plugin, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`payloadSize.count`	Un conteggio della distribuzione delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`payloadSize.sum`	La distribuzione totale delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`payloadSize.max`	La distribuzione massima delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`

Metriche Http

Le seguenti metriche si applicano all'origine HTTP. Ogni metrica è preceduta dal nome della pipeline secondaria e. `http` Ad esempio, *`sub_pipeline_name`*`.http.requestsReceived.count`.

Suffisso metrico	Description
`requestsReceived.count`	Il numero di richieste ricevute dall'endpoint. `/log/ingest` Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestsRejected.count`	Il numero di richieste rifiutate (codice di stato della risposta 429) dal plugin. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`successRequests.count`	Il numero di richieste elaborate con successo (200 codice di stato della risposta) dal plugin.

Suffisso metrico	Description
	Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`badRequests.count`	Il numero di richieste con tipo o formato di contenuto non valido (codice di stato della risposta 400) elaborate dal plug-in. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestTimeouts.count`	Il numero di richieste che scadono nel server di origine HTTP (codice di stato della risposta 415). Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestsTooLarge.count`	Il numero di richieste la cui dimensione degli eventi nel contenuto è superiore alla capacità del buffer (codice di stato della risposta 413). Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`internalServerError.count`	Il numero di richieste elaborate dal plugin con un tipo di eccezione personalizzato (codice di stato della risposta 500). Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`requestProcessDuration.count`	Un conteggio della latenza delle richieste elaborate dal plugin, in secondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestProcessDuration.sum`	La latenza totale delle richieste elaborate dal plugin, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestProcessDuration.max`	La latenza massima delle richieste elaborate dal plugin, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`payloadSize.count`	Un conteggio della distribuzione delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`payloadSize.sum`	La distribuzione totale delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`payloadSize.max`	La distribuzione massima delle dimensioni del payload delle richieste in entrata, in byte. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`

Metriche S3

Le seguenti metriche si applicano alla fonte S3. Ogni metrica è preceduta dal nome della pipeline secondaria e. `s3` Ad esempio, *`sub_pipeline_name`*`.s3.s3ObjectsFailed.count`.

Suffisso metrico	Description
`s3ObjectsFailed.count`	Il numero totale di oggetti S3 che il plugin non è riuscito a leggere. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3ObjectsNotFound.count`	Il numero di oggetti S3 che il plugin non è riuscito a leggere a causa di un `Not Found` errore di S3. Queste metriche contano anche ai fini della metrica. `s3ObjectsFailed` Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3ObjectsAccessDenied.count`	Il numero di oggetti S3 che il plugin non è riuscito a leggere a causa di un `Forbidden` errore `Access Denied` or di S3. Queste metriche contano anche ai fini della metrica. `s3ObjectsFailed` Statistiche rilevanti: Sum (Somma)

Suffisso metrico	Description
	Dimensione: `PipelineName`
`s3ObjectReadTimeElapsed.count`	Il tempo impiegato dal plugin per eseguire una richiesta GET per un oggetto S3, analizzarlo e scrivere eventi nel buffer. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3ObjectReadTimeElapsed.sum`	Il tempo totale impiegato dal plugin per eseguire una richiesta GET per un oggetto S3, analizzarlo e scrivere eventi nel buffer, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3ObjectReadTimeElapsed.max`	Il tempo massimo impiegato dal plugin per eseguire una richiesta GET per un oggetto S3, analizzarlo e scrivere eventi nel buffer, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`s3ObjectSizeBytes.count`	Il conteggio della distribuzione delle dimensioni degli oggetti S3, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3ObjectSizeBytes.sum`	La distribuzione totale delle dimensioni degli oggetti S3, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`s3ObjectSizeBytes.max`	La distribuzione massima delle dimensioni degli oggetti S3, in byte. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`s3ObjectProcessedBytes.count`	Il conteggio della distribuzione degli oggetti S3 elaborati dal plugin, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3ObjectProcessedBytes.sum`	La distribuzione totale degli oggetti S3 elaborati dal plugin, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3ObjectProcessedBytes.max`	La distribuzione massima degli oggetti S3 elaborati dal plugin, in byte. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`s3ObjectsEvents.count`	Il conteggio della distribuzione degli eventi S3 ricevuti dal plugin. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3ObjectsEvents.sum`	La distribuzione totale degli eventi S3 ricevuti dal plugin. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`s3ObjectsEvents.max`	La distribuzione massima degli eventi S3 ricevuti dal plugin. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`sqsMessageDelay.count`	Un conteggio dei punti dati registrati mentre S3 registra l'ora dell'evento per la creazione di un oggetto fino al momento in cui è stato completamente analizzato. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`sqsMessageDelay.sum`	Il periodo di tempo totale tra il momento in cui S3 registra l'ora di un evento per la creazione di un oggetto e il momento in cui viene analizzato completamente, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`sqsMessageDelay.max`	Il periodo di tempo massimo tra il momento in cui S3 registra l'ora di un evento per la creazione di un oggetto e il momento in cui viene analizzato completamente, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`s3ObjectsSucceeded.count`	Il numero di oggetti S3 che il plugin ha letto con successo. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`sqsMessagesReceived.count`	Il numero di messaggi Amazon SQS ricevuti dalla coda dal plug-in. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`sqsMessagesDeleted.count`	Il numero di messaggi Amazon SQS eliminati dalla coda dal plug-in. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`sqsMessagesFailed.count`	Il numero di messaggi Amazon SQS che il plug-in non è riuscito ad analizzare. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Parametro Aggregate

Le seguenti metriche si applicano al processore Aggregate. Ogni metrica è preceduta dal nome della pipeline secondaria e. `aggregate` Ad esempio, *`sub_pipeline_name`*`.aggregate.actionHandleEventsOut.count`.

Suffisso metrico	Description
`actionHandleEvents Out.count`	Il numero di eventi che sono stati restituiti dalla `handleEvent` chiamata all'azione configurata. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`actionHandleEvents Dropped.count`	Il numero di eventi che sono stati restituiti dalla `handleEvent` chiamata all'azione configurata.

Suffisso metrico	Description
	Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`actionHandleEvents ProcessingErrors.count`	Il numero di chiamate effettuate `handleEvent` per l'azione configurata che ha provocato un errore. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`actionConcludeGrou pEventsOut.count`	Il numero di eventi che sono stati restituiti dalla `concludeGroup` chiamata all'azione configurata. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`actionConcludeGrou pEventsDropped.count`	Il numero di eventi che non sono stati restituiti dalla `condludeGroup` chiamata all'azione configurata. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`actionConcludeGrou pEventsProcessingE rrors.count`	Il numero di chiamate effettuate `concludeGroup` per l'azione configurata che ha provocato un errore. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`currentAggregateGr oups.value`	Il numero attuale di gruppi. Questo indicatore diminuisce quando i gruppi vengono conclusi e aumenta quando un evento avvia la creazione di un nuovo gruppo. Statistiche pertinenti: media Dimensione: `PipelineName`

Metriche relative alla data

Le seguenti metriche si applicano al processore di dati. Ogni metrica è preceduta dal nome della pipeline secondaria e. `date` Ad esempio, *`sub_pipeline_name`*`.date.dateProcessingMatchSuccess.count`.

Suffisso metrico	Description
`dateProcessingMatchSuccess.count`	Il numero di record che corrispondono ad almeno uno dei modelli specificati nell'opzione di `match` configurazione. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`dateProcessingMatchFailure.count`	Il numero di record che non corrispondono a nessuno dei modelli specificati nell'opzione `match` di configurazione. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

parametri Lambda

Le seguenti metriche si applicano al AWS Lambdaprocessore. Ogni metrica è preceduta dal nome della sottopipeline e. `lambda` Ad esempio, *`sub_pipeline_name`*`.lambda.recordsSuccessfullySentToLambda.count`.

Suffisso metrico	Description
`recordsSuccessfullySentToLambda.count`	Il numero di record elaborati con successo dalla funzione Lambda. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`recordsFailedToSendToLambda.count`	Il numero di record che non sono stati inviati alla funzione Lambda.

Suffisso metrico	Description
	Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

`lambdaFunctionLatency.avg`
 	
`lambdaFunctionLatency.max`
	La latenza delle chiamate alla funzione Lambda.
	
	Statistiche pertinenti: media e massima
	
	Dimensione: `PipelineName`

`numberOfRequestsSucceeded.count`	Il numero totale di richieste di invocazione Lambda riuscite. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`numberOfRequestsFailed.count`	Il numero totale di richieste di chiamata Lambda non riuscite. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`requestPayloadSize.avg`	La dimensione dei payload delle richieste inviati a Lambda. Statistiche pertinenti: media Dimensione: `PipelineName`
`responsePayloadSize.avg`	La dimensione dei payload di risposta ricevuti da Lambda. Statistiche pertinenti: media Dimensione: `PipelineName`

Metriche Grok

Le seguenti metriche si applicano al processore Grok. Ogni metrica è preceduta dal nome della sottopipeline e. `grok` Ad esempio, *`sub_pipeline_name`*`.grok.grokProcessingMatch.count`.

Suffisso metrico	Description
`grokProcessingMatch.count`	Il numero di record che hanno trovato almeno uno schema corrisponde all'opzione di `match` configurazione. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`grokProcessingMismatch.count`	Il numero di record che non corrispondono a nessuno dei modelli specificati nell'opzione `match` di configurazione. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`grokProcessingErrors.count`	Il numero di errori di elaborazione dei record. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`grokProcessingTimeouts.count`	Il numero di record scaduti durante la corrispondenza. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`grokProcessingTime.count`	Un conteggio dei punti dati registrati mentre un singolo record corrispondeva ai modelli dell'opzione di `match` configurazione. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`grokProcessingTime.sum`	Il tempo totale impiegato da ogni singolo record per corrispondere ai modelli dell'opzione di `match` configurazione, in millisecondi. Statistiche rilevanti: Sum (Somma)

Suffisso metrico	Description
	Dimensione: `PipelineName`
`grokProcessingTime.max`	Il tempo massimo impiegato da ogni singolo record per corrispondere ai modelli dell'opzione di `match` configura zione, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`

Metriche grezze di Otel Trace

Le seguenti metriche si applicano al processore OTel trace raw. Ogni metrica è preceduta dal nome della sottopipeline e. `otel_trace_raw` Ad esempio, *`sub_pipeline_name`*`.otel_trace_raw.traceGroupCacheCount.value`.

Suffisso metrico	Description
`traceGroupCacheCou nt.value`	Il numero di gruppi di tracce nella cache dei gruppi di tracce. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`spanSetCount.value`	Il numero di span set nella raccolta di span set. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Metriche del gruppo Otel trace

Le seguenti metriche si applicano al processore del gruppo di OTel traccia. Ogni metrica è preceduta dal nome della sottopipeline e. `otel_trace_group` Ad esempio, *`sub_pipeline_name`*`.otel_trace_group.recordsInMissingTraceGroup.count`.

Suffisso metrico	Description
`recordsInMissingTraceGroup.count`	Il numero di record di ingresso mancanti nei campi del gruppo di traccia. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`recordsOutFixedTraceGroup.count`	Il numero di record in uscita con campi del gruppo di traccia che sono stati compilati correttamente. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`recordsOutMissingTraceGroup.count`	Il numero di record in uscita mancanti nei campi del gruppo di traccia. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Metriche statiche della mappa dei servizi

Le seguenti metriche si applicano al processore stateful di Service-map. Ogni metrica è preceduta dal nome della pipeline secondaria e. `service-map-stateful` Ad esempio, *`sub_pipeline_name`*`.service-map-stateful.spansDbSize.count`.

Suffisso metrico	Description
`spansDbSize.value`	Le dimensioni dei byte in memoria si estendono in MapDB nella durata corrente e in quella precedente della finestra. Statistiche pertinenti: media Dimensione: `PipelineName`

Suffisso metrico	Description
`traceGroupDbSize.value`	Le dimensioni dei byte in memoria dei gruppi di tracce in MapDB per la durata corrente e precedente della finestra. Statistiche pertinenti: media Dimensione: `PipelineName`
`spansDbCount.value`	Il conteggio degli intervalli in MapDB tra la durata corrente e quella precedente della finestra. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`traceGroupDbCount.value`	Il conteggio dei gruppi di tracce in MapDB nella durata attuale e precedente della finestra. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`relationshipCount.value`	Il numero di relazioni memorizzate nella durata corrente e in quella precedente della finestra. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

OpenSearch metriche

Le seguenti metriche si applicano al lavandino. OpenSearch Ogni metrica è preceduta dal nome della pipeline secondaria e. `opensearch` Ad esempio, *`sub_pipeline_name`*`.opensearch.bulkRequestErrors.count`.

Suffisso metrico	Description
`bulkRequestErrors.count`	Il numero totale di errori riscontrati durante l'invio di richieste in blocco.

Suffisso metrico	Description
	Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`documentsSuccess.count`	Il numero di documenti inviati correttamente al OpenSearch Servizio tramite richieste in blocco, compresi i nuovi tentativi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`documentsSuccessFirstAttempt.count`	Il numero di documenti inviati correttamente al OpenSearch Servizio tramite richiesta in blocco al primo tentativo. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`documentErrors.count`	Il numero di documenti che non sono stati inviati tramite richieste in blocco. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestFailed.count`	Il numero di richieste in blocco non riuscite. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestNumberOfRetries.count`	Il numero di tentativi di richieste in blocco non riuscite. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`bulkBadRequestErrors.count`	Il numero di `Bad Request` errori riscontrati durante l'invio di richieste in blocco. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestNotAllowedErrors.count`	Il numero di `Request Not Allowed` errori riscontrati durante l'invio di richieste in blocco. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestInvalidInputErrors.count`	Il numero di `Invalid Input` errori riscontrati durante l'invio di richieste in blocco. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestNotFoundErrors.count`	Il numero di `Request Not Found` errori riscontrati durante l'invio di richieste in blocco. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestTimeoutErrors.count`	Il numero di `Request Timeout` errori riscontrati durante l'invio di richieste in blocco. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`bulkRequestServerErrors.count`	Il numero di `Server Error` errori riscontrati durante l'invio di richieste in blocco. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestSizeBytes.count`	Un conteggio della distribuzione delle dimensioni del payload delle richieste di massa, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestSizeBytes.sum`	La distribuzione totale delle dimensioni del payload delle richieste in blocco, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestSizeBytes.max`	La distribuzione massima delle dimensioni del payload delle richieste in blocco, in byte. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`bulkRequestLatency.count`	Un conteggio dei punti dati registrati durante l'invio delle richieste al plug-in, inclusi i nuovi tentativi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`bulkRequestLatency.sum`	La latenza totale delle richieste inviate al plugin, inclusi i nuovi tentativi, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`bulkRequestLatency.max`	La latenza massima delle richieste inviate al plugin, inclusi i nuovi tentativi, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`s3.dlqS3RecordsSuccess.count`	Il numero di record inviati con successo alla coda di lettere morte di S3. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3.dlqS3RecordsFailed.count`	Il numero di ricorsi che non sono stati inviati alla coda di lettere morte di S3. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3.dlqS3RequestSuccess.count`	Il numero di richieste riuscite alla coda di lettere morte di S3. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`s3.dlqS3RequestFailed.count`	Il numero di richieste non riuscite alla coda di lettere morte di S3. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3.dlqS3RequestLatency.count`	Conteggio dei punti dati registrati durante l'invio delle richieste alla coda di lettere morte di S3, inclusi i nuovi tentativi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3.dlqS3RequestLatency.sum`	La latenza totale delle richieste inviate alla coda di lettere morte di S3, inclusi i nuovi tentativi, in millisecondi. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3.dlqS3RequestLatency.max`	La latenza massima delle richieste inviate alla coda di lettere morte di S3, inclusi i nuovi tentativi, in millisecondi. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`
`s3.dlqS3RequestSizeBytes.count`	Un conteggio della distribuzione delle dimensioni del payload delle richieste nella coda di lettere morte di S3, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`

Suffisso metrico	Description
`s3.dlqS3RequestSizeBytes.sum`	La distribuzione totale delle dimensioni del payload delle richieste nella coda di lettere morte di S3, in byte. Statistiche rilevanti: Sum (Somma) Dimensione: `PipelineName`
`s3.dlqS3RequestSizeBytes.max`	La distribuzione massima delle dimensioni del payload delle richieste nella coda di lettere morte di S3, in byte. Statistiche rilevanti: Max (Massimo) Dimensione: `PipelineName`

Metriche di sistema e misurazione

Le seguenti metriche si applicano all'intero OpenSearch sistema di ingestione. Queste metriche non sono precedute da nulla.

Metrica	Description
`system.cpu.usage.value`	La percentuale di utilizzo della CPU disponibile per tutti i nodi di dati. Statistiche pertinenti: media Dimensione:`PipelineName` ,`area`, `id`
`system.cpu.count.value`	La quantità totale di utilizzo della CPU per tutti i nodi di dati. Statistiche pertinenti: media Dimensione:`PipelineName` ,`area`, `id`
`jvm.memory.max.value`	La quantità massima di memoria che può essere utilizzata per la gestione della memoria, espressa in byte.

Metrica	Description
	Statistiche pertinenti: media Dimensione:`PipelineName` ,`area`, `id`
`jvm.memory.used.value`	La quantità totale di memoria utilizzata, in byte. Statistiche pertinenti: media Dimensione:`PipelineName` ,`area`, `id` signa
`jvm.memory.committed.value`	La quantità di memoria impegnata per l'uso da parte della macchina virtuale Java (JVM), espressa in byte. Statistiche pertinenti: media Dimensione:`PipelineName` ,`area`, `id`
`computeUnits`	Il numero di Ingestion OpenSearch Compute Units (Ingestion OCUs) utilizzate da una pipeline. Statistiche pertinenti: Max, Sum, Average Dimensione: `PipelineName`

Le migliori pratiche per Amazon OpenSearch Ingestion

Questo argomento fornisce le best practice per la creazione e la gestione delle pipeline di Amazon OpenSearch Ingestion e include linee guida generali che si applicano a molti casi d'uso. Ogni carico di lavoro è unico, con caratteristiche uniche, quindi nessun suggerimento generico è adatto per ogni caso d'uso.

Argomenti

- Best practice generali
- Allarmi consigliati CloudWatch

Best practice generali

Le seguenti best practice generali si applicano alla creazione e alla gestione di pipeline.

- Per garantire un'elevata disponibilità, configura le pipeline VPC con due o tre sottoreti. Se distribuisci una pipeline solo in una sottorete e la zona di disponibilità non funziona, non sarai in grado di importare dati.
- All'interno di ogni pipeline, consigliamo di limitare il numero di sotto-pipeline a 5 o meno.
- Se utilizzi il plug-in sorgente S3, utilizza file S3 di dimensioni uniformi per prestazioni ottimali.
- Se utilizzi il plug-in sorgente S3, aggiungi 30 secondi di timeout di visibilità aggiuntivo per ogni 0,25 GB di dimensione del file nel bucket S3 per prestazioni ottimali.
- Includi una dead-letter queue (DLQ) nella configurazione della pipeline in modo da poter scaricare gli eventi non riusciti e renderli accessibili per l'analisi. Se i tuoi sink rifiutano i dati a causa di mappature errate o altri problemi, puoi indirizzare i dati al DLQ per risolvere il problema e risolverlo.
- Se utilizzi un processore aggregato in una pipeline, ti consigliamo di utilizzare flag per prestazioni ottimali della pipeline. `“local_mode: true”`

Allarmi consigliati CloudWatch

CloudWatch gli allarmi eseguono un'azione quando una CloudWatch metrica supera un valore specificato per un certo periodo di tempo. Ad esempio, potresti voler AWS inviarti un'e-mail se lo stato di salute del cluster dura più `red` di un minuto. Questa sezione include alcuni allarmi consigliati per Amazon OpenSearch Ingestion e come rispondere ad essi.

Per ulteriori informazioni sulla configurazione degli allarmi, consulta Creating Amazon CloudWatch Alarms nella Amazon CloudWatch User Guide.

Allarme	Problema
`computeUnits` il massimo è = quello configurato `maxUnits` per 15 minuti, 3 volte consecutive	La pipeline ha raggiunto la capacità massima e potrebbe richiedere un `maxUnits` aggiornamento. Aumenta la capacità massima della tua pipeline
`opensearch.documen`	La pipeline non è in grado di scrivere nel OpenSearch sink. Controlla le autorizzazioni della pipeline e conferma che il dominio o la raccolta

Allarme	Problema
`tErrors.c ount` sum is = *`{sub_pipe line_name }`* `.opensear ch.record sIn.count` somma per 1 minuto, 1 volta consecutiva	siano integri. Puoi anche controllare la presenza di eventi non riusciti nella coda delle lettere morte (DLQ), se è configurata.
`bulkReque stLatency .max` max è >= x per 1 minuto, 1 volta consecutiva	La pipeline presenta un'elevata latenza nell'invio dei dati al sink. OpenSearch Ciò è probabilmente dovuto al fatto che il sink è sottodime nsionato o a una strategia di sharding inadeguata, che sta facendo sì che il sink rimanga indietro. Una latenza elevata e sostenuta può influire sulle prestazioni della pipeline e probabilmente portare a una contropre ssione sui client.
`httpAuthF ailure.co unt` somma >= 1 per 1 minuto, 1 volta consecutiva	Le richieste di ingestione non vengono autenticate. Verifica che l'autenti cazione Signature Version 4 sia abilitata correttamente per tutti i client.
`system.cp u.usage.v alue` media >= 80% per 15 minuti, 3 volte consecutive	Un utilizzo elevato e prolungato della CPU può essere problematico. Valuta la possibilità di aumentare la capacità massima della pipeline.
`bufferUsa ge.value` media >= 80% per 15 minuti, 3 volte consecutive	Un utilizzo prolungato e elevato del buffer può essere problematico. Valuta la possibilità di aumentare la capacità massima della pipeline.

Altri allarmi che potresti prendere in considerazione

Valuta la possibilità di configurare i seguenti allarmi a seconda delle funzionalità di Amazon OpenSearch Ingestion che utilizzi regolarmente.

Allarme	Problema
`dynamodb.exportJobFailure.count` somma 1	Il tentativo di attivare un'esportazione in Amazon S3 non è riuscito.
`opensearch.EndtoEndLatency.avg` media > X per 15 minuti, 4 volte consecutive	`EndtoEndLatency` È superiore a quello desiderato per la lettura da flussi DynamoDB. Ciò potrebbe essere causato da un OpenSearch cluster sottodimensionato o da una capacità OCU massima della pipeline troppo bassa per il throughput WCU sulla tabella DynamoDB. `EndtoEndLatency` sarà più alto dopo un'esportazione, ma dovrebbe diminuire nel tempo man mano che raggiunge gli ultimi stream DynamoDB.
`dynamodb.changeEventsProcessed.count` somma == 0 per X minuti	Nessun record viene raccolto dai flussi DynamoDB. Ciò potrebbe essere causato dall'assenza di attività sulla tabella o da un problema di accesso ai flussi DynamoDB.
`opensearch.s3.dlqS3RecordsSuccess.count` somma >= `opensearch.documentSuccess.count` somma per 1 minuto, 1 volta consecutiva	Al DLQ viene inviato un numero maggiore di record rispetto al OpenSearch sink. Esamina le metriche del plug-in OpenSearch sink per indagare e determinare la causa principale.

Allarme	Problema
`grok.grok Processin gTimeouts .count` sum = RecordsIn.count somma per 1 minuto, 5 volte consecutive	Il timeout di tutti i dati si verifica mentre il processore Grok tenta di creare una corrispondenza tra i modelli. È probabile che ciò influisca sulle prestazioni e rallenti la pipeline. Valuta la possibilità di modificare i tuoi schemi per ridurre i timeout.
`grok.grok Processin gErrors.count` la somma è >= 1 per 1 minuto, 1 volta consecutiva	Il processore Grok non riesce ad abbinare i modelli ai dati nella pipeline, con conseguenti errori. Rivedi i dati e le configurazioni del plug-in Grok per assicurarti che sia prevista la corrispondenza dei modelli.
`grok.grok Processin gMismatch .count` sum = RecordsIn.count somma per 1 minuto, 5 volte consecutive	Il processore Grok non è in grado di abbinare i modelli ai dati nella pipeline. Rivedi i dati e le configurazioni del plug-in Grok per assicurarti che sia prevista la corrispondenza dei modelli.
`date.date Processin gMatchFai lure.count` sum = RecordsIn.count = somma per 1 minuto, 5 volte consecutive	Il processore Date non è in grado di abbinare alcun modello ai dati nella pipeline. Controlla le configurazioni dei dati e del plug-in Date per assicurarti che il modello sia previsto.

Allarme	Problema
`s3.s3ObjectsFailed.count` somma >= 1 per 1 minuto, 1 volta consecutiva	Questo problema si verifica perché l'oggetto S3 non esiste o la pipeline non dispone di privilegi sufficienti. Esamina le `s3ObjectsAccessDenied.count` metriche `s3ObjectsNotFound.count` and per determinare la causa principale. Conferma che l'oggetto S3 esiste and/or e aggiorna le autorizzazioni.
`s3.sqsMessagesFailed.count` somma >= 1 per 1 minuto, 1 volta consecutiva	Il plug-in S3 non è riuscito a elaborare un messaggio Amazon SQS. Se hai un DLQ abilitato sulla coda SQS, esamina il messaggio di errore. La coda potrebbe ricevere dati non validi che la pipeline sta tentando di elaborare.
`http.badRequests.count` somma >= 1 per 1 minuto, 1 volte consecutive	Il client sta inviando una richiesta errata. Verifica che tutti i client stiano inviando il payload corretto.
`http.requestsTooLarge.count` somma >= 1 per 1 minuto, 1 volta consecutiva	Le richieste provenienti dal plugin sorgente HTTP contengono troppi dati, il che supera la capacità del buffer. Regola la dimensione del batch per i tuoi clienti.
`http.internalServerError.count` somma >= 0 per 1 minuto, 1 volta consecutiva	Il plugin di origine HTTP non riesce a ricevere gli eventi.

Allarme	Problema
`http.requestTimeouts.count` somma >= 0 per 1 minuto, 1 volta consecutiva	I timeout di origine sono probabilmente il risultato di un approvvigionamento insufficiente della pipeline. Valuta la possibilità di aumentare la pipeline `maxUnits` per gestire un carico di lavoro aggiuntivo.
`otel_trace.badRequests.count` somma >= 1 per 1 minuto, 1 volta consecutiva	Il client sta inviando una richiesta errata. Verifica che tutti i client stiano inviando il payload corretto.
`otel_trace.requestsTooLarge.count` somma >= 1 per 1 minuto, 1 volta consecutiva	Le richieste provenienti dal plugin sorgente di Otel Trace contengono troppi dati, il che supera la capacità del buffer. Regola la dimensione del batch per i tuoi clienti.
`otel_trace.internalServerError.count` somma >= 0 per 1 minuto, 1 volta consecutiva	Il plugin sorgente di Otel Trace non riesce a ricevere gli eventi.
`otel_trace.requestTimeouts.count` somma >= 0 per 1 minuto, 1 volta consecutiva	I timeout di origine sono probabilmente il risultato di un approvvigionamento insufficiente della pipeline. Valuta la possibilità di aumentare la pipeline `maxUnits` per gestire un carico di lavoro aggiuntivo.

Allarme	Problema
`otel_metrics.requestTimeouts.count` somma >= 0 per 1 minuto, 1 volta consecutiva	I timeout di origine sono probabilmente il risultato di un approvvigionamento insufficiente della pipeline. Valuta la possibilità di aumentare la pipeline `maxUnits` per gestire un carico di lavoro aggiuntivo.

Amazon OpenSearch Serverless

Amazon OpenSearch Serverless è una configurazione on-demand con scalabilità automatica per Amazon Service. OpenSearch A differenza dei OpenSearch domini forniti, che richiedono la gestione manuale della capacità, una raccolta OpenSearch Serverless ridimensiona automaticamente le risorse di elaborazione in base alle esigenze dell'applicazione.

OpenSearch Serverless offre una soluzione conveniente per carichi di lavoro poco frequenti, intermittenti o imprevedibili. Ottimizza i costi scalando automaticamente la capacità di elaborazione in base all'utilizzo dell'applicazione. Le raccolte serverless utilizzano lo stesso volume di storage ad alta capacità, distribuito e ad alta disponibilità dei domini di servizio forniti. OpenSearch

OpenSearch Le raccolte serverless sono sempre crittografate. Puoi scegliere la chiave di crittografia, ma non puoi disabilitare la crittografia. Per ulteriori informazioni, consulta the section called “Encryption (Crittografia)”

Vantaggi

OpenSearch Serverless offre i seguenti vantaggi:

- Più semplice del provisioning: OpenSearch Serverless elimina gran parte della complessità della gestione OpenSearch dei cluster e della capacità. Dimensiona e ottimizza automaticamente i cluster e si occupa della gestione del ciclo di vita delle partizioni e degli indici. Gestisce inoltre gli aggiornamenti del software di servizio e gli upgrade delle versioni. OpenSearch Tutti gli aggiornamenti e gli upgrade non comportano interruzioni.
- Conveniente: quando utilizzi OpenSearch Serverless, paghi solo per le risorse che consumi. Ciò elimina la necessità di un provisioning anticipato e di un overprovisioning per gestire i picchi dei carichi di lavoro.
- Alta disponibilità: OpenSearch Serverless supporta i carichi di lavoro di produzione con ridondanza per proteggere dalle interruzioni della Availability Zone e dai guasti dell'infrastruttura.
- Scalabile: OpenSearch Serverless ridimensiona automaticamente le risorse per mantenere tassi di ingestione dei dati e tempi di risposta alle query costantemente elevati.

Cos'è Amazon OpenSearch Serverless?

Amazon OpenSearch Serverless è un'opzione serverless on-demand per Amazon OpenSearch Service che elimina la complessità operativa del provisioning, della configurazione e del tuning dei cluster. OpenSearch È ideale per le organizzazioni che preferiscono non gestire autonomamente i propri cluster o che non dispongono delle risorse e delle competenze dedicate per gestire implementazioni su larga scala. Con OpenSearch Serverless, puoi cercare e analizzare grandi volumi di dati senza gestire l'infrastruttura sottostante.

Una raccolta OpenSearch Serverless è un gruppo di OpenSearch indici che interagiscono per supportare un carico di lavoro o un caso d'uso specifici. Le raccolte semplificano le operazioni rispetto ai OpenSearch cluster autogestiti, che richiedono il provisioning manuale.

Le raccolte utilizzano lo stesso storage ad alta capacità, distribuito e ad alta disponibilità dei domini di OpenSearch servizio forniti, ma riducono ulteriormente la complessità eliminando la configurazione e l'ottimizzazione manuali. I dati all'interno di una raccolta vengono crittografati durante il transito. OpenSearch Serverless supporta anche OpenSearch le dashboard, che forniscono un'interfaccia per l'analisi dei dati.

Attualmente, le raccolte serverless eseguono OpenSearch la versione 2.17.x. Man mano che vengono rilasciate nuove versioni, OpenSearch Serverless aggiorna automaticamente le raccolte per incorporare nuove funzionalità, correzioni di bug e miglioramenti delle prestazioni.

OpenSearch Serverless supporta le stesse operazioni API di acquisizione e interrogazione della suite OpenSearch open source, in modo da poter continuare a utilizzare i client e le applicazioni esistenti. I tuoi client devono essere compatibili con OpenSearch 2.x per poter funzionare con Serverless. OpenSearch Per ulteriori informazioni, consulta the section called “Importazione dei dati nelle raccolte”.

Argomenti

- Casi d'uso per Serverless OpenSearch
- Come funziona
- Scelta di un tipo di raccolta
- Prezzi
- Supportato Regioni AWS
- Limitazioni

- Service e Serverless a confronto OpenSearch OpenSearch

Casi d'uso per Serverless OpenSearch

OpenSearch Serverless supporta due casi d'uso principali:

- Analisi dei log: l'opzione dell'analisi dei log si concentra sull'analisi di grandi volumi di dati di serie temporali semistrutturati generati da macchine per approfondimenti sull'aspetto operativo e sul comportamento degli utenti.
- Ricerca full-text: l'opzione della ricerca full-text supporta le applicazioni nelle reti interne (ad esempio, per i sistemi di gestione dei contenuti, i documenti legali, ecc.) e le applicazioni connesse a Internet (ad esempio, per la ricerca di contenuti nei siti di e-commerce).

Quando crei una raccolta, scegli uno di questi casi d'uso. Per ulteriori informazioni, consulta the section called “Scelta di un tipo di raccolta”.

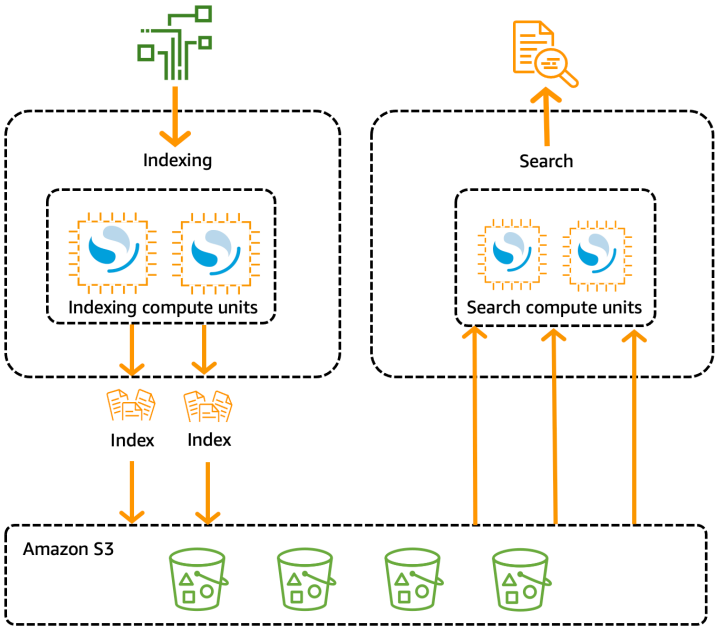
Come funziona

OpenSearch I cluster tradizionali dispongono di un unico set di istanze che eseguono sia operazioni di indicizzazione che di ricerca e lo storage degli indici è strettamente associato alla capacità di elaborazione. Al contrario, OpenSearch Serverless utilizza un'architettura nativa del cloud che separa i componenti di indicizzazione (inserimento) dai componenti di ricerca (query), con Amazon S3 come storage di dati principale per gli indici.

Questa architettura disaccoppiata consente di scalare le funzioni di ricerca e indicizzazione indipendentemente l'una dall'altra, e indipendentemente dai dati indicizzati in S3. Inoltre, l'architettura fornisce l'isolamento per le operazioni di importazione e interrogazione in modo che possano essere eseguite contemporaneamente, senza conflitti tra le risorse.

Quando scrivi dati su una raccolta, Serverless li distribuisce alle unità di calcolo di indicizzazione. OpenSearch Le unità di calcolo di indicizzazione importano i dati in entrata e spostano gli indici su S3. Quando si esegue una ricerca sui dati della raccolta, OpenSearch Serverless indirizza le richieste alle unità di calcolo di ricerca che contengono i dati interrogati. Le unità di calcolo di ricerca scaricano i dati indicizzati direttamente da S3 (se non sono già memorizzati nella cache locale), eseguono operazioni di ricerca ed effettuano aggregazioni.

L'immagine seguente illustra questa architettura disaccoppiata:



OpenSearch La capacità di elaborazione serverless per l'inserimento, la ricerca e l'interrogazione dei dati viene misurata in unità di calcolo (). OpenSearch OCUs Ogni OCU è una combinazione di 6 GiB di memoria e della CPU virtuale (vCPU) corrispondente, oltre al trasferimento dei dati su Amazon S3. Ogni OCU include una memoria temporanea a caldo sufficiente per 120 GiB di dati di indice.

Quando create la prima raccolta, OpenSearch Serverless ne crea due istanze, una per l'indicizzazione e una per la ricerca. OCUs Inoltre, avvia anche un set di nodi in standby in un'altra Zona di disponibilità per garantire un'elevata disponibilità. Per scopi di sviluppo e test, è possibile disabilitare l'impostazione Enable redondancy per una raccolta, che elimina le due repliche in standby e ne crea solo due. OCUs Per impostazione predefinita, le repliche attive ridondanti sono abilitate, il che significa che ne vengono create un totale di quattro OCUs per la prima raccolta in un account.

OCUs Esistono anche quando non c'è attività su alcun endpoint di raccolta. Tutte le raccolte successive li OCUs condividono. Quando crei raccolte aggiuntive nello stesso account, OpenSearch Serverless ne aggiunge altre OCUs per la ricerca e l'inserimento solo se necessario per supportare le raccolte, in base ai limiti di capacità specificati. La capacità si riduce man mano che l'utilizzo dell'elaborazione diminuisce.

Per informazioni su come ti vengono fatturati, consulta. OCUs the section called “Prezzi”

Scelta di un tipo di raccolta

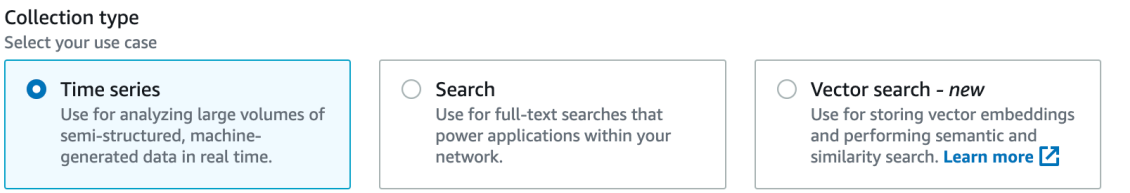
OpenSearch Serverless supporta tre tipi di raccolta principali:

Serie temporali: il segmento di analisi dei log che analizza grandi volumi di dati semistrutturati generati automaticamente in tempo reale, fornendo approfondimenti su operazioni, sicurezza, comportamento degli utenti e prestazioni aziendali.

Ricerca: ricerca di testo completo che abilita le applicazioni all'interno delle reti interne, come i sistemi di gestione dei contenuti e gli archivi di documenti legali, nonché le applicazioni connesse a Internet come la ricerca di siti di e-commerce e l'individuazione di contenuti.

Ricerca vettoriale: la ricerca semantica sugli incorporamenti vettoriali semplifica la gestione dei dati vettoriali e consente esperienze di ricerca aumentate basate sull'apprendimento automatico (ML). Supporta applicazioni di intelligenza artificiale generativa come chatbot, assistenti personali e rilevamento delle frodi.

Quando all'inizio crei una raccolta, scegli un tipo di raccolta:



Il tipo di raccolta scelto dipende dal tipo di dati che intendi importare nella raccolta e dal modo in cui intendi interrogarli. Non puoi modificare il tipo di raccolta dopo averla creata.

I tipi di raccolta presentano le seguenti differenze rilevanti:

- Per le raccolte di ricerca e di ricerca vettoriale, tutti i dati vengono archiviati in hot storage per garantire tempi di risposta rapidi alle query. Le raccolte di serie temporali utilizzano una combinazione di archiviazione ad accesso frequente e a caldo, in cui i dati più recenti vengono conservati in un'archiviazione ad accesso frequente per ottimizzare i tempi di risposta alle interrogazioni per i dati a cui, come suggerisce il nome, si accede più frequentemente.
- Per le raccolte di serie temporali e di ricerca vettoriale, non è possibile indicizzarle in base a un ID di documento personalizzato o aggiornarle tramite richieste upsert. Questa operazione è riservata ai casi d'uso della ricerca. Puoi invece eseguire l'aggiornamento in base all'ID del documento. Per ulteriori informazioni, consulta the section called “Operazioni e autorizzazioni API supportate OpenSearch”.
- Per le raccolte di ricerche e serie temporali, non puoi utilizzare indici di tipo k-NN.

Prezzi

AWS ti addebita i seguenti componenti Serverless: OpenSearch

- Calcolo dell'importazione dei dati
- Elaborazione di ricerche e query
- Archiviazione conservata in Amazon S3

Una OCU comprende 6 GB di RAM, vCPU GP3 , storage e trasferimento dati corrispondenti su Amazon S3. L'unità più piccola che ti può essere fatturata è 0,5 OCU. AWS fattura l'OCU su base oraria, con granularità di secondo livello. Nel tuo estratto conto, vedi una voce per il calcolo in OCU-hours con un'etichetta per l'inserimento dei dati e un'etichetta per la ricerca. AWS inoltre ti fattura su base mensile per i dati archiviati in Amazon S3. Non ti addebita alcun costo per l'utilizzo OpenSearch delle dashboard.

Quando crei una raccolta con repliche attive ridondanti, ti viene addebitato un minimo di 1 OCU (0,5 OCU × 2) per l'ingestione, comprese le repliche principali e di standby, e 1 OCU (0,5 OCU × 2) per la ricerca:

- 1 OCU (0,5 OCU × 2) per l'ingestione, inclusi quelli primari e quelli di standby
- 1 OCU (0,5 OCU × 2) per la ricerca

Se disabiliti le repliche attive ridondanti, ti verrà addebitato un minimo di 1 OCU (0,5 OCU x 2) per la prima raccolta nel tuo account. Tutte le raccolte successive possono condividerle. OCUs

OpenSearch Serverless ne aggiunge altre OCUs in incrementi di 1 OCU in base alla potenza di elaborazione e allo storage necessari per supportare le raccolte. Puoi configurare un numero massimo di OCUs per il tuo account per controllare i costi.

> **Note**
>
> Le raccolte con unique non AWS KMS keys possono essere condivise OCUs con altre raccolte.

OpenSearch Serverless tenta di utilizzare le risorse minime richieste per tenere conto dei cambiamenti dei carichi di lavoro. Il numero di unità OCUs fornite in qualsiasi momento può variare e

non è esatto. Nel tempo, l'algoritmo utilizzato da OpenSearch Serverless continuerà a migliorare per ridurre al minimo l'utilizzo del sistema.

Per i dettagli completi sui prezzi, consulta i prezzi OpenSearch di Amazon Service.

Supportato Regioni AWS

OpenSearch Serverless è disponibile in un sottoinsieme di Regioni AWS tale OpenSearch servizio disponibile in. Per un elenco delle regioni supportate, consulta gli endpoint e le quote di Amazon OpenSearch Service nel. Riferimenti generali di AWS

Limitazioni

OpenSearch Serverless presenta le seguenti limitazioni:

- Alcune operazioni OpenSearch API non sono supportate. Per informazioni, consulta the section called “Operazioni e autorizzazioni API supportate OpenSearch”.
- Alcuni OpenSearch plugin non sono supportati. Per informazioni, consulta the section called “OpenSearch Plugin supportati”.
- Al momento non è possibile migrare automaticamente i dati da un dominio di OpenSearch servizio gestito a una raccolta serverless. È necessario reindicizzare i dati da un dominio a una raccolta.
- L'accesso multi-account alle raccolte non è supportato. Non è possibile includere raccolte da altri account nelle tue policy di crittografia o accesso ai dati.
- I OpenSearch plugin personalizzati non sono supportati.
- Le istantanee automatiche sono supportate per le raccolte OpenSearch Serverless. Le istantanee manuali non sono supportate. Per ulteriori informazioni, consulta the section called “Backup delle raccolte tramite istantanee”.
- La ricerca e la replica tra Regioni non sono supportate.
- Ci sono dei limiti al numero di risorse serverless di cui è possibile disporre in un singolo account e in una sola regione. Vedi OpenSearch Quote serverless.
- L'intervallo di aggiornamento per gli indici nelle raccolte di ricerca vettoriale è di circa 60 secondi. L'intervallo di aggiornamento per gli indici nelle raccolte di ricerca e di serie temporali è di circa 10 secondi.
- Il numero di shard, il numero di intervalli e l'intervallo di aggiornamento non sono modificabili e vengono gestiti da Serverless. OpenSearch La strategia di sharding si basa sul tipo di raccolta e sul traffico. Ad esempio, una raccolta di serie temporali ridimensiona gli shard primari in base ai colli di bottiglia relativi al traffico di scrittura.

- Sono supportate le funzionalità geospaziali disponibili nelle versioni fino alla 2.1. OpenSearch

Service e Serverless a confronto OpenSearch OpenSearch

In OpenSearch Serverless, alcuni concetti e funzionalità sono diversi dalla funzionalità corrispondente per un dominio di servizio fornito. OpenSearch Ad esempio, una differenza importante è che OpenSearch Serverless non ha il concetto di cluster o nodo.

La tabella seguente descrive in che modo le funzionalità e i concetti importanti di OpenSearch Serverless differiscono dalla funzionalità equivalente in un dominio di servizio fornito. OpenSearch

Funzionalità	OpenSearch Servizio	OpenSearch Senza server
Domini invece di raccolte	Gli indici sono contenuti in domini, che sono cluster preimpostati. OpenSearch Per ulteriori informazioni, consulta Creazione e gestione dei domini.	Gli indici sono contenuti in raccolte, ossia raggruppamenti logici di indici che rappresen tano un carico di lavoro o un caso d'uso specifico. Per ulteriori informazioni, consulta the section called “Gestione delle raccolte”.
Tipi di nodi e gestione della capacità	Crei un cluster con i tipi di nodi che soddisfano le tue specifiche di costi e prestazio ni. Devi calcolare i tuoi requisiti di archiviazione e scegliere un tipo di istanza per il tuo dominio. Per ulteriori informazioni, consulta the section called “Dimensionamento dei domini”.	OpenSearch Serverless ridimensiona e fornisce automaticamente unità di calcolo aggiuntive per l'account in base all'utilizzo della capacità. Per ulteriori informazioni, consulta the section called “Gestione dei limiti di capacità”.
Fatturazione	Pagherai per ogni ora d'uso di un'istanza EC2 e per le dimensioni complessive di tutti i volumi di archiviazione EBS collegati alle istanze.	Per l'elaborazione relativa all'importazione dei dati, l'elaborazione per la ricerca e l'interro gazione, e lo spazio di archiviazione conservat o in S3, ricevi un addebito espresso in OCU all'ora.

Funzionalità	OpenSearch Servizio	OpenSearch Senza server
	Per ulteriori informazioni, consulta the section called “Prezzi”.	Per ulteriori informazioni, consulta the section called “Prezzi”.
Encryption (Crittografia)	La crittografia dei dati inattivi è facoltativa per i domini. Per ulteriori informazioni, consulta the section called “Crittografia dei dati a riposo”.	La crittografia dei dati inattivi è obbligatoria per le raccolte. Per ulteriori informazioni, consulta the section called “Encryption (Crittografia)”.
Controllo dell'accesso ai dati	L'accesso ai dati all'interno dei domini è determinato dalle policy IAM e dal controllo granulare degli accessi.	L'accesso ai dati all'interno delle raccolte è determinato dalle policy di accesso ai dati.
Operazioni supportate OpenSearch	OpenSearch Il servizio supporta un sottoinsieme di tutte le operazioni OpenSearch API. Per ulteriori informazioni, consulta the section called “Operazioni supportate”.	OpenSearch Serverless supporta un sottoinsi eme diverso di OpenSearch operazioni API. Per ulteriori informazioni, consulta the section called “Operazioni e plug-in supportati”.
Accesso ai pannelli di controllo	Accedi con nome utente e password. Per ulteriori informazioni, consulta the section called “Accesso alle OpenSearc h dashboard come utente principale”.	Se hai effettuato l'accesso alla AWS console e accedi all'URL della dashboard, accederai automaticamente. Per ulteriori informazioni, consulta the section called “Accesso ai OpenSearch pannelli di controllo”.

Funzionalità	OpenSearch Servizio	OpenSearch Senza server
APIs	Interagisci in modo programmatico con il OpenSearch servizio utilizzando le operazioni dell'API del OpenSearch servizio.	Interagisci a livello di codice con OpenSearch Serverless utilizzando le operazioni dell'API Serverless. OpenSearch
Accesso alla rete	Le impostazioni di rete per un dominio si applicano all'endpoint del dominio e all'endpoint Dashboards. OpenSearch L'accesso alla rete per entrambi è strettamente accoppiato.	Le impostazioni di rete per l'endpoint del dominio e l'endpoint OpenSearch Dashboards sono disaccoppiate. Puoi scegliere di non configurare l'accesso alla rete per i dashboard. OpenSearch Per ulteriori informazioni, consulta the section called “Accesso alla rete”.
Firma delle richieste	Utilizza i client REST di OpenSearch alto e basso livello per firmare le richieste. Specifica il nome del servizio come `es`.	Al momento, OpenSearch Serverless supporta un sottoinsieme di client supportati OpenSearch da Service. Quando firmi richieste, specifica il nome del servizio come `aoss`. L'intestazione `x-amz-content-sha256` è obbligatoria. Per ulteriori informazioni, consulta the section called “Altri clienti”.
OpenSearch aggiornamenti di versione	Aggiorna manualmente i tuoi domini non appena OpenSearch diventano disponibili nuove versioni. Sei responsabile di assicurarti che il tuo dominio soddisfi i requisiti di aggiornamento e di aver apportato eventuali modifiche importanti.	OpenSearch Serverless aggiorna automaticamente le tue raccolte a nuove versioni. OpenSearch Gli aggiornamenti non devono avvenire necessariamente appena diventa disponibile una nuova versione.

Funzionalità	OpenSearch Servizio	OpenSearch Senza server
Aggiornamenti del software del servizio	Non appena diventano disponibili, puoi applicare manualmente gli aggiornamenti del software del servizio al tuo dominio.	OpenSearch Serverless aggiorna automaticamente le tue raccolte per utilizzare le correzioni di bug, le funzionalità e i miglioramenti delle prestazioni più recenti.
Accesso VPC	Puoi effettuare il provisioning del tuo dominio all'interno di un VPC. Puoi anche creare endpoint OpenSearch VPC aggiuntivi gestiti dal servizio per accedere al dominio.	Crei uno o più endpoint VPC OpenSearch gestiti senza server per il tuo account. Quindi, puoi includere questi endpoint nelle policy di rete.
Autenticazione SAML	Puoi abilitare l'autenticazione SAML in base al dominio. Per ulteriori informazioni, consulta the section called “Autenticazione SAML per dashboard OpenSearch”.	Configurate uno o più provider SAML a livello di account, quindi includete l'utente e il gruppo associati nelle politiche di accesso ai dati. IDs Per ulteriori informazioni, consulta the section called “Autenticazione SAML”.
Transport Layer Security (TLS)	OpenSearch Il servizio supporta TLS 1.2 ma si consiglia di utilizzare TLS 1.3.	OpenSearch Serverless supporta TLS 1.2 ma si consiglia di utilizzare TLS 1.3.

Tutorial: Guida introduttiva ad Amazon OpenSearch Serverless

Questo tutorial illustra i passaggi di base per avviare rapidamente una raccolta di ricerche Amazon OpenSearch Serverless. Una raccolta di ricerche ti consente di potenziare le applicazioni nelle tue reti interne e le applicazioni connesse a Internet, come la ricerca di siti di e-commerce e la ricerca di contenuti.

Per informazioni su come utilizzare una raccolta di ricerca vettoriale, consulta. the section called “Lavorare con le raccolte di ricerca vettoriale” Per informazioni più dettagliate sull'uso delle raccolte, consulta the section called “Gestione delle raccolte” e gli altri argomenti di questa guida.

In questo tutorial completerai le seguenti fasi:

1. Configurazione delle autorizzazioni
2. Creazione di una raccolta
3. Caricamento e ricerca dei dati
4. Eliminazione della raccolta

> **Note**
>
> Ti consigliamo di utilizzare solo caratteri ASCII per il tuo`IndexName`. Se non utilizzi caratteri ASCII per i tuoi caratteri`IndexName`, le CloudWatch metriche `IndexName` in verranno convertite in un formato con codifica URL per caratteri non ASCII.

Fase 1: configurazione delle autorizzazioni

Per completare questo tutorial e per utilizzare OpenSearch Serverless in generale, devi disporre delle autorizzazioni IAM corrette. In questo tutorial creerai una raccolta, caricherai i dati e farai una ricerca, quindi eliminerai la raccolta.

L'utente o il ruolo devono avere una policy basata sull'identità allegata con le seguenti autorizzazioni minime:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Action": [
        "aoss:CreateCollection",
        "aoss:ListCollections",
        "aoss:BatchGetCollection",
        "aoss:DeleteCollection",
        "aoss:CreateAccessPolicy",
```

```
        "aoss:ListAccessPolicies",
        "aoss:UpdateAccessPolicy",
        "aoss:CreateSecurityPolicy",
        "aoss:GetSecurityPolicy",
        "aoss:UpdateSecurityPolicy",
        "iam:ListUsers",
        "iam:ListRoles"
      ],
      "Effect": "Allow",
      "Resource": "*"
    }
  ]
}
```

Per ulteriori informazioni sulle autorizzazioni IAM OpenSearch Serverless, consulta. the section called "Identity and Access Management"

Fase 2: creazione di una raccolta

Una raccolta è un gruppo di OpenSearch indici che interagiscono per supportare un carico di lavoro o un caso d'uso specifici.

Per creare una raccolta Serverless OpenSearch

1. Apri la console OpenSearch di Amazon Service a https://console.aws.amazon.com/aos/casa.
2. Scegli Collections (Raccolte) nel pannello di navigazione a sinistra e scegli Create collection (Crea raccolta).
3. Nomina la raccolta movies (film).
4. Per il tipo di raccolta, scegli Search (Cerca). Per ulteriori informazioni, consulta Scelta di un tipo di raccolta.
5. Per Sicurezza, scegli Standard create.
6. In Crittografia, seleziona Usa Chiave di proprietà di AWS. Questo è AWS KMS key quello che OpenSearch Serverless utilizzerà per crittografare i tuoi dati.
7. In Network (Rete), configura le impostazioni di rete per la raccolta.
   - Per il tipo di accesso, seleziona Public (Pubblico).

- Per il tipo di risorsa, scegli sia Abilita l'accesso agli OpenSearch endpoint che Abilita l'accesso ai dashboard. OpenSearch Poiché caricherai e cercherai dati utilizzando OpenSearch le dashboard, devi abilitare entrambe.

8. Scegli Next (Successivo).
9. Per Configure data access (Configura l'accesso ai dati), configura le impostazioni di accesso per la raccolta. Le policy di accesso ai dati consentono a utenti e ruoli di accedere ai dati all'interno di una raccolta. In questo tutorial, forniremo a un singolo utente le autorizzazioni necessarie per indicizzare e cercare i dati nella raccolta denominata movies.

   Crea una singola regola che fornisca l'accesso alla raccolta movies. Nomina la regola Accesso alla raccolta movies.
10. Scegli Aggiungi responsabili, utenti e ruoli IAM e seleziona l'utente o il ruolo che utilizzerai per accedere alle OpenSearch dashboard e indicizzare i dati. Scegli Save (Salva).
11. In Index permissions (Autorizzazioni relative all'indice), seleziona tutte le autorizzazioni.
12. Scegli Next (Successivo).
13. Per le impostazioni della policy di accesso, scegli Create a new data access policy (Crea una nuova policy di accesso ai dati) e assegna alla policy il nome movies.
14. Scegli Next (Successivo).
15. Controlla le impostazioni della raccolta e scegli Submit (Invia). Attendi alcuni minuti affinché lo stato della raccolta diventi `Active`.

Fase 3: Caricamento e ricerca dei dati

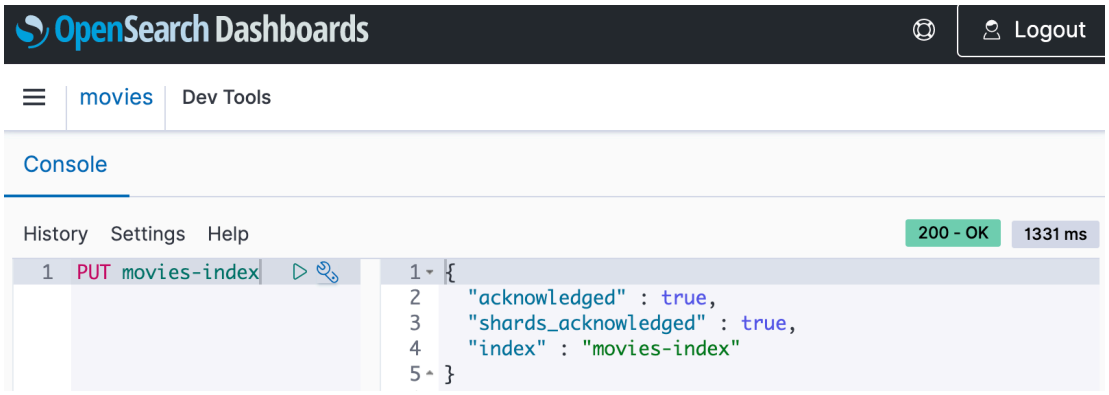
Puoi caricare dati in una raccolta OpenSearch Serverless utilizzando Postman o cURL. Per brevità, questi esempi utilizzano Dev Tools all'interno della console Dashboards. OpenSearch

Indicizzazione e ricerca dei dati nella raccolta movies

1. Scegli Collections (Raccolte) nel pannello di navigazione a sinistra e seleziona la raccolta movies per aprire la rispettiva pagina dei dettagli.
2. Scegli l'URL delle OpenSearch dashboard per la raccolta. L'URL assume il formato `https://dashboards.`*`{region}`*`.aoss.amazonaws.com/_login/?collectionId=`*`{collection-id}`*.
3. All'interno di OpenSearch Dashboards, apri il riquadro di navigazione a sinistra e scegli Dev Tools.

4. Per creare un singolo indice denominato movies-index, invia la seguente richiesta:

```
PUT movies-index
```



5. Per indicizzare un singolo documento in movies-index, invia la seguente richiesta:

```
PUT movies-index/_doc/1
{
  "title": "Shawshank Redemption",
  "genre": "Drama",
  "year": 1994
}
```

6. Per cercare dati nei OpenSearch dashboard, devi configurare almeno un modello di indice. OpenSearch utilizza questi modelli per identificare gli indici da analizzare. Apri il pannello di navigazione a sinistra, scegli Stack Management (Gestione stack), scegli Index Patterns (Modelli di indice), quindi scegli Create index pattern (Crea modello di indice). Per questo tutorial, digita movies.
7. Scegliere Fase successiva quindi selezionare Crea modello di indice. Dopo aver creato il modello, è possibile visualizzare i vari campi del documento, ad esempio `title` e `genre`.
8. Per iniziare a cercare i dati, apri di nuovo il pannello di navigazione a sinistra e scegli Discover (Rileva) o utilizza l'API di ricerca nei Dev Tools.

Gestione degli errori

Quando si eseguono operazioni di indicizzazione e ricerca, è possibile che vengano visualizzate le seguenti risposte di errore:

- `HTTP 507`— Indica che si è verificato un errore interno del server. Questo errore indica in genere che le unità di OpenSearch calcolo (OCUs) sono sovraccaricate dal volume o dalla complessità delle richieste. Sebbene OpenSearch Serverless sia scalabile automaticamente per gestire il carico, può verificarsi un ritardo nella distribuzione di risorse aggiuntive.

  Per mitigare questo errore, implementa una politica esponenziale di backoff e ripetizione dei tentativi. Questo approccio riduce temporaneamente la frequenza delle richieste per gestire efficacemente il carico. Per maggiori dettagli, consulta il comportamento di riprova nella Guida di riferimento agli strumenti AWS SDKs e agli strumenti.
- `HTTP 402`— Indica che è stato raggiunto il limite di capacità dell'unità di OpenSearch calcolo massima (OCU). Ottimizza il carico di lavoro per ridurre l'utilizzo dell'OCU o richiedere un aumento della quota.

Fase 4: Eliminazione della raccolta

Poiché la raccolta movies è stata creata per finalità di prova, è consigliabile eliminarla una volta terminate le prove in modo da evitare costi aggiuntivi.

Per eliminare una raccolta Serverless OpenSearch

1. Torna alla console di Amazon OpenSearch Service.
2. Scegli Collections (Raccolte) nel pannello di navigazione a sinistra e seleziona la raccolta movies.
3. Scegli Elimina e conferma l'eliminazione.

Fasi successive

Adesso che è chiaro come creare una raccolta e i dati di indice, è possibile provare a completare alcuni degli esercizi seguenti:

- Scopri ulteriori opzioni avanzate per la creazione di una raccolta. Per ulteriori informazioni, consulta the section called “Gestione delle raccolte”.
- Scopri come configurare le policy di sicurezza per gestire la sicurezza delle raccolte su larga scala. Per ulteriori informazioni, consulta the section called “Sicurezza in modalità serverless OpenSearch ”.

- Scopri altri modi per indicizzare i dati nelle raccolte. Per ulteriori informazioni, consulta the section called “Importazione dei dati nelle raccolte”.

Raccolte Amazon OpenSearch Serverless

Una raccolta in Amazon OpenSearch Serverless è un raggruppamento logico di uno o più indici che rappresentano un carico di lavoro di analisi. OpenSearch Serverless gestisce e ottimizza automaticamente la raccolta, richiedendo un input manuale minimo.

Argomenti

- Gestione delle raccolte Amazon OpenSearch Serverless
- Lavorare con le raccolte di ricerca vettoriale
- Utilizzo delle politiche del ciclo di vita dei dati con Amazon Serverless OpenSearch
- Utilizzo di AWS SDKs per interagire con Amazon OpenSearch Serverless
- Utilizzo CloudFormation per creare raccolte Amazon OpenSearch Serverless
- Backup delle raccolte tramite istantanee
- Supporto codec standard in Amazon Serverless OpenSearch
- Salva spazio di archiviazione utilizzando una fonte derivata

Gestione delle raccolte Amazon OpenSearch Serverless

Una raccolta in Amazon OpenSearch Serverless è un raggruppamento logico di uno o più indici che rappresentano un carico di lavoro di analisi. OpenSearch Serverless gestisce e ottimizza automaticamente la raccolta, richiedendo un input manuale minimo.

Argomenti

- Configurazione delle autorizzazioni per le raccolte
- Arricchimento semantico automatico per Serverless
- Creazione di raccolte
- Accesso ai OpenSearch pannelli di controllo
- Visualizzazione delle raccolte
- Eliminazione di raccolte

Configurazione delle autorizzazioni per le raccolte

OpenSearch Serverless utilizza le seguenti autorizzazioni AWS Identity and Access Management (IAM) per creare e gestire raccolte. È possibile specificare le condizioni IAM per limitare gli utenti a raccolte specifiche.

- `aoss:CreateCollection`: crea una raccolta.
- `aoss:ListCollections`: elenca le raccolte nell'account corrente.
- `aoss:BatchGetCollection`: ottieni dettagli su una o più raccolte.
- `aoss:UpdateCollection`: modifica una raccolta.
- `aoss:DeleteCollection`: elimina una raccolta.

La seguente policy di accesso basata sull'identità di esempio fornisce le autorizzazioni minime necessarie a un utente per gestire una singola raccolta denominata `Logs`:

```
[
   {
      "Sid":"Allows managing logs collections",
      "Effect":"Allow",
      "Action":[
         "aoss:CreateCollection",
         "aoss:ListCollections",
         "aoss:BatchGetCollection",
         "aoss:UpdateCollection",
         "aoss:DeleteCollection",
         "aoss:CreateAccessPolicy",
         "aoss:CreateSecurityPolicy"
      ],
      "Resource":"*",
      "Condition":{
         "StringEquals":{
            "aoss:collection":"Logs"
         }
      }
   }
]
```

`aoss:CreateAccessPolicy` e `aoss:CreateSecurityPolicy` sono inclusi perché le policy di crittografia, rete e accesso ai dati sono necessarie per il corretto funzionamento di una raccolta. Per ulteriori informazioni, consulta the section called “Identity and Access Management”.

> **Note**
>
> Se stai creando la prima raccolta nel tuo account, hai bisogno anche dell'autorizzazione `iam:CreateServiceLinkedRole`. Per ulteriori informazioni, consulta the section called "Ruolo di creazione della raccolta".

Arricchimento semantico automatico per Serverless

Introduzione

La funzionalità di arricchimento semantico automatico può contribuire a migliorare la pertinenza della ricerca fino al 20% rispetto alla ricerca lessicale. L'arricchimento semantico automatico elimina il peso indifferenziato della gestione dell'infrastruttura del modello ML (machine learning) e dell'integrazione con il motore di ricerca. La funzionalità è disponibile per tutti e tre i tipi di raccolta serverless: Search, Time Series e Vector.

Cos'è la ricerca semantica

I motori di ricerca tradizionali si basano sulla word-to-word corrispondenza (denominata ricerca lessicale) per trovare i risultati delle query. Sebbene funzioni bene per domande specifiche come i numeri di modello dei televisori, ha difficoltà con ricerche più astratte. Ad esempio, quando si cerca «scarpe per la spiaggia», una ricerca lessicale corrisponde semplicemente alle singole parole «scarpe», «spiaggia», «per» e «la» negli articoli del catalogo, e potenzialmente mancano prodotti pertinenti come «sandali impermeabili» o «calzature da surf» che non contengono i termini di ricerca esatti.

La ricerca semantica restituisce risultati di query che incorporano non solo la corrispondenza delle parole chiave, ma anche l'intento e il significato contestuale della ricerca dell'utente. Ad esempio, se un utente cerca «come curare un mal di testa», un sistema di ricerca semantico potrebbe restituire i seguenti risultati:

- Rimedi contro l'emicrania
- Tecniche di gestione del dolore
- Over-the-counter antidolorifici

Dettagli del modello e benchmark delle prestazioni

Sebbene questa funzionalità gestisca le complessità tecniche dietro le quinte senza esporre il modello sottostante, forniamo trasparenza attraverso una breve descrizione del modello e risultati di benchmark per aiutarvi a prendere decisioni informate sull'adozione delle funzionalità nei vostri carichi di lavoro critici.

L'arricchimento semantico automatico utilizza un modello sparso preaddestrato e gestito dai servizi che funziona in modo efficace senza richiedere regolazioni di precisione personalizzate. Il modello analizza i campi specificati, espandendoli in vettori sparsi basati su associazioni apprese da diversi dati di addestramento. I termini estesi e i relativi pesi di significatività vengono archiviati nel formato di indice Lucene nativo per un recupero efficiente. Abbiamo ottimizzato questo processo utilizzando la modalità solo documento, in cui la codifica avviene solo durante l'inserimento dei dati. Le query di ricerca vengono semplicemente tokenizzate anziché elaborate tramite il modello sparso, il che rende la soluzione conveniente e performante.

La nostra convalida delle prestazioni durante lo sviluppo delle funzionalità ha utilizzato il set di dati MS MARCO per il recupero dei passaggi, con passaggi della lunghezza media di 334 caratteri. Per quanto riguarda il punteggio di pertinenza, abbiamo misurato il guadagno cumulativo medio Normalized Discounted Cumulative Gain (NDCG) per i primi 10 risultati di ricerca (ndcg @10) sul benchmark BEIR per i contenuti in inglese e il ndcg @10 medio su MIRACL per i contenuti multilingue. Abbiamo valutato la latenza mediante misurazioni lato client del 90° percentile (p90) e la risposta di ricerca p90 ha rilevato dei valori. Questi benchmark forniscono indicatori prestazionali di base sia per la pertinenza della ricerca che per i tempi di risposta. Ecco i principali numeri di riferimento:

- Lingua inglese - Miglioramento della pertinenza del 20% rispetto alla ricerca lessicale. Ha inoltre ridotto la latenza di ricerca P90 del 7,7% rispetto alla ricerca lessicale (BM25 è di 26 ms e l'arricchimento semantico automatico è di 24 ms).
- Multilingue: miglioramento della pertinenza del 105% rispetto alla ricerca lessicale, mentre la latenza di ricerca P90 è aumentata del 38,4% rispetto alla ricerca lessicale (è di 26 ms e l'arricchimento semantico automatico è di 36 ms). BM25

Data la natura unica di ogni carico di lavoro, ti invitiamo a valutare questa funzionalità nel tuo ambiente di sviluppo utilizzando i tuoi criteri di benchmarking prima di prendere decisioni di implementazione.

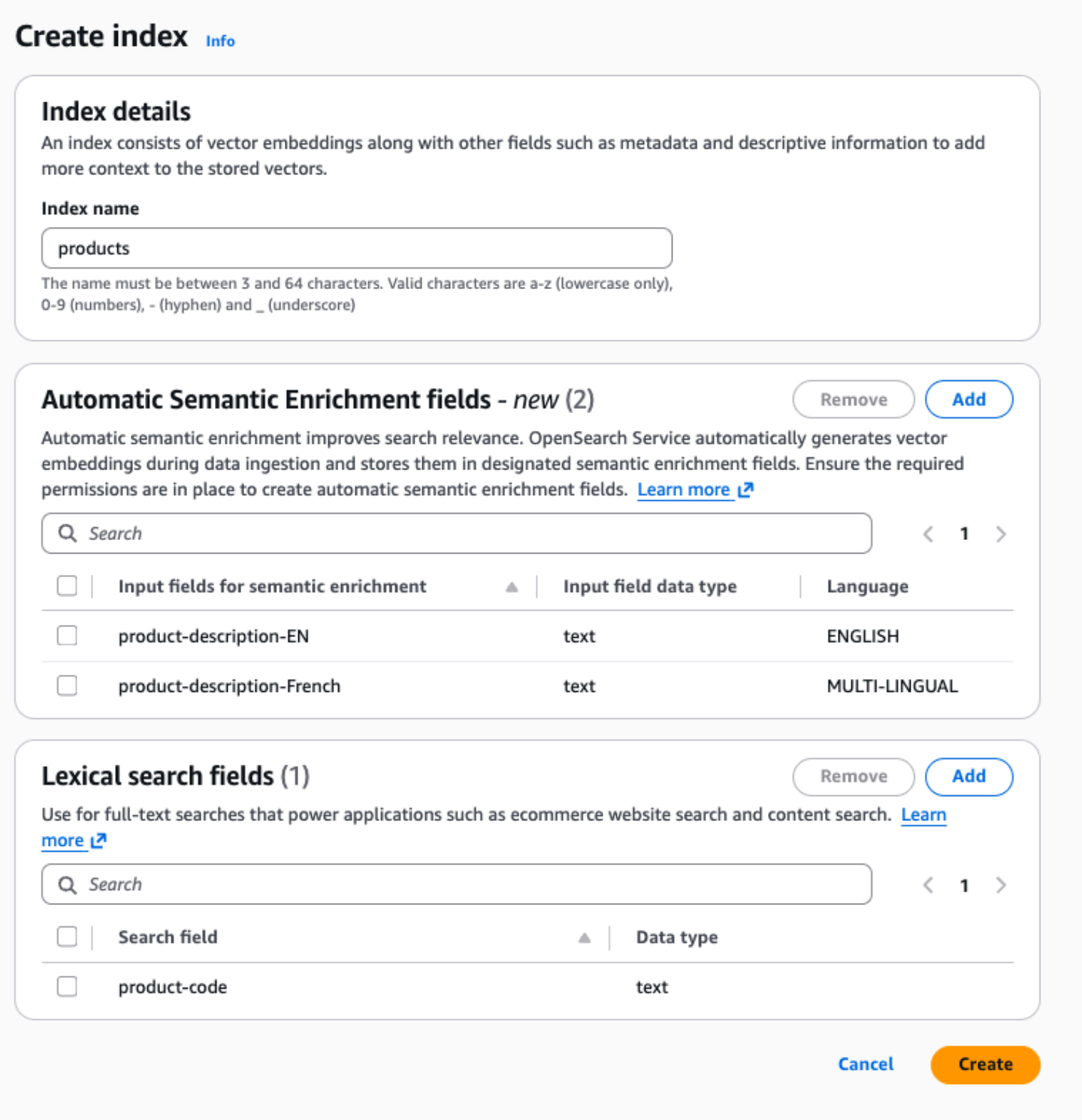
Lingue supportate

La funzionalità supporta l'inglese. Inoltre, il modello supporta anche arabo, bengalese, cinese, finlandese, francese, hindi, indonesiano, giapponese, coreano, persiano, russo, spagnolo, swahili e telugu.

Imposta un indice di arricchimento semantico automatico per le raccolte serverless

La configurazione di un indice con l'arricchimento semantico automatico abilitato per i campi di testo è semplice e può essere gestito tramite la console e i CloudFormation modelli durante la creazione di APIs nuovi indici. Per abilitarlo per un indice esistente, è necessario ricreare l'indice con l'arricchimento semantico automatico abilitato per i campi di testo.

Esperienza da console: la AWS console consente di creare facilmente un indice con campi di arricchimento semantico automatico. Dopo aver selezionato una raccolta, troverai il pulsante di creazione dell'indice nella parte superiore della console. Dopo aver fatto clic sul pulsante di creazione dell'indice, troverai le opzioni per definire campi di arricchimento semantico automatico. In un indice, puoi avere combinazioni di arricchimento semantico automatico per l'inglese e campi multilingue, oltre che lessicali.



Esperienza API - Per creare un indice di arricchimento semantico automatico utilizzando l'interfaccia a riga di comando ()AWS CLI, usa il AWS comando create-index:

```
aws opensearchserverless create-index \
--id [collection_id] \
```

```
--index-name [index_name] \
--index-schema [index_body] \
```

Nell'esempio seguente index-schema, il campo title_semantic ha un tipo di campo impostato su text e il parametro semantic_richment è impostato sullo stato ENABLED. L'impostazione del parametro semantic_enrichment abilita l'arricchimento semantico automatico nel campo title_semantic. Puoi usare il campo language_options per specificare l'inglese o il multilingue.

```
    aws opensearchserverless create-index \
    --id XXXXXXXXX \
    --index-name 'product-catalog' \
    --index-schema '{
    "mappings": {
        "properties": {
            "product_id": {
                "type": "keyword"
            },
            "title_semantic": {
                "type": "text",
                "semantic_enrichment": {
                    "status": "ENABLED",
                    "language_options": "english"
                }
            },
            "title_non_semantic": {
                "type": "text"
            }
        }
    }
}'
```

Per descrivere l'indice creato, usa il seguente comando:

```
aws opensearchserverless get-index \
--id [collection_id] \
--index-name [index_name] \
```

È inoltre possibile utilizzare CloudFormation i modelli (Tipo:AWS::OpenSearchServerless::CollectionIndex) per creare una ricerca semantica durante il provisioning della raccolta e dopo la creazione della raccolta.

Inserimento e ricerca di dati

Dopo aver creato un indice con l'arricchimento semantico automatico abilitato, la funzionalità funziona automaticamente durante il processo di inserimento dei dati, senza bisogno di configurazioni aggiuntive.

Inserimento di dati: quando aggiungi documenti all'indice, il sistema automaticamente:

- Analizza i campi di testo che hai designato per l'arricchimento semantico
- Genera codifiche semantiche utilizzando il modello sparse gestito dal servizio OpenSearch
- Memorizza queste rappresentazioni arricchite insieme ai dati originali

Questo processo utilizza i OpenSearch connettori ML e le pipeline di importazione integrati, che vengono creati e gestiti automaticamente dietro le quinte.

Ricerca: i dati di arricchimento semantico sono già indicizzati, quindi le query vengono eseguite in modo efficiente senza richiamare nuovamente il modello ML. Ciò significa che ottieni una maggiore pertinenza della ricerca senza alcun sovraccarico di latenza di ricerca aggiuntivo.

Configurazione delle autorizzazioni per l'arricchimento semantico automatico

Prima di creare un indice di arricchimento semantico automatizzato, è necessario configurare le autorizzazioni richieste. Questa sezione spiega le autorizzazioni necessarie e come configurarle.

Autorizzazioni della policy IAM

Utilizza la seguente policy AWS Identity and Access Management (IAM) per concedere le autorizzazioni necessarie per lavorare con l'arricchimento semantico automatico:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "AutomaticSemanticEnrichmentPermissions",
            "Effect": "Allow",
```

```
            "Action": [
                "aoss:CreateIndex",
                "aoss:GetIndex",
                "aoss:UpdateIndex",
                "aoss:DeleteIndex",
                "aoss:APIAccessAll"
            ],
            "Resource": "*"
        }
    ]
}
```

Autorizzazioni chiave

- Le `aoss:*Index` autorizzazioni consentono la gestione degli indici
- L'`aoss:APIAccessAll`autorizzazione consente le operazioni OpenSearch API
- Per limitare le autorizzazioni a una raccolta specifica, sostituiscile `"Resource": "*"` con l'ARN della raccolta

Configura le autorizzazioni di accesso ai dati

Per configurare un indice per l'arricchimento semantico automatico, è necessario disporre di politiche di accesso ai dati appropriate che consentano l'accesso a risorse di indicizzazione, pipeline e raccolta di modelli. Per ulteriori informazioni sulle politiche di accesso ai dati, vedere. Controllo dell'accesso ai dati per Amazon OpenSearch Serverless Per la procedura di configurazione di una politica di accesso ai dati, vedereCreazione di policy di accesso ai dati (console).

Autorizzazioni di accesso ai dati

```
[
    {
        "Description": "Create index permission",
        "Rules": [
            {
                "ResourceType": "index",
                "Resource": ["index/collection_name/*"],
                "Permission": [
                  "aoss:CreateIndex",
                  "aoss:DescribeIndex",
                  "aoss:UpdateIndex",
                  "aoss:DeleteIndex"
```

```
                ]
            }
        ],
        "Principal": [
            "arn:aws:iam::account_id:role/role_name"
        ]
    },
    {
        "Description": "Create pipeline permission",
        "Rules": [
            {
                "ResourceType": "collection",
                "Resource": ["collection/collection_name"],
                "Permission": [
                  "aoss:CreateCollectionItems",
                  "aoss:DescribeCollectionItems"
                ]
            }
        ],
        "Principal": [
            "arn:aws:iam::account_id:role/role_name"
        ]
    },
    {
        "Description": "Create model permission",
        "Rules": [
            {
                "ResourceType": "model",
                "Resource": ["model/collection_name/*"],
                "Permission": ["aoss:CreateMLResource"]
            }
        ],
        "Principal": [
            "arn:aws:iam::account_id:role/role_name"
        ]
    },
]
```

Autorizzazioni di accesso alla rete

Per consentire al servizio APIs di accedere alle raccolte private, è necessario configurare politiche di rete che consentano l'accesso richiesto tra l'API del servizio e la raccolta. Per ulteriori informazioni sulle politiche di rete, consulta Accesso alla rete per Amazon OpenSearch Serverless.

```
[
   {
      "Description":"Enable automatic semantic enrichment in a private collection",
      "Rules":[
         {
            "ResourceType":"collection",
            "Resource":[
               "collection/collection_name"
            ]
         }
      ],
      "AllowFromPublic":false,
      "SourceServices":[
         "aoss.amazonaws.com"
      ],
   }
]
```

Per configurare le autorizzazioni di accesso alla rete per una raccolta privata

1. Accedi alla console di OpenSearch servizio da https://console.aws.amazon.com/aos/casa.
2. Nella barra di navigazione a sinistra, scegli Politiche di rete. Poi esegui una delle seguenti operazioni:
   - Scegli il nome di una politica esistente e scegli Modifica
   - Scegli Crea politica di rete e configura i dettagli della politica
3. Nell'area Tipo di accesso, scegli Privato (consigliato), quindi seleziona Accesso privato al AWS servizio.
4. Nel campo di ricerca, scegli Servizio, quindi scegli aoss.amazonaws.com.
5. Nell'area Tipo di risorsa, seleziona la casella Abilita l'accesso all'endpoint. OpenSearch
6. Per Cerca raccolte o inserisci termini di prefisso specifici, nel campo di ricerca, seleziona Nome raccolta. Quindi inserisci o seleziona il nome delle raccolte da associare alla politica di rete.
7. Scegli Crea per una nuova politica di rete o Aggiorna per una politica di rete esistente.

La query riscrive

L'arricchimento semantico automatico converte automaticamente le query «match» esistenti in query di ricerca semantiche senza richiedere modifiche alle query. Se una query di corrispondenza fa parte

di una query composta, il sistema analizza la struttura delle query, trova le query di corrispondenza e le sostituisce con query sparse neurali. Attualmente, la funzionalità supporta solo la sostituzione delle query «match», indipendentemente dal fatto che si tratti di una query autonoma o di parte di una query composta. «multi_match» non è supportato. Inoltre, la funzionalità supporta tutte le query composte per sostituire le relative query di corrispondenza annidate. Le query composte includono: bool, boosting, constant_score, dis_max, function_score e hybrid.

Limitazioni dell'arricchimento semantico automatico

La ricerca semantica automatica è più efficace se applicata a campi di small-to-medium grandi dimensioni contenenti contenuti in linguaggio naturale, come titoli di film, descrizioni di prodotti, recensioni e riepiloghi. Sebbene la ricerca semantica aumenti la pertinenza per la maggior parte dei casi d'uso, potrebbe non essere ottimale per determinati scenari. Prendi in considerazione le seguenti limitazioni quando decidi se implementare l'arricchimento semantico automatico per il tuo caso d'uso specifico.

- Documenti molto lunghi: l'attuale modello sparso elabora solo i primi 8.192 token di ogni documento in inglese. Per i documenti multilingue, sono 512 token. Per articoli lunghi, prendi in considerazione l'implementazione della suddivisione in blocchi dei documenti per garantire l'elaborazione completa dei contenuti.
- Carichi di lavoro di analisi dei log: l'arricchimento semantico aumenta in modo significativo la dimensione dell'indice, il che potrebbe non essere necessario per l'analisi dei log, dove in genere è sufficiente una corrispondenza esatta. Il contesto semantico aggiuntivo raramente migliora l'efficacia della ricerca nei log abbastanza da giustificare i maggiori requisiti di archiviazione.
- L'arricchimento semantico automatico non è compatibile con la funzionalità Derived Source.
- Limitazione: le richieste di inferenza di indicizzazione sono attualmente limitate a 100 TPS per Serverless. OpenSearch Si tratta di un limite non vincolante; contatta il AWS servizio Support per ottenere limiti più elevati.

Prezzi

OpenSearch Serverless fattura l'arricchimento semantico automatico basato sulle unità di OpenSearch calcolo (OCUs) utilizzate durante la generazione di vettori sparsi al momento dell'indicizzazione. Ti viene addebitato solo l'utilizzo effettivo durante l'indicizzazione. Puoi monitorare questo consumo utilizzando la CloudWatch metrica SemanticSearch OCU di Amazon. Per dettagli specifici sui limiti dei token del modello, sul throughput di volume per OCU ed esempi di calcolo di esempio, visita Service Pricing. OpenSearch

Creazione di raccolte

Puoi utilizzare la console o il AWS CLI per creare una raccolta serverless. Questi passaggi spiegano come creare una raccolta di ricerche o di serie temporali. Per creare una raccolta di ricerca vettoriale, vedithe section called “Lavorare con le raccolte di ricerca vettoriale”.

Argomenti

- Creazione di una raccolta (console)
- Creazione di una raccolta (CLI)

Creazione di una raccolta (console)

Utilizzate le procedure descritte in questa sezione per creare una raccolta utilizzando. Console di gestione AWS Questi passaggi spiegano come creare una raccolta di ricerche o di serie temporali. Per creare una raccolta di ricerca vettoriale, vedithe section called “Lavorare con le raccolte di ricerca vettoriale”.

Argomenti

- Configurare le impostazioni della raccolta
- Configura campi di ricerca aggiuntivi

Configurare le impostazioni della raccolta

Utilizza la seguente procedura per configurare le informazioni sulla tua raccolta.

Per configurare le impostazioni della raccolta utilizzando la console

1. Accedi alla console di Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa/.
2. Espandi Serverless nel pannello di navigazione a sinistra e scegli Collections (Raccolte).
3. Scegli Create collection (Crea raccolta).
4. Fornisci un nome e una descrizione per la raccolta. Il nome deve soddisfare i seguenti criteri:
    - È unico per il tuo account e Regione AWS
    - Contiene solo lettere minuscole a-z, i numeri da 0 a 9 e i trattini (-)
    - Contiene da 3 a 32 caratteri

5. Scegli un tipo di raccolta:

   - Serie temporali: segmento di analisi dei log che si concentra sull'analisi di grandi volumi di dati semistrutturati generati da macchine. Almeno 24 ore di dati vengono archiviate su indici caldi e il resto rimane in una memoria a caldo.
   - Ricerca: l'opzione della ricerca full-text supporta le applicazioni nelle reti interne e le applicazioni connesse a Internet. Tutti i dati di ricerca vengono memorizzati in un'archiviazione ad accesso frequente per garantire tempi di risposta rapidi alle query.

     > **Note**
     >
     > Scegliete questa opzione se state abilitando la ricerca semantica automatica, come descritto in. Configurare le impostazioni della raccolta

   - Ricerca vettoriale: ricerca semantica su incorporamenti vettoriali che semplifica la gestione dei dati vettoriali. Potenzia le esperienze di ricerca aumentata di machine learning (ML) e le applicazioni di intelligenza artificiale generativa come chatbot, assistenti personali e rilevamento delle frodi.

   Per ulteriori informazioni, consulta the section called “Scelta di un tipo di raccolta”.
6. Per il tipo di implementazione, scegli l'impostazione di ridondanza per la tua raccolta. Per impostazione predefinita, ogni raccolta presenta ridondanza, il che significa che le unità di OpenSearch calcolo di indicizzazione e ricerca (OCUs) hanno ciascuna le proprie repliche in standby in una zona di disponibilità diversa. Per scopi di sviluppo e test, puoi scegliere di disabilitare la ridondanza, in modo da ridurre a due il numero di file presenti nella raccolta. OCUs Per ulteriori informazioni, consulta the section called “Come funziona”.
7. Per la sicurezza, scegli Standard create.
8. Per la crittografia, scegli una AWS KMS chiave con cui crittografare i tuoi dati. OpenSearch Serverless ti avvisa se il nome della raccolta che hai inserito corrisponde a uno schema definito in una politica di crittografia. Puoi scegliere di mantenere tale corrispondenza o di sostituirla con impostazioni di crittografia specifiche. Per ulteriori informazioni, consulta the section called “Encryption (Crittografia)”.
9. Per le impostazioni di accesso alla rete, configura l'accesso alla rete per la raccolta.

   - Per Tipo di accesso, seleziona pubblico o privato.

     Se scegli privato, specifica quali endpoint VPC Servizi AWS possono accedere alla raccolta.

    - Endpoint VPC per l'accesso: specifica uno o più endpoint VPC per consentire l'accesso. Per creare un endpoint VPC, consulta la sezione the section called “Endpoint VPC del piano dati”.
    - Servizio AWS accesso privato: seleziona uno o più servizi supportati a cui consentire l'accesso.
- Per Tipo di risorsa, seleziona se gli utenti possono accedere alla raccolta tramite il relativo OpenSearchendpoint (per effettuare chiamate API tramite CURL, Postman e così via), tramite OpenSearch l'endpoint Dashboards (per lavorare con le visualizzazioni ed effettuare chiamate API tramite la console) o entrambi.

> **Note**
>
> Servizio AWS l'accesso privato si applica solo all'endpoint, non all' OpenSearchendpoint Dashboards. OpenSearch

OpenSearch Serverless ti avvisa se il nome della raccolta che hai inserito corrisponde a uno schema definito in una politica di rete. Puoi scegliere di mantenere tale corrispondenza o di sostituirla con impostazioni di rete personalizzate. Per ulteriori informazioni, consulta the section called “Accesso alla rete”.

10. (Facoltativo) Aggiungi uno o più tag alla raccolta. Per ulteriori informazioni, consulta the section called “Assegnazione di tag alle raccolte”.
11. Scegli Next (Successivo).

Configura campi di ricerca aggiuntivi

Le opzioni visualizzate a pagina due del flusso di lavoro per la creazione di una raccolta dipendono dal tipo di raccolta che state creando. Questa sezione descrive come configurare campi di ricerca aggiuntivi per ogni tipo di raccolta. Questa sezione descrive anche come configurare l'arricchimento semantico automatico. Salta qualsiasi sezione che non si applica al tuo tipo di collezione.

Argomenti

- Configura l'arricchimento semantico automatico
- Configura i campi di ricerca delle serie temporali
- Configura i campi di ricerca lessicali

- Configura i campi di ricerca vettoriale

Configura l'arricchimento semantico automatico

Quando crei o modifichi una raccolta, puoi configurare l'arricchimento semantico automatico, che semplifica l'implementazione e le funzionalità della ricerca semantica in Amazon Service. OpenSearch La ricerca semantica restituisce risultati di query che incorporano non solo la corrispondenza delle parole chiave, ma anche l'intento e il significato contestuale della ricerca dell'utente. Per ulteriori informazioni, consulta Arricchimento semantico automatico per Serverless.

Per configurare l'arricchimento semantico automatico

1. Nella sezione Dettagli dell'indice, per Nome dell'indice, specifica un nome.
2. Nella sezione Campi di arricchimento semantico automatico, scegli Aggiungi campo di ricerca semantico.
3. Nel campo Nome del campo di input per l'arricchimento semantico, inserisci il nome di un campo che desideri arricchire.
4. Il tipo di dati è Testo. Non puoi modificare questo valore.
5. Per Lingua, scegli Inglese o Multilingue.
6. Scegli Aggiungi campo.
7. Dopo aver completato la configurazione dei campi opzionali per la raccolta, scegli Avanti. Controlla le modifiche e scegli Invia per creare la raccolta.

Configura i campi di ricerca delle serie temporali

Le opzioni nella sezione Campi di ricerca delle serie temporali riguardano i dati delle serie temporali e i flussi di dati. Per ulteriori informazioni su questi argomenti, vedere. Gestione dei dati di serie temporali in Amazon OpenSearch Service con flussi di dati

Per configurare i campi di ricerca delle serie temporali

1. Nella sezione Campi di ricerca delle serie temporali, scegli Aggiungi campo delle serie temporali.
2. Per Nome campo, inserisci un nome.
3. Per Tipo di dati, scegli un tipo dall'elenco.
4. Scegli Aggiungi campo

5. Dopo aver completato la configurazione dei campi opzionali per la tua raccolta, scegli Avanti. Controlla le modifiche e scegli Invia per creare la raccolta.

Configura i campi di ricerca lessicali

La ricerca lessicale cerca una corrispondenza esatta tra una query di ricerca e termini o parole chiave indicizzati.

Per configurare i campi di ricerca lessicali

1. Nella sezione Campi di ricerca lessicali, scegli Aggiungi campo di ricerca.
2. Per Nome campo, inserisci un nome.
3. Per Tipo di dati, scegli un tipo dall'elenco.
4. Scegli Aggiungi campo
5. Dopo aver completato la configurazione dei campi opzionali per la tua raccolta, scegli Avanti. Controlla le modifiche e scegli Invia per creare la raccolta.

Configura i campi di ricerca vettoriale

Per configurare i campi di ricerca vettoriali

1. Nella sezione Campi vettoriali, scegli Aggiungi campo vettoriale.
2. Per Nome campo, inserisci un nome.
3. Per Engine, scegli un tipo dall'elenco.
4. Immettete il numero di dimensioni.
5. Per Distance Metric, scegliete un tipo dall'elenco.
6. Dopo aver completato la configurazione dei campi opzionali per la raccolta, scegli Avanti.
7. Controlla le modifiche e scegli Invia per creare la raccolta.

Creazione di una raccolta (CLI)

Utilizza le procedure in questa sezione per creare una raccolta OpenSearch Serverless utilizzando. AWS CLI

Argomenti

- Prima di iniziare

- Creazione di una raccolta
- Creazione di una raccolta con un indice di arricchimento semantico automatico

Prima di iniziare

Prima di creare una raccolta utilizzando la AWS CLI, utilizzate la procedura seguente per creare le politiche richieste per la raccolta.

> **Note**
>
> In ciascuna delle seguenti procedure, quando si specifica un nome per una raccolta, il nome deve soddisfare i seguenti criteri:
>
> - È unico per il tuo account e Regione AWS
> - Contiene solo lettere minuscole a-z, i numeri da 0 a 9 e i trattini (-)
> - Contiene da 3 a 32 caratteri

Per creare le politiche necessarie per una raccolta

1. Aprite AWS CLI ed eseguite il comando seguente per creare una politica di crittografia con un modello di risorse che corrisponda al nome previsto della raccolta.

```
&aws opensearchserverless create-security-policy \
  --name policy name \
  --type encryption --policy "{\"Rules\":[{\"ResourceType\":\"collection\",
\"Resource\":[\"collection\/collection name\"]}],\"AWSOwnedKey\":true}"
```

   Ad esempio, se intendi dare un nome all'applicazione dei log della raccolta, potresti creare una policy di crittografia come questa:

```
&aws opensearchserverless create-security-policy \
  --name logs-policy \
  --type encryption --policy "{\"Rules\":[{\"ResourceType\":\"collection\",
\"Resource\":[\"collection\/logs-application\"]}],\"AWSOwnedKey\":true}"
```

   Se prevedi di utilizzare la policy per raccolte aggiuntive, puoi rendere la regola più ampia, come `collection/logs*` o `collection/*`.

2. Eseguite il comando seguente per configurare le impostazioni di rete per la raccolta utilizzando una politica di rete. È possibile creare policy di rete dopo aver creato una raccolta, tuttavia consigliamo di farlo prima.

```
&aws opensearchserverless create-security-policy \
  --name policy name \
  --type network --policy "[{\"Description\":\"description\",\"Rules\":
[{\"ResourceType\":\"dashboard\",\"Resource\":[\"collection\/collection name\"]},
{\"ResourceType\":\"collection\",\"Resource\":[\"collection\/collection name\"]}],
\"AllowFromPublic\":true}]"
```

Utilizzando l'esempio precedente applicazione-log, è possibile creare la seguente policy di rete:

```
&aws opensearchserverless create-security-policy \
  --name logs-policy \
  --type network --policy "[{\"Description\":\"Public access for logs collection
\",\"Rules\":[{\"ResourceType\":\"dashboard\",\"Resource\":[\"collection\/logs-
application\"]},{\"ResourceType\":\"collection\",\"Resource\":[\"collection\/logs-
application\"]}],\"AllowFromPublic\":true}]"
```

Creazione di una raccolta

La procedura seguente utilizza l'azione CreateCollectionAPI per creare una raccolta di tipo `SEARCH` o`TIMESERIES`. Se non specifichi un tipo di raccolta nella richiesta, il valore predefinito sarà `TIMESERIES`. Per ulteriori informazioni su questi tipi, vedereScelta di un tipo di raccolta. Per creare una raccolta di ricerca vettoriale, vederethe section called “Lavorare con le raccolte di ricerca vettoriale”.

Se la raccolta è crittografata con un Chiave di proprietà di AWS, `kmsKeyArn` è `auto` anziché un ARN.

> **Important**
>
> Dopo aver creato una raccolta, non potrai accedervi a meno che non corrisponda a una policy di accesso ai dati. Per ulteriori informazioni, consulta Controllo dell'accesso ai dati per Amazon OpenSearch Serverless.

Come creare una raccolta

1. Verifica di aver creato le politiche richieste descritte inPrima di iniziare.
2. Eseguire il seguente comando seguente. Per `type` specificare `SEARCH` o`TIMESERIES`.

```
&aws opensearchserverless create-collection --name "collection name" --
type collection type --description "description"
```

Creazione di una raccolta con un indice di arricchimento semantico automatico

Utilizzare la procedura seguente per creare una nuova raccolta OpenSearch Serverless con un indice configurato per l'arricchimento semantico automatico. La procedura utilizza l'azione OpenSearch Serverless CreateIndexAPI.

Per creare una nuova raccolta con un indice configurato per l'arricchimento semantico automatico

Eseguite il comando seguente per creare la raccolta e un indice.

```
&aws opensearchserverless create-index \
--region Region ID \
--id collection name --index-name index name \
--index-schema \
'mapping in json'
```

Ecco un esempio:

```
&aws opensearchserverless create-index \
--region us-east-1 \
--id conversation_history --index-name conversation_history_index \
--index-schema \
'{
    "mappings": {
        "properties": {
            "age": {
                "type": "integer"
            },
            "name": {
                "type": "keyword"
            },
            "user_description": {
                "type": "text"
```

```
            },
            "conversation_history": {
                "type": "text",
                "semantic_enrichment": {
                    "status": "ENABLED",
                    // Specifies the sparse tokenizer for processing multi-lingual text
                    "language_option": "MULTI-LINGUAL",
                    // If embedding_field is provided, the semantic embedding field
 will be set to the given name rather than original field name + "_embedding"
                    "embedding_field": "conversation_history_user_defined"
                }
            },
            "book_title": {
                "type": "text",
                "semantic_enrichment": {
                    // No embedding_field is provided, so the semantic embedding field
 is set to "book_title_embedding"
                    "status": "ENABLED",
                    "language_option": "ENGLISH"
                }
            },
            "abstract": {
                "type": "text",
                "semantic_enrichment": {
                    // If no language_option is provided, it will be set to English.
                    // No embedding_field is provided, so the semantic embedding field
 is set to "abstract_embedding"
                    "status": "ENABLED"
                }
            }
        }
    }
}'
```

Accesso ai OpenSearch pannelli di controllo

Dopo aver creato una raccolta con Console di gestione AWS, puoi accedere all'URL delle OpenSearch dashboard della raccolta. Puoi trovare l'URL delle dashboard scegliendo Raccolte nel riquadro di navigazione a sinistra e selezionando la raccolta per aprirne la pagina dei dettagli. L'URL assume il formato `https://dashboards.`*`us-east-1`*`.aoss.amazonaws.com/_login/?collectionId=`*`07tjusf2h91cunochc`*. Una volta che accedi all'URL, accederai automaticamente alle dashboard.

Se hai già l'URL delle OpenSearch dashboard disponibile ma non lo sei Console di gestione AWS, chiamando l'URL delle dashboard dal browser verrai reindirizzato alla console. Una volta inserite le AWS credenziali, accederai automaticamente a Dashboards. Per informazioni sull'accesso alle raccolte per SAML, consulta Accesso ai OpenSearch dashboard con SAML.

Il timeout della console OpenSearch Dashboards è di un'ora e non è configurabile.

> **Note**
>
> Il 10 maggio 2023, OpenSearch ha introdotto un endpoint globale comune per Dashboards. OpenSearch Ora puoi accedere alle OpenSearch dashboard nel browser con un URL che assume il formato. `https://dashboards.us-east-1.aoss.amazonaws.com/_login/?collectionId=07tjusf2h91cunochc` Per garantire la compatibilità con le versioni precedenti, continueremo a supportare gli endpoint OpenSearch Dashboards specifici della raccolta esistenti con il formato. `https://07tjusf2h91cunochc.us-east-1.aoss.amazonaws.com/_dashboards`

Visualizzazione delle raccolte

Puoi visualizzare le raccolte esistenti nella tua Account AWS scheda Raccolte della console di Amazon OpenSearch Service.

Per elencare le raccolte insieme alle relative IDs, invia una ListCollectionsrichiesta.

```
&aws opensearchserverless list-collections
```

Risposta di esempio

```
{
   "collectionSummaries":[
      {
         "arn":"arn:aws:aoss:us-east-1:123456789012:collection/07tjusf2h91cunochc",
         "id":"07tjusf2h91cunochc",
         "name":"my-collection",
         "status":"CREATING"
      }
   ]
}
```

Per limitare i risultati della ricerca, utilizza i filtri della raccolta. Questa richiesta filtra la risposta alle raccolte nello stato `ACTIVE`:

```
&aws opensearchserverless list-collections --collection-filters '{ "status":
 "ACTIVE" }'
```

Per ottenere informazioni più dettagliate su una o più raccolte, inclusi l' OpenSearch endpoint e l'endpoint OpenSearch Dashboards, invia una richiesta: BatchGetCollection

```
&aws opensearchserverless batch-get-collection --ids "07tjusf2h91cunochc"
 "1iu5usc4rame"
```

> **Note**
>
> Nella richiesta puoi includere `--names` o `--ids`, ma non entrambi.

Risposta di esempio


```
{
   "collectionDetails":[
      {
         "id": "07tjusf2h91cunochc",
         "name": "my-collection",
         "status": "ACTIVE",
         "type": "SEARCH",
         "description": "",
         "arn": "arn:aws:aoss:us-east-1:123456789012:collection/07tjusf2h91cunochc",
         "kmsKeyArn": "arn:aws:kms:us-
east-1:123456789012:key/1234abcd-12ab-34cd-56ef-1234567890ab",
         "createdDate": 1667446262828,
         "lastModifiedDate": 1667446300769,
         "collectionEndpoint": "https://07tjusf2h91cunochc.us-
east-1.aoss.amazonaws.com",
         "dashboardEndpoint": "https://07tjusf2h91cunochc.us-east-1.aoss.amazonaws.com/
_dashboards"
      },
      {
         "id": "178ukvtg3i82dvopdid",
         "name": "another-collection",
         "status": "ACTIVE",
```

```
         "type": "TIMESERIES",
         "description": "",
         "arn": "arn:aws:aoss:us-east-1:123456789012:collection/178ukvtg3i82dvopdid",
         "kmsKeyArn": "arn:aws:kms:us-
east-1:123456789012:key/1234abcd-12ab-34cd-56ef-1234567890ab",
         "createdDate": 1667446262828,
         "lastModifiedDate": 1667446300769,
         "collectionEndpoint": "https://178ukvtg3i82dvopdid.us-
east-1.aoss.amazonaws.com",
         "dashboardEndpoint": "https://178ukvtg3i82dvopdid.us-
east-1.aoss.amazonaws.com/_dashboards"
      }
   ],
   "collectionErrorDetails":[]
}
```

Eliminazione di raccolte

Eliminando una raccolta vengono eliminati anche tutti i relativi dati e gli indici. Non è possibile recuperare le raccolte dopo averle eliminate.

Per eliminare una raccolta tramite la console

1. Dal pannello Raccolte della console Amazon OpenSearch Service, seleziona la raccolta che desideri eliminare.
2. Scegli Elimina e conferma l'eliminazione.

Per eliminare una raccolta utilizzando il AWS CLI, invia una DeleteCollectionrichiesta:

```
&aws opensearchserverless delete-collection --id 07tjusf2h91cunochc
```

Risposta di esempio

```
{
   "deleteCollectionDetail":{
      "id":"07tjusf2h91cunochc",
      "name":"my-collection",
      "status":"DELETING"
   }
}
```

Lavorare con le raccolte di ricerca vettoriale

Il tipo di raccolta di ricerca vettoriale in OpenSearch Serverless offre una funzionalità di ricerca per similarità scalabile e ad alte prestazioni. Semplifica la creazione di moderne esperienze di ricerca aumentata di machine learning (ML) e applicazioni generative di intelligenza artificiale (AI) senza dover gestire l'infrastruttura di database vettoriale sottostante.

I casi d'uso per le raccolte di ricerche vettoriali includono la ricerca di immagini, la ricerca di documenti, il recupero di musica, i consigli sui prodotti, le ricerche video, le ricerche basate sulla posizione, il rilevamento di frodi e il rilevamento di anomalie.

Poiché il motore vettoriale per OpenSearch Serverless è alimentato dalla funzione di ricerca k-Nearest Neighbor (k-NN) in, ottieni le stesse funzionalità con la semplicità di un ambiente serverless. OpenSearch Il motore supporta l'API del plug-in k-NN. Con queste operazioni, puoi sfruttare la ricerca full-text, il filtraggio avanzato, le aggregazioni, le query geospaziali, le query annidate per un recupero più rapido dei dati e risultati di ricerca migliorati.

Il motore vettoriale fornisce metriche di distanza come la distanza euclidea, la somiglianza del coseno e la somiglianza del prodotto scalare e può contenere 16.000 dimensioni. È possibile memorizzare campi con vari tipi di dati per i metadati, ad esempio numeri, valori booleani, date, parole chiave e punti geografici. Puoi anche memorizzare campi con testo per informazioni descrittive per aggiungere più contesto ai vettori memorizzati. La colocazione dei tipi di dati riduce la complessità, aumenta la manutenibilità ed evita la duplicazione dei dati, problemi di compatibilità delle versioni e problemi di licenza.

> **Note**
>
> Prendi nota delle seguenti informazioni:
>
> - Amazon OpenSearch Serverless supporta la quantizzazione scalare Faiss a 16 bit, che può essere utilizzata per eseguire conversioni tra vettori mobili a 32 bit e vettori a 16 bit. Per ulteriori informazioni, consulta Quantizzazione scalare a 16 bit di Faiss. È inoltre possibile utilizzare vettori binari per ridurre i costi di memoria. Per ulteriori informazioni, consulta Vettori binari.
> - Amazon OpenSearch Serverless supporta la ricerca vettoriale basata su disco. La ricerca vettoriale basata su disco riduce in modo significativo i costi operativi per i carichi di lavoro vettoriali in ambienti con poca memoria. Per ulteriori informazioni, consulta Ricerca vettoriale basata su disco.

Guida introduttiva alle raccolte di ricerca vettoriale

In questo tutorial, completerai i seguenti passaggi per archiviare, cercare e recuperare gli incorporamenti vettoriali in tempo reale:

1. Configurazione delle autorizzazioni
2. Creazione di una raccolta
3. Caricamento e ricerca dei dati
4. Eliminazione della raccolta

Fase 1: configurazione delle autorizzazioni

Per completare questo tutorial (e per utilizzare OpenSearch Serverless in generale), devi disporre delle autorizzazioni AWS Identity and Access Management (IAM) corrette. In questo tutorial, crei una raccolta, carichi e cerchi dati, quindi elimini la raccolta.

L'utente o il ruolo devono avere una policy basata sull'identità allegata con le seguenti autorizzazioni minime:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Action": [
        "aoss:CreateCollection",
        "aoss:ListCollections",
        "aoss:BatchGetCollection",
        "aoss:DeleteCollection",
        "aoss:CreateAccessPolicy",
        "aoss:ListAccessPolicies",
        "aoss:UpdateAccessPolicy",
        "aoss:CreateSecurityPolicy",
        "iam:ListUsers",
        "iam:ListRoles"
      ],
      "Effect": "Allow",
      "Resource": "*"
```

```
    }
  ]
}
```

Per ulteriori informazioni sulle autorizzazioni IAM OpenSearch Serverless, consulta. the section called “Identity and Access Management”

Fase 2: creazione di una raccolta

Una raccolta è un gruppo di OpenSearch indici che interagiscono per supportare un carico di lavoro o un caso d'uso specifici.

Per creare una raccolta Serverless OpenSearch

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Scegli Collections (Raccolte) nel pannello di navigazione a sinistra e scegli Create collection (Crea raccolta).
3. Assegna un nome all'alloggiamento della raccolta.
4. Per il tipo di raccolta, scegli Ricerca vettoriale. Per ulteriori informazioni, consulta the section called “Scelta di un tipo di raccolta”.
5. In Tipo di distribuzione, deseleziona Abilita ridondanza (repliche attive). In questo modo viene creata una raccolta in modalità sviluppo o test e il numero di unità di OpenSearch calcolo (OCUs) nella raccolta viene ridotto a due. Se vuoi creare un ambiente di produzione in questo tutorial, lascia selezionata la casella di controllo.
6. In Sicurezza, seleziona Easy create per semplificare la configurazione di sicurezza. Per impostazione predefinita, tutti i dati nel motore vettoriale sono crittografati in transito e inattivi. Il motore vettoriale supporta autorizzazioni IAM granulari in modo da poter definire chi può creare, aggiornare ed eliminare crittografie, reti, raccolte e indici.
7. Scegli Next (Successivo).
8. Controlla le impostazioni della raccolta e scegli Submit (Invia). Attendi alcuni minuti affinché lo stato della raccolta diventi `Active`.

Fase 3: Caricamento e ricerca dei dati

Un indice è una raccolta di documenti con uno schema di dati comune che consente di archiviare, cercare e recuperare gli incorporamenti vettoriali e altri campi. Puoi creare e caricare dati negli indici

di una raccolta OpenSearch Serverless utilizzando la console Dev Tools in OpenSearch Dashboards o uno strumento HTTP come Postman o awscurl. Questo tutorial utilizza Dev Tools.

Per indicizzare e cercare i dati nella collezione di alloggi

1. Per creare un singolo indice per la tua nuova collezione, invia la seguente richiesta nella console Dev Tools. Per impostazione predefinita, questo crea un indice con un `nmslib` motore e una distanza euclidea.

```
PUT housing-index
{
   "settings": {
      "index.knn": true
   },
   "mappings": {
      "properties": {
         "housing-vector": {
            "type": "knn_vector",
            "dimension": 3
         },
         "title": {
            "type": "text"
         },
         "price": {
            "type": "long"
         },
         "location": {
            "type": "geo_point"
         }
      }
   }
}
```

2. Per indicizzare un singolo documento in housing-index, invia la seguente richiesta:

```
POST housing-index/_doc
{
  "housing-vector": [
    10,
    20,
    30
  ],
  "title": "2 bedroom in downtown Seattle",
```

```
  "price": "2800",
  "location": "47.71, 122.00"
}
```

3. Per cercare proprietà simili a quelle del tuo indice, invia la seguente query:

```
GET housing-index/_search
{
    "size": 5,
    "query": {
        "knn": {
            "housing-vector": {
                "vector": [
                    10,
                    20,
                    30
                ],
                "k": 5
            }
        }
    }
}
```

Fase 4: Eliminazione della raccolta

Poiché la collezione di alloggi è a scopo di test, assicurati di eliminarla quando hai finito di sperimentare.

Per eliminare una raccolta OpenSearch Serverless

1. Torna alla console di Amazon OpenSearch Service.
2. Scegli Raccolte nel riquadro di navigazione a sinistra e seleziona la raccolta delle proprietà.
3. Scegli Elimina e conferma l'eliminazione.

Ricerca filtrata

Puoi utilizzare i filtri per affinare i risultati della ricerca semantica. Per creare un indice ed eseguire una ricerca filtrata sui tuoi documenti, sostituisci Carica e cerca dati nel tutorial precedente con le seguenti istruzioni. Gli altri passaggi rimangono invariati. Per ulteriori informazioni sui filtri, consulta K-nn search with filters.

Per indicizzare e cercare i dati nella collezione di alloggi

1. Per creare un indice singolo per la tua raccolta, invia la seguente richiesta nella console Dev Tools:

```
PUT housing-index-filtered
{
  "settings": {
    "index.knn": true
  },
  "mappings": {
    "properties": {
      "housing-vector": {
        "type": "knn_vector",
        "dimension": 3,
        "method": {
          "engine": "faiss",
          "name": "hnsw"
        }
      },
      "title": {
        "type": "text"
      },
      "price": {
        "type": "long"
      },
      "location": {
        "type": "geo_point"
      }
    }
  }
}
```

2. Per indicizzare un singolo documento housing-index-filtered, invia la seguente richiesta:

```
POST housing-index-filtered/_doc
{
  "housing-vector": [
    10,
    20,
    30
  ],
  "title": "2 bedroom in downtown Seattle",
```

```
  "price": "2800",
  "location": "47.71, 122.00"
}
```

3. Per cercare i tuoi dati relativi a un appartamento a Seattle a un determinato prezzo ed entro una determinata distanza da un punto geografico, invia la seguente richiesta:

```
GET housing-index-filtered/_search
{
  "size": 5,
  "query": {
    "knn": {
      "housing-vector": {
        "vector": [
          0.1,
          0.2,
          0.3
        ],
        "k": 5,
        "filter": {
          "bool": {
            "must": [
              {
                "query_string": {
                  "query": "Find me 2 bedroom apartment in Seattle under $3000 ",
                  "fields": [
                    "title"
                  ]
                }
              },
              {
                "range": {
                  "price": {
                    "lte": 3000
                  }
                }
              },
              {
                "geo_distance": {
                  "distance": "100miles",
                  "location": {
                    "lat": 48,
                    "lon": 121
```

```
                    }
                  }
                }
              ]
            }
          }
        }
      }
    }
  }
}
```

Carichi di lavoro su scala miliardaria

Le raccolte di ricerca vettoriale supportano carichi di lavoro con miliardi di vettori. Non è necessario reindicizzare per scopi di ridimensionamento perché il ridimensionamento automatico lo fa per te. Se hai milioni di vettori (o più) con un numero elevato di dimensioni e ne hai bisogno di più di 200 OCUs, contatta l'AWS assistenza per aumentare il numero massimo di unità di OpenSearch calcolo (OCUs) per il tuo account.

Limitazioni

Le raccolte di ricerca vettoriale presentano le seguenti limitazioni:

- Le raccolte di ricerca vettoriale non supportano il motore Apache Lucene ANN.
- Le raccolte di ricerca vettoriale supportano solo l'algoritmo HNSW con Faiss e non supportano IVF e IVFQ.
- Le raccolte di ricerca vettoriale non supportano le operazioni API di warmup, stats e model training.
- Le raccolte di ricerca vettoriale non supportano script in linea o memorizzati.
- Le informazioni sul conteggio degli indici non sono disponibili nelle raccolte Console di gestione AWS per la ricerca vettoriale.
- L'intervallo di aggiornamento per gli indici nelle raccolte di ricerca vettoriale è di 60 secondi.

Fasi successive

Ora che sapete come creare una raccolta di ricerca vettoriale e indicizzare i dati, potreste provare alcuni dei seguenti esercizi:

- Usa il client OpenSearch Python per lavorare con raccolte di ricerca vettoriale. Guarda questo tutorial su. GitHub
- Usa il client OpenSearch Java per lavorare con raccolte di ricerca vettoriale. Guarda questo tutorial su. GitHub
- Configurato LangChain per essere utilizzato OpenSearch come archivio vettoriale. LangChain è un framework open source per lo sviluppo di applicazioni basate su modelli linguistici. Per ulteriori informazioni, consulta la documentazione relativa ad LangChain .

Utilizzo delle politiche del ciclo di vita dei dati con Amazon Serverless OpenSearch

Una policy sul ciclo di vita dei dati in Amazon OpenSearch Serverless definisce per quanto tempo OpenSearch Serverless conserva i dati in una raccolta di serie temporali. Ad esempio, puoi impostare una politica per conservare i dati di registro per 30 giorni prima che Serverless li elimini. OpenSearch

Puoi configurare una politica separata per ogni indice all'interno di ogni raccolta di serie temporali presente nel tuo. Account AWS OpenSearch Serverless conserva i documenti almeno per la durata specificata nella policy. Quindi elimina automaticamente i documenti nel miglior modo possibile, in genere entro 48 ore o il 10% del periodo di conservazione, a seconda di quale sia il più lungo.

Solo le raccolte di serie temporali supportano le politiche relative al ciclo di vita dei dati. Le raccolte di ricerca e di ricerca vettoriale non lo fanno.

Argomenti

- Politiche relative al ciclo di vita dei dati
- Autorizzazioni richieste
- Priorità delle policy
- Sintassi delle policy
- Creazione di politiche relative al ciclo di vita dei dati
- Aggiornamento delle politiche relative al ciclo di vita dei dati
- Eliminazione delle politiche relative al ciclo di vita dei dati

Politiche relative al ciclo di vita dei dati

In una politica del ciclo di vita dei dati, si specifica una serie di regole. La politica del ciclo di vita dei dati consente di gestire il periodo di conservazione dei dati associati agli indici o alle raccolte che soddisfano queste regole. Queste regole definiscono il periodo di conservazione dei dati in un indice o in un gruppo di indici. Ogni regola è composta da un tipo di risorsa (`index`), un periodo di conservazione e un elenco di risorse (indici) a cui si applica il periodo di conservazione.

Il periodo di conservazione viene definito con uno dei seguenti formati:

- `"MinIndexRetention": "24h"`— OpenSearch Serverless conserva i dati dell'indice per il periodo specificato in ore o giorni. È possibile impostare questo periodo in modo che sia compreso tra`24h`. `3650d`
- `"NoMinIndexRetention": true`— OpenSearch Serverless conserva i dati dell'indice a tempo indeterminato.

Nella seguente politica di esempio, la prima regola specifica un periodo di conservazione di 15 giorni per tutti gli indici della raccolta. `marketing` La seconda regola specifica che per tutti i nomi di indice che iniziano con `log` la `finance` raccolta non è impostato alcun periodo di conservazione e verranno conservati a tempo indeterminato.

```
{
   "lifeCyclePolicyDetail": {
      "type": "retention",
      "name": "my-policy",
      "policyVersion": "MTY4ODI0NTM2OTk1N18x",
      "policy": {
         "Rules": [
            {
            "ResourceType":"index",
            "Resource":[
               "index/marketing/*"
            ],
            "MinIndexRetention": "15d"
         },
         {
            "ResourceType":"index",
            "Resource":[
               "index/finance/log*"
            ],
```

```
            "NoMinIndexRetention": true
         }
         ]
      },
      "createdDate": 1688245369957,
      "lastModifiedDate": 1688245369957
   }
}
```

Nel seguente esempio di regola politica, OpenSearch Serverless conserva a tempo indeterminato i dati in tutti gli indici per tutte le raccolte all'interno dell'account.

```
{
   "Rules": [
      {
         "ResourceType": "index",
         "Resource": [
            "index/*/*"
         ]
      }
   ],
   "NoMinIndexRetention": true
}
```

Autorizzazioni richieste

Le policy del ciclo di vita per OpenSearch Serverless utilizzano le seguenti autorizzazioni (IAM). AWS Identity and Access Management Puoi specificare le condizioni IAM per limitare gli utenti alle politiche del ciclo di vita dei dati associate a raccolte e indici specifici.

- `aoss:CreateLifecyclePolicy`— Creare una policy sul ciclo di vita dei dati.
- `aoss:ListLifecyclePolicies`— Elenca tutte le politiche relative al ciclo di vita dei dati nell'account corrente.
- `aoss:BatchGetLifecyclePolicy`— Visualizza una politica sul ciclo di vita dei dati associata a un account o al nome di una policy.
- `aoss:BatchGetEffectiveLifecyclePolicy`— Visualizza una politica del ciclo di vita dei dati per una determinata risorsa (`index`è l'unica risorsa supportata).
- `aoss:UpdateLifecyclePolicy`— Modificare una determinata politica del ciclo di vita dei dati e modificarne l'impostazione o la risorsa di conservazione.

- `aoss:DeleteLifecyclePolicy`— Eliminare una politica sul ciclo di vita dei dati.

La seguente politica di accesso basata sull'identità consente a un utente di visualizzare tutte le politiche relative al ciclo di vita dei dati e di aggiornare le politiche in base al modello di risorse: `index/application-logs`

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "aoss:UpdateLifecyclePolicy"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aoss:collection": "application-logs"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "aoss:ListLifecyclePolicies",
                "aoss:BatchGetLifecyclePolicy"
            ],
            "Resource": "*"
        }
    ]
}
```

Priorità delle policy

Possono verificarsi situazioni in cui le regole relative al ciclo di vita dei dati si sovrappongono, all'interno o tra le policy. Quando ciò accade, una regola con un nome o uno schema di risorsa più specifico per un indice sostituisce una regola con un nome o uno schema di risorsa più generale per tutti gli indici comuni a entrambe le regole.

Ad esempio, nella seguente politica, due regole si applicano a un indice. `index/sales/logstash` In questa situazione, la seconda regola ha la precedenza perché `index/sales/log*` è la corrispondenza più lunga a. `index/sales/logstash` Pertanto, OpenSearch Serverless non imposta alcun periodo di conservazione per l'indice.

```
{
      "Rules":[
         {
            "ResourceType":"index",
            "Resource":[
               "index/sales/*",
            ],
            "MinIndexRetention": "15d"
         },
         {
            "ResourceType":"index",
            "Resource":[
               "index/sales/log*",
            ],
            "NoMinIndexRetention": true
         }
      ]
   }
```

Sintassi delle policy

Fornisci una o più regole. Queste regole definiscono le impostazioni del ciclo di vita dei dati per gli indici Serverless. OpenSearch

Ogni regola contiene i seguenti elementi. È possibile fornire `MinIndexRetention` o `NoMinIndexRetention` in ciascuna regola, ma non in entrambe.

Elemento	Description
Tipo di risorsa	Il tipo di risorsa a cui si applica la regola. L'unica opzione supportata per le politiche del ciclo di vita dei dati è. `index`
Risorsa	Un elenco di modelli di nomi and/or di risorse. I pattern sono costituiti da un prefisso e da un carattere jolly (*), che consentono l'applica

Elemento	Description
	zione delle autorizzazioni associate a più risorse. Ad esempio, `index/<collection-name|pattern> /<index-name|pattern>` .
MinIndexRetention	Il periodo minimo, in giorni (d) o ore (h), per conservare il documento nell'indice. Il limite inferiore è `24h` e il limite superiore è`3650d`.
NoMinIndexRetention	Se`true`, OpenSearch Serverless conserva i documenti a tempo indeterminato.

Nell'esempio seguente, la prima regola si applica a tutti gli indici inclusi nel `autoparts-inventory` pattern (`index/autoparts-inventory/*`) e richiede che i dati vengano conservati per almeno 20 giorni prima che possa verificarsi qualsiasi azione, come l'eliminazione o l'archiviazione.

La seconda regola si rivolge agli indici che corrispondono al `auto*/gear` pattern (`index/auto*/gear`), impostando un periodo di conservazione minimo di 24 ore.

La terza regola si applica specificamente all'`tires`indice e non prevede un periodo di conservazione minimo, il che significa che i dati in questo indice possono essere eliminati o archiviati immediatamente o in base ad altri criteri. Queste regole aiutano a gestire la conservazione dei dati dell'indice con tempi di conservazione variabili o senza restrizioni di conservazione.

```
{
  "Rules": [
    {
      "ResourceType": "index",
      "Resource": [
        "index/autoparts-inventory/*"
      ],
      "MinIndexRetention": "20d"
    },
    {
      "ResourceType": "index",
      "Resource": [
        "index/auto*/gear"
      ],
```

```
      "MinIndexRetention": "24h"
    },
    {
      "ResourceType": "index",
      "Resource": [
        "index/autoparts-inventory/tires"
      ],
      "NoMinIndexRetention": true
    }
  ]
}
```

Creazione di politiche relative al ciclo di vita dei dati

Per creare una politica del ciclo di vita dei dati, si definiscono regole che gestiscono la conservazione e l'eliminazione dei dati in base a criteri specifici.

Console

Per creare una politica sul ciclo di vita dei dati

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli Politiche del ciclo di vita dei dati.
3. Scegli Crea una politica sul ciclo di vita dei dati.
4. Inserisci un nome descrittivo per la politica.
5. Per Ciclo di vita dei dati, scegli Aggiungi e seleziona le raccolte e gli indici per la politica.

   Inizia scegliendo le raccolte a cui appartengono gli indici. Quindi, scegli l'indice dall'elenco o inserisci uno schema di indice. Per selezionare tutte le raccolte come fonti, inserisci un asterisco (*).
6. Per la conservazione dei dati, puoi scegliere di conservare i dati a tempo indeterminato oppure deselezionare Unlimited (mai eliminare) e specificare un periodo di tempo dopo il quale OpenSearch Serverless elimina automaticamente i dati da Amazon S3.
7. Scegli Salva, quindi Crea.

AWS CLI

Per creare una politica del ciclo di vita dei dati utilizzando il AWS CLI, utilizza il create-lifecycle-policycomando con le seguenti opzioni:

- `--name`— Il nome della politica.
- `--type`— Il tipo di politica. Attualmente, l'unico valore disponibile è`retention`.
- `--policy`— La politica del ciclo di vita dei dati. Questo parametro accetta sia le policy in linea che i file.json. È necessario codificare le politiche in linea come stringa di escape JSON. Per fornire la politica in un file, usa il formato. `--policy file://my-policy.json`

Example

```
aws opensearchserverless create-lifecycle-policy \
  --name my-policy \
  --type retention \
  --policy "{\"Rules\":[{\"ResourceType\":\"index\",\"Resource\":[\"index/autoparts-
inventory/*\"],\"MinIndexRetention\": \"81d\"},{\"ResourceType\":\"index\",\"Resource
\":[\"index/sales/orders*\"],\"NoMinIndexRetention\":true}]}"
```

Aggiornamento delle politiche relative al ciclo di vita dei dati

Per aggiornare una politica sul ciclo di vita dei dati, puoi modificare le regole esistenti per riflettere le modifiche ai requisiti di conservazione o eliminazione dei dati. Ciò consente di adattare le politiche all'evolversi delle esigenze di gestione dei dati.

Potrebbero trascorrere alcuni minuti tra l'aggiornamento della policy e il momento in cui OpenSearch Serverless inizia a far rispettare i nuovi periodi di conservazione.

Console

Per aggiornare una policy sul ciclo di vita dei dati

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli Politiche del ciclo di vita dei dati.
3. Seleziona la politica del ciclo di vita dei dati che desideri aggiornare, quindi scegli Modifica.
4. Modifica la politica utilizzando l'editor visivo o l'editor JSON.
5. Scegli Save (Salva).

AWS CLI

Per aggiornare una politica del ciclo di vita dei dati utilizzando il AWS CLI, usa il comando. update-lifecycle-policy

È necessario includere il `--policy-version` parametro nella richiesta. È possibile recuperare la versione della policy utilizzando i comandi list-lifecycle-policies o batch-get-lifecycle-policy. Ti consigliamo di includere la versione più recente della policy per evitare di sovrascrivere accidentalmente le modifiche apportate da altri.

La seguente richiesta aggiorna una policy sul ciclo di vita dei dati con un nuovo documento JSON di policy.

Example

```
aws opensearchserverless update-lifecycle-policy \
  --name my-policy \
  --type retention \
  --policy-version MTY2MzY5MTY1MDA3Ml8x \
  --policy file://my-new-policy.json
```

Eliminazione delle politiche relative al ciclo di vita dei dati

Quando elimini una policy sul ciclo di vita dei dati, OpenSearch Serverless non la applica più sugli indici corrispondenti.

Console

Per eliminare una policy sul ciclo di vita dei dati

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli Politiche del ciclo di vita dei dati.
3. Seleziona la politica che desideri eliminare, quindi scegli Elimina e conferma l'eliminazione.

AWS CLI

Per eliminare una politica del ciclo di vita dei dati utilizzando il AWS CLI, usa il delete-lifecycle-policycomando.

Example

```
aws opensearchserverless delete-lifecycle-policy \
  --name my-policy \
  --type retention
```

Utilizzo di AWS SDKs per interagire con Amazon OpenSearch Serverless

Questa sezione include esempi di come utilizzarlo AWS SDKs per interagire con Amazon OpenSearch Serverless. Questi esempi di codice mostrano come creare raccolte e policy di sicurezza, e come eseguire query sulle raccolte.

> **Note**
>
> Attualmente stiamo sviluppando questi esempi di codice. Se vuoi contribuire con un esempio di codice (Java, Go, ecc.), apri una pull request direttamente all'interno del GitHub repository.

Argomenti

- Python
- JavaScript

Python

Il seguente script di esempio utilizza il client AWS SDK per Python (Boto3), così come opensearch-py per Python per creare policy di crittografia, rete e accesso ai dati, creare una raccolta corrispondente e indicizzare alcuni dati di esempio.

Per installare le dipendenze richieste, eseguire i seguenti comandi:

```
pip install opensearch-py
pip install boto3
pip install botocore
pip install requests-aws4auth
```

All'interno dello script, sostituire l'elemento `Principal` con il nome della risorsa Amazon (ARN) dell'utente o il ruolo che sta firmando la richiesta. Facoltativamente, è possibile anche modificare la `region`.

```
from opensearchpy import OpenSearch, RequestsHttpConnection
from requests_aws4auth import AWS4Auth
import boto3
import botocore
import time

# Build the client using the default credential configuration.
# You can use the CLI and run 'aws configure' to set access key, secret
# key, and default region.

client = boto3.client('opensearchserverless')
service = 'aoss'
region = 'us-east-1'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key,
                   region, service, session_token=credentials.token)


def createEncryptionPolicy(client):
    """Creates an encryption policy that matches all collections beginning with tv-"""
    try:
        response = client.create_security_policy(
            description='Encryption policy for TV collections',
            name='tv-policy',
            policy="""
                {
                    \"Rules\":[
                        {
                            \"ResourceType\":\"collection\",
                            \"Resource\":[
                                \"collection\/tv-*\"
                            ]
                        }
                    ],
                    \"AWSOwnedKey\":true
                }
                """,
            type='encryption'
        )
        print('\nEncryption policy created:')
        print(response)
    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ConflictException':
```

```
            print(
                '[ConflictException] The policy name or rules conflict with an existing
 policy.')
        else:
            raise error


def createNetworkPolicy(client):
    """Creates a network policy that matches all collections beginning with tv-"""
    try:
        response = client.create_security_policy(
            description='Network policy for TV collections',
            name='tv-policy',
            policy="""
                [{
                    \"Description\":\"Public access for TV collection\",
                    \"Rules\":[
                        {
                            \"ResourceType\":\"dashboard\",
                            \"Resource\":[\"collection\/tv-*\"]
                        },
                        {
                            \"ResourceType\":\"collection\",
                            \"Resource\":[\"collection\/tv-*\"]
                        }
                    ],
                    \"AllowFromPublic\":true
                }]
                """,
            type='network'
        )
        print('\nNetwork policy created:')
        print(response)
    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ConflictException':
            print(
                '[ConflictException] A network policy with this name already exists.')
        else:
            raise error


def createAccessPolicy(client):
    """Creates a data access policy that matches all collections beginning with tv-"""
    try:
```

```
        response = client.create_access_policy(
            description='Data access policy for TV collections',
            name='tv-policy',
            policy="""
                [{
                    \"Rules\":[
                        {
                            \"Resource\":[
                                \"index\/tv-*\/*\"
                            ],
                            \"Permission\":[
                                \"aoss:CreateIndex\",
                                \"aoss:DeleteIndex\",
                                \"aoss:UpdateIndex\",
                                \"aoss:DescribeIndex\",
                                \"aoss:ReadDocument\",
                                \"aoss:WriteDocument\"
                            ],
                            \"ResourceType\": \"index\"
                        },
                        {
                            \"Resource\":[
                                \"collection\/tv-*\"
                            ],
                            \"Permission\":[
                                \"aoss:CreateCollectionItems\"
                            ],
                            \"ResourceType\": \"collection\"
                        }
                    ],
                    \"Principal\":[
                        \"arn:aws:iam::123456789012:role\/Admin\"
                    ]
                }]
                """,
            type='data'
        )
        print('\nAccess policy created:')
        print(response)
    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ConflictException':
            print(
                '[ConflictException] An access policy with this name already exists.')
        else:
```

```
            raise error


def createCollection(client):
    """Creates a collection"""
    try:
        response = client.create_collection(
            name='tv-sitcoms',
            type='SEARCH'
        )
        return(response)
    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ConflictException':
            print(
                '[ConflictException] A collection with this name already exists. Try
 another name.')
        else:
            raise error


def waitForCollectionCreation(client):
    """Waits for the collection to become active"""
    response = client.batch_get_collection(
        names=['tv-sitcoms'])
    # Periodically check collection status
    while (response['collectionDetails'][0]['status']) == 'CREATING':
        print('Creating collection...')
        time.sleep(30)
        response = client.batch_get_collection(
            names=['tv-sitcoms'])
    print('\nCollection successfully created:')
    print(response["collectionDetails"])
    # Extract the collection endpoint from the response
    host = (response['collectionDetails'][0]['collectionEndpoint'])
    final_host = host.replace("https://", "")
    indexData(final_host)


def indexData(host):
    """Create an index and add some sample data"""
    # Build the OpenSearch client
    client = OpenSearch(
        hosts=[{'host': host, 'port': 443}],
        http_auth=awsauth,
```

```
        use_ssl=True,
        verify_certs=True,
        connection_class=RequestsHttpConnection,
        timeout=300
    )
    # It can take up to a minute for data access rules to be enforced
    time.sleep(45)

    # Create index
    response = client.indices.create('sitcoms-eighties')
    print('\nCreating index:')
    print(response)

    # Add a document to the index.
    response = client.index(
        index='sitcoms-eighties',
        body={
            'title': 'Seinfeld',
            'creator': 'Larry David',
            'year': 1989
        },
        id='1',
    )
    print('\nDocument added:')
    print(response)


def main():
    createEncryptionPolicy(client)
    createNetworkPolicy(client)
    createAccessPolicy(client)
    createCollection(client)
    waitForCollectionCreation(client)


if __name__ == "__main__":
    main()
```

JavaScript

Il seguente script di esempio utilizza l'SDK di Node.js, oltre al client opensearch-js per creare politiche di crittografia JavaScript, rete e accesso ai dati, creare una raccolta corrispondente, creare un indice e indicizzare alcuni dati di esempio. JavaScript

Per installare le dipendenze richieste, eseguire i seguenti comandi:

```
npm i aws-sdk
npm i aws4
npm i @opensearch-project/opensearch
```

All'interno dello script, sostituire l'elemento `Principal` con il nome della risorsa Amazon (ARN) dell'utente o il ruolo che sta firmando la richiesta. Facoltativamente, è possibile anche modificare la `region`.

```
var AWS = require('aws-sdk');
var aws4 = require('aws4');
var {
    Client,
    Connection
} = require("@opensearch-project/opensearch");
var {
    OpenSearchServerlessClient,
    CreateSecurityPolicyCommand,
    CreateAccessPolicyCommand,
    CreateCollectionCommand,
    BatchGetCollectionCommand
} = require("@aws-sdk/client-opensearchserverless");
var client = new OpenSearchServerlessClient();

async function execute() {
    await createEncryptionPolicy(client)
    await createNetworkPolicy(client)
    await createAccessPolicy(client)
    await createCollection(client)
    await waitForCollectionCreation(client)
}

async function createEncryptionPolicy(client) {
    // Creates an encryption policy that matches all collections beginning with 'tv-'
    try {
        var command = new CreateSecurityPolicyCommand({
            description: 'Encryption policy for TV collections',
            name: 'tv-policy',
            type: 'encryption',
            policy: " \
        { \
            \"Rules\":[ \
```

```
                { \
                    \"ResourceType\":\"collection\", \
                    \"Resource\":[ \
                        \"collection\/tv-*\" \
                    ] \
                } \
            ], \
            \"AWSOwnedKey\":true \
        }"
        });
        const response = await client.send(command);
        console.log("Encryption policy created:");
        console.log(response['securityPolicyDetail']);
    } catch (error) {
        if (error.name === 'ConflictException') {
            console.log('[ConflictException] The policy name or rules conflict with an
 existing policy.');
        } else
            console.error(error);
    };
}

async function createNetworkPolicy(client) {
    // Creates a network policy that matches all collections beginning with 'tv-'
    try {
        var command = new CreateSecurityPolicyCommand({
            description: 'Network policy for TV collections',
            name: 'tv-policy',
            type: 'network',
            policy: " \
            [{ \
                \"Description\":\"Public access for television collection\", \
                \"Rules\":[ \
                    { \
                        \"ResourceType\":\"dashboard\", \
                        \"Resource\":[\"collection\/tv-*\"] \
                    }, \
                    { \
                        \"ResourceType\":\"collection\", \
                        \"Resource\":[\"collection\/tv-*\"] \
                    } \
                ], \
                \"AllowFromPublic\":true \
            }]"
```

```
        });
        const response = await client.send(command);
        console.log("Network policy created:");
        console.log(response['securityPolicyDetail']);
    } catch (error) {
        if (error.name === 'ConflictException') {
            console.log('[ConflictException] A network policy with that name already
 exists.');
        } else
            console.error(error);
    };
}

async function createAccessPolicy(client) {
    // Creates a data access policy that matches all collections beginning with 'tv-'
    try {
        var command = new CreateAccessPolicyCommand({
            description: 'Data access policy for TV collections',
            name: 'tv-policy',
            type: 'data',
            policy: " \
            [{ \
                \"Rules\":[ \
                    { \
                        \"Resource\":[ \
                            \"index\/tv-*\/*\" \
                        ], \
                        \"Permission\":[ \
                            \"aoss:CreateIndex\", \
                            \"aoss:DeleteIndex\", \
                            \"aoss:UpdateIndex\", \
                            \"aoss:DescribeIndex\", \
                            \"aoss:ReadDocument\", \
                            \"aoss:WriteDocument\" \
                        ], \
                        \"ResourceType\": \"index\" \
                    }, \
                    { \
                        \"Resource\":[ \
                            \"collection\/tv-*\" \
                        ], \
                        \"Permission\":[ \
                            \"aoss:CreateCollectionItems\" \
                        ], \
```

```
                        \"ResourceType\": \"collection\" \
                    } \
                ], \
                \"Principal\":[ \
                    \"arn:aws:iam::123456789012:role\/Admin\" \
                ] \
            }]"
        });
        const response = await client.send(command);
        console.log("Access policy created:");
        console.log(response['accessPolicyDetail']);
    } catch (error) {
        if (error.name === 'ConflictException') {
            console.log('[ConflictException] An access policy with that name already
 exists.');
        } else
            console.error(error);
    };
}

async function createCollection(client) {
    // Creates a collection to hold TV sitcoms indexes
    try {
        var command = new CreateCollectionCommand({
            name: 'tv-sitcoms',
            type: 'SEARCH'
        });
        const response = await client.send(command);
        return (response)
    } catch (error) {
        if (error.name === 'ConflictException') {
            console.log('[ConflictException] A collection with this name already
 exists. Try another name.');
        } else
            console.error(error);
    };
}

async function waitForCollectionCreation(client) {
    // Waits for the collection to become active
    try {
        var command = new BatchGetCollectionCommand({
            names: ['tv-sitcoms']
        });
```

```
        var response = await client.send(command);
        while (response.collectionDetails[0]['status'] == 'CREATING') {
            console.log('Creating collection...')
            await sleep(30000) // Wait for 30 seconds, then check the status again
            function sleep(ms) {
                return new Promise((resolve) => {
                    setTimeout(resolve, ms);
                });
            }
            var response = await client.send(command);
        }
        console.log('Collection successfully created:');
        console.log(response['collectionDetails']);
        // Extract the collection endpoint from the response
        var host = (response.collectionDetails[0]['collectionEndpoint'])
        // Pass collection endpoint to index document request
        indexDocument(host)
    } catch (error) {
        console.error(error);
    };
}

async function indexDocument(host) {

    var client = new Client({
        node: host,
        Connection: class extends Connection {
            buildRequestObject(params) {
                var request = super.buildRequestObject(params)
                request.service = 'aoss';
                request.region = 'us-east-1'; // e.g. us-east-1
                var body = request.body;
                request.body = undefined;
                delete request.headers['content-length'];
                request.headers['x-amz-content-sha256'] = 'UNSIGNED-PAYLOAD';
                request = aws4.sign(request, AWS.config.credentials);
                request.body = body;

                return request
            }
        }
    });

    // Create an index
```

```
    try {
        var index_name = "sitcoms-eighties";

        var response = await client.indices.create({
            index: index_name
        });

        console.log("Creating index:");
        console.log(response.body);

        // Add a document to the index
        var document = "{ \"title\": \"Seinfeld\", \"creator\": \"Larry David\", \"year
\": \"1989\" }\n";

        var response = await client.index({
            index: index_name,
            body: document
        });

        console.log("Adding document:");
        console.log(response.body);
    } catch (error) {
        console.error(error);
    };
}

execute()
```

Utilizzo CloudFormation per creare raccolte Amazon OpenSearch Serverless

Puoi utilizzarlo CloudFormation per creare risorse Amazon OpenSearch Serverless come raccolte, policy di sicurezza ed endpoint VPC. Per un CloudFormation riferimento completo OpenSearch su Serverless, consulta Amazon OpenSearch Serverless nella Guida per l'CloudFormation utente.

Il seguente CloudFormation modello di esempio crea una semplice politica di accesso ai dati, una politica di rete e una politica di sicurezza, oltre a una raccolta corrispondente. È un buon modo per iniziare rapidamente a usare Amazon OpenSearch Serverless e fornire gli elementi necessari per creare e utilizzare una raccolta.

> **Important**
>
> Questo esempio utilizza l'accesso alla rete pubblica, sconsigliato per carichi di lavoro di produzione. È preferibile utilizzare l'accesso VPC per proteggere le raccolte. Per ulteriori informazioni, consultare AWS::OpenSearchServerless::VpcEndpoint e the section called "Endpoint VPC del piano dati".

```
AWSTemplateFormatVersion: 2010-09-09
Description: 'Amazon OpenSearch Serverless template to create an IAM user, encryption
 policy, data access policy and collection'
Resources:
  IAMUSer:
    Type: 'AWS::IAM::User'
    Properties:
      UserName:  aossadmin
  DataAccessPolicy:
    Type: 'AWS::OpenSearchServerless::AccessPolicy'
    Properties:
      Name: quickstart-access-policy
      Type: data
      Description: Access policy for quickstart collection
      Policy: !Sub >-
        [{"Description":"Access for cfn user","Rules":
[{"ResourceType":"index","Resource":["index/*/*"],"Permission":["aoss:*"]},
        {"ResourceType":"collection","Resource":["collection/quickstart"],"Permission":
["aoss:*"]}],
        "Principal":["arn:aws:iam::${AWS::AccountId}:user/aossadmin"]}]
  NetworkPolicy:
    Type: 'AWS::OpenSearchServerless::SecurityPolicy'
    Properties:
      Name: quickstart-network-policy
      Type: network
      Description: Network policy for quickstart collection
      Policy: >-
        [{"Rules":[{"ResourceType":"collection","Resource":["collection/
quickstart"]}, {"ResourceType":"dashboard","Resource":["collection/
quickstart"]}],"AllowFromPublic":true}]
  EncryptionPolicy:
    Type: 'AWS::OpenSearchServerless::SecurityPolicy'
    Properties:
      Name: quickstart-security-policy
```

```
      Type: encryption
      Description: Encryption policy for quickstart collection
      Policy: >-
        {"Rules":[{"ResourceType":"collection","Resource":["collection/
quickstart"]}],"AWSOwnedKey":true}
  Collection:
    Type: 'AWS::OpenSearchServerless::Collection'
    Properties:
      Name: quickstart
      Type: TIMESERIES
      Description: Collection to holds timeseries data
    DependsOn: EncryptionPolicy
Outputs:
  IAMUser:
    Value: !Ref IAMUSer
  DashboardURL:
    Value: !GetAtt Collection.DashboardEndpoint
  CollectionARN:
    Value: !GetAtt Collection.Arn
```

Backup delle raccolte tramite istantanee

Le istantanee sono point-in-time backup delle tue raccolte Amazon OpenSearch Serverless che forniscono funzionalità di disaster recovery. OpenSearch Serverless crea e gestisce automaticamente istantanee delle tue raccolte, garantendo la continuità aziendale e la protezione dei dati. Ogni istantanea contiene metadati dell'indice (impostazioni e mappature per gli indici), metadati del cluster (modelli di indice e alias) e dati dell'indice (tutti i documenti e i dati archiviati negli indici).

OpenSearch Serverless offre backup orari automatici senza configurazione manuale, zero costi di manutenzione, costi di storage aggiuntivi, ripristino rapido in caso di perdita accidentale di dati e la possibilità di ripristinare indici specifici da un'istantanea.

Prima di lavorare con le istantanee, è necessario comprendere queste importanti considerazioni. La creazione di un'istantanea richiede tempo e non è istantanea. I nuovi documenti o gli aggiornamenti durante la creazione dell'istantanea non verranno inclusi nell'istantanea. È possibile ripristinare le istantanee solo nella raccolta originale e non in una nuova. Una volta ripristinati, gli indici ne ricevono di nuovi UUIDs che differiscono dalle versioni originali. Il ripristino di un indice aperto esistente in OpenSearch Serverless sovrascriverà i dati di tale indice, a condizione che non venga fornito un nuovo nome di indice o un modello di prefisso. Ciò differisce dal comportamento principale. OpenSearch È possibile eseguire una sola operazione di ripristino alla volta e non è possibile avviare più operazioni di ripristino contemporaneamente sulla stessa raccolta. Il tentativo di ripristinare

gli indici durante un'operazione di ripristino attiva causa l'esito negativo dell'operazione. Durante un'operazione di ripristino, le richieste agli indici hanno esito negativo.

Autorizzazioni richieste

Per utilizzare le istantanee, configura le seguenti autorizzazioni nella tua politica di accesso ai dati. Per ulteriori informazioni sulle politiche di accesso ai dati, consulta. Policy di accesso ai dati rispetto alle policy IAM

Politica di accesso ai dati	APIs
aoss: DescribeSnapshot	GET /_ -automatizzato cat/snapshots/aoss OTTIENI _/snapshot/aoss-automated/snapshot
aos: RestoreSnapshot	POST /_ /_restore snapshot/aoss-automated/snapshot
aos: DescribeCollectionItems	OTTIENI /_cat/recovery

È possibile configurare le politiche utilizzando i seguenti comandi: AWS CLI

1. create-access-policy
2. delete-access-policy
3. get-access-policy
4. update-access-policy

Ecco un esempio di comando CLI per creare una politica di accesso. Nel comando, sostituisci il *example* contenuto con le tue informazioni specifiche.

```
aws opensearchserverless create-access-policy \
--type data \
--name Example-data-access-policy \
--region aws-region \
--policy '[
  {
    "Rules": [
      {
        "Resource": [
```

```
          "collection/Example-collection"
        ],
        "Permission": [
          "aoss:DescribeSnapshot",
          "aoss:RestoreSnapshot",
          "aoss:DescribeCollectionItems"
        ],
        "ResourceType": "collection"
      }
    ],
    "Principal": [
      "arn:aws:iam::111122223333:user/UserName"
    ],
    "Description": "Data policy to support snapshot operations."
  }
]'
```

Utilizzo degli snapshot

Per impostazione predefinita, quando crei una nuova raccolta, OpenSearch Serverless crea automaticamente istantanee ogni ora. Non è necessario eseguire nessuna operazione. Ogni istantanea include tutti gli indici della raccolta. Dopo che OpenSearch Serverless ha creato le istantanee, è possibile elencarle e rivederne i dettagli utilizzando le seguenti procedure.

Elenca le istantanee

Utilizzate le seguenti procedure per elencare tutte le istantanee di una raccolta e rivederne i dettagli.

Console

1. Apri la console Amazon OpenSearch Service all'indirizzo https://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Serverless, quindi scegli Collezioni.
3. Scegli il nome della tua collezione per aprirne la pagina dei dettagli.
4. Scegli la scheda Istantanee per visualizzare tutte le istantanee generate.
5. Esamina le informazioni sulle istantanee, tra cui:
   - ID istantanea: identificatore univoco per l'istantanea
   - Stato: stato attuale (disponibile, in corso)
   - Ora di creazione: quando è stata scattata l'istantanea

AWS CLI

- Utilizzate il seguente comando per elencare tutte le istantanee di una raccolta:

```
GET /_cat/snapshots/aoss-automated
```

OpenSearch Serverless restituisce una risposta simile alla seguente:

```
id                                  status  start_epoch start_time end_epoch
 end_time    duration    indexes successful_shards failed_shards total_shards
snapshot-ExampleSnapshotID1     SUCCESS 1737964331  07:52:11   1737964382
 07:53:02    50.4s       1
snapshot-ExampleSnapshotID2     SUCCESS 1737967931  08:52:11   1737967979
 08:52:59    47.7s       2
snapshot-ExampleSnapshotID3     SUCCESS 1737971531  09:52:11   1737971581
 09:53:01    49.1s       3
snapshot-ExampleSnapshotID4 IN_PROGRESS 1737975131  10:52:11   -          -
       4.8d       3
```

Ottieni i dettagli dell'istantanea

Utilizza le seguenti procedure per recuperare informazioni dettagliate su un'istantanea specifica.

Console

1. Apri la console Amazon OpenSearch Service all'indirizzo https://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Serverless, quindi scegli Collezioni.
3. Scegli il nome della tua collezione per aprirne la pagina dei dettagli.
4. Selezionare la scheda Snapshots (Snapshot).
5. Scegli l'ID del lavoro dell'istantanea per visualizzare informazioni dettagliate sull'istantanea, inclusi i metadati, gli indici inclusi e le informazioni sulla tempistica.

AWS CLI

- Utilizzate il comando seguente per recuperare informazioni su un'istantanea. Nel comando, sostituisci il *example* contenuto con le tue informazioni specifiche.

```
GET _snapshot/aoss-automated/snapshot/
```

Richiesta di esempio:

```
GET _snapshot/aoss-automated/snapshot-ExampleSnapshotID1/
```

Risposta di esempio:

```
{
    "snapshots": [
        {
            "snapshot": "snapshot-ExampleSnapshotID1-5e01-4423-9833Example",
            "uuid": "Example-5e01-4423-9833-9e9eb757Example",
            "version_id": 136327827,
            "version": "2.11.0",
            "remote_store_index_shallow_copy": true,
            "indexes": [
                "Example-index-0117"
            ],
            "data_streams": [],
            "include_global_state": true,
            "metadata": {},
            "state": "SUCCESS",
            "start_time": "2025-01-27T09:52:11.953Z",
            "start_time_in_millis": 1737971531953,
            "end_time": "2025-01-27T09:53:01.062Z",
            "end_time_in_millis": 1737971581062,
            "duration_in_millis": 49109,
            "failures": [],
            "shards": {
                "total": 0,
                "failed": 0,
                "successful": 0
            }
        }
    ]
}
```

La risposta istantanea include diversi campi chiave: `id` fornisce un identificatore univoco per l'operazione di istantanea, `status` restituisce lo stato corrente `SUCCESS` oppure `IN_PROGRESS`

`duration` indica il tempo impiegato per completare l'operazione di snapshot e `indexes` restituisce il numero di indici inclusi nell'istantanea.

Ripristino da uno snapshot

Il ripristino da un'istantanea consente di recuperare i dati da un backup eseguito in precedenza. Questo processo è fondamentale per il disaster recovery e la gestione dei dati in Serverless. OpenSearch Prima del ripristino, tenete presente che gli indici ripristinati avranno caratteristiche diverse UUIDs rispetto alle versioni originali, il ripristino su un indice aperto esistente in OpenSearch Serverless sovrascriverà i dati di quell'indice a condizione che non venga fornito un nuovo nome di indice o uno schema di prefisso, le istantanee possono essere ripristinate solo nella raccolta originale (il ripristino tra raccolte non è supportato) e le operazioni di ripristino influiranno sulle prestazioni del cluster, quindi pianifica di conseguenza.

Utilizzare le seguenti procedure per ripristinare gli indici di backup da un'istantanea.

Console

1. Apri la console Amazon OpenSearch Service all'indirizzo https://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Serverless, quindi scegli Collezioni.
3. Scegli il nome della tua collezione per aprirne la pagina dei dettagli.
4. Scegli la scheda Istantanee per visualizzare le istantanee disponibili.
5. Scegli l'istantanea da cui desideri ripristinare, quindi scegli Ripristina da istantanea.
6. Nella finestra di dialogo Ripristina da istantanea:
   - Per il nome dell'istantanea, verifica l'ID dell'istantanea selezionato.
   - Per Snapshot scope, scegli una delle seguenti opzioni:
     - Tutti gli indici della raccolta: ripristina tutti gli indici dall'istantanea
     - Indici specifici: seleziona i singoli indici da ripristinare
   - Per Destinazione, scegli la raccolta in cui eseguire il ripristino.
   - (Facoltativo) Configura le impostazioni di ridenominazione per rinominare gli indici ripristinati:
     - Non rinominare: mantieni i nomi degli indici originali
     - Aggiungi prefisso ai nomi degli indici ripristinati: aggiungi un prefisso per evitare conflitti
     - Rinomina utilizzando espressioni regolari: utilizza modelli di ridenominazione avanzati

- (Facoltativo) Configura le impostazioni di notifica per ricevere una notifica quando il ripristino viene completato o riscontra errori.

7. Scegli Salva per avviare l'operazione di ripristino.

OpenSearch API

1. Esegui il comando seguente per identificare l'istantanea appropriata.

```
GET /_snapshot/aoss-automated/_all
```

   Per un elenco più piccolo di istantanee, esegui il comando seguente.

```
GET /_cat/snapshots/aoss-automated
```

2. Esegui il comando seguente per verificare i dettagli dell'istantanea prima del ripristino. Nel comando, sostituisci il *example* contenuto con le tue informazioni specifiche.

```
GET _snapshot/aoss-automated/snapshot-ExampleSnapshotID1/
```

3. Esegui il comando seguente per ripristinare da un'istantanea specifica.

```
POST /_snapshot/aoss-automated/snapshot-ID/_restore
```

   È possibile personalizzare l'operazione di ripristino includendo un corpo della richiesta. Ecco un esempio:

```
POST /_snapshot/aoss-automated/snapshot-
ExampleSnapshotID1-5e01-4423-9833Example/_restore
{
  "indices": "opensearch-dashboards*,my-index*",
  "ignore_unavailable": true,
  "include_global_state": false,
  "include_aliases": false,
  "rename_pattern": "opensearch-dashboards(.+)",
  "rename_replacement": "restored-opensearch-dashboards$1"
}
```

4. Esegui il comando seguente per visualizzare l'avanzamento del ripristino.

```
GET /_cat/recovery
```

> **Note**
>
> Quando si ripristina un'istantanea con un comando che include un corpo della richiesta, è possibile utilizzare diversi parametri per controllare il comportamento di ripristino. Il `indices` parametro specifica quali indici ripristinare e supporta i pattern jolly. Impostato `ignore_unavailable` per continuare l'operazione di ripristino anche se manca un indice nell'istantanea. Utilizzato `include_global_state` per determinare se ripristinare lo stato del cluster e `include_aliases` per controllare se ripristinare gli alias associati. I `rename_replacement` parametri `rename_pattern` and rinominano gli indici durante l'operazione di ripristino.

Supporto codec standard in Amazon Serverless OpenSearch

I codec di indice determinano il modo in cui i campi memorizzati di un indice vengono compressi e archiviati su disco e in S3. Il codec dell'indice è controllato dall'`index.codec`impostazione statica che specifica l'algoritmo di compressione. Questa impostazione influisce sia sulla dimensione dei frammenti di indice che sulle prestazioni del funzionamento dell'indice.

Per impostazione predefinita, gli indici in OpenSearch Serverless utilizzano il codec predefinito con l'algoritmo di compressione. LZ4 OpenSearch Serverless supporta `zstd` e `zstd_no_dict` codifica anche con livelli di compressione configurabili da 1 a 6.

> **Important**
>
> Poiché `index.codec` si tratta di un'impostazione statica, non può essere modificata dopo la creazione dell'indice.

Per maggiori dettagli, consulta la documentazione OpenSearch Index Codecs.

Creazione di un indice con il codec ZSTD

È possibile specificare il codec ZSTD durante la creazione dell'indice utilizzando l'impostazione: `index.codec`

```
PUT /your_index
{
  "settings": {
    "index.codec": "zstd"
  }
}
```

Livelli di compressione

I codec ZSTD supportano livelli di compressione opzionali tramite l'`index.codec.compression_level`impostazione, accettando numeri interi nell'intervallo [1, 6]. Livelli di compressione più elevati si traducono in rapporti di compressione migliori (spazio di archiviazione più piccolo) ma velocità di compressione e decompressione più lente. Il livello di compressione predefinito è 3.

```
PUT /your_index
{
  "settings": {
    "index.codec": "zstd",
    "index.codec.compression_level": 2
  }
}
```

Analisi comparativa delle prestazioni

Sulla base di test di benchmark con il set di dati nyc_taxi, la compressione ZSTD ha ottenuto una compressione migliore del 26-32% rispetto alla linea di base in diverse combinazioni di e livelli di compressione. `zstd` `zstd_no_dict`

Metrica	ZSTD L1	ZSTD L6	ZSTD_NO_DICT L1	ZSTD_NO_DICT L6
Riduzione delle dimensioni dell'indice	28,10%	32%	26,90%	28,70%
Modifica del throughput di indicizzazione	-0,50%	-23,80%	-0,50%	-5,30%

Metrica	ZSTD L1	ZSTD L6	ZSTD_NO_DICT L1	ZSTD_NO_DICT L6
Miglioramento della latenza di Match-all Query p90	-16,40%	29,50%	-16,40%	23,40%
Miglioramento della latenza di Range Query p90	90,90%	92,40%	-282,90%	92,50%
Quantità di distanza (p90 Agg) Miglioramento della latenza	2%	24,70%	2%	13,80%

Per maggiori dettagli, consulta il AWS OpenSearch blog.

Salva spazio di archiviazione utilizzando una fonte derivata

Per impostazione predefinita, OpenSearch Serverless archivia ogni documento importato nel `_source` campo, che contiene il corpo del documento JSON originale, e indicizza i singoli campi per la ricerca. Sebbene il `_source` campo non sia ricercabile, viene mantenuto in modo da poter restituire il documento completo durante l'esecuzione di richieste di recupero, come get e search. Quando la fonte derivata è abilitata, OpenSearch Serverless salta la memorizzazione del `_source` campo e lo ricostruisce invece dinamicamente su richiesta, ad esempio durante le operazioni di ricerca, get, mget, reindex o aggiornamento. L'utilizzo dell'impostazione della fonte derivata può ridurre l'utilizzo dello storage fino al 50%.

Configurazione

Per configurare la fonte derivata per il tuo indice, crea l'indice utilizzando l'`index.derived_source.enabled`impostazione:

```
PUT my-index1
{
  "settings": {
    "index": {
      "derived_source": {
        "enabled": true
      }
    }
  }
```

```
}
```

Considerazioni importanti

- Sono supportati solo alcuni tipi di campo. Per un elenco dei campi e delle limitazioni supportati, consulta la OpenSearch documentazione. Se crei un indice con origine derivata e un campo non supportato, la creazione dell'indice avrà esito negativo. Se si tenta di importare un documento con un campo non supportato in un indice derivato abilitato ai sorgenti, l'importazione avrà esito negativo. Utilizzate questa funzionalità solo quando conoscete i tipi di campo che verranno aggiunti all'indice.
- L'impostazione `index.derived_source.enabled` è statica. Questa impostazione non può essere modificata dopo la creazione dell'indice.

Limitazioni relative alle risposte alle interrogazioni

Quando l'origine derivata è abilitata, impone alcune limitazioni al modo in cui le risposte alle query vengono generate e restituite.

- I campi data con più formati specificati utilizzano sempre il primo formato dell'elenco per tutti i documenti richiesti, indipendentemente dal formato originale importato.
- I valori Geopoint vengono restituiti in un `{"lat": lat_val, "lon": lon_val}` formato fisso e possono perdere una certa precisione.
- Gli array multivalore possono essere ordinati e i campi di parole chiave possono essere deduplicati.

Per maggiori dettagli, consulta il blog. OpenSearch

Analisi comparativa delle prestazioni

Sulla base di test di benchmark con il set di dati nyc_taxi, la fonte derivata ha ottenuto una riduzione del 58% della dimensione dell'indice rispetto al valore di base.

Metrica	Fonte derivata
Riduzione delle dimensioni dell'indice	58,3%
Modifica del throughput di indicizzazione	3,7%

Metrica	Fonte derivata
Indicizzazione della modifica della latenza p90	6,9%
Miglioramento della latenza di Match-all Query p90	19%
Miglioramento della latenza di Range Query p90	-18,8%
Quantità di distanza (p90 Agg): miglioramento della latenza	-7,3%

Per maggiori dettagli, consulta il OpenSearch blog.

Gruppi di raccolta Amazon OpenSearch Serverless

I gruppi di raccolta in Amazon OpenSearch Serverless organizzano più raccolte e consentono la condivisione delle risorse di calcolo tra raccolte con diverse chiavi KMS. Questo modello di elaborazione condiviso riduce i costi eliminando la necessità di unità di OpenSearch calcolo separate (OCUs) per ogni chiave KMS.

Ogni raccolta OpenSearch Serverless creata è protetta con la crittografia dei dati inattivi, utilizzata AWS KMS per archiviare e gestire le chiavi di crittografia. Le raccolte all'interno dello stesso gruppo di raccolta condividono le risorse di calcolo e lo spazio di memoria OCU, anche quando utilizzano chiavi KMS diverse per la crittografia.

I gruppi di raccolta forniscono l'isolamento per i requisiti di sicurezza e prestazioni. È possibile raggruppare le raccolte con la stessa chiave KMS in un unico gruppo di raccolta per l'isolamento della sicurezza o combinare raccolte con diverse chiavi KMS nello stesso gruppo per ottimizzare i costi. Questa flessibilità consente di bilanciare i requisiti di sicurezza con l'efficienza delle risorse.

Quando aggiungi una raccolta a un gruppo di raccolta, OpenSearch Serverless la assegna alle risorse di elaborazione condivise del gruppo. Il sistema gestisce automaticamente la distribuzione dei carichi di lavoro tra queste risorse, mantenendo al contempo la sicurezza crittografando i dati di ciascuna raccolta con la chiave KMS designata. I controlli di accesso continuano ad essere applicati a livello di raccolta e le risorse di elaborazione condivise accedono a più chiavi KMS in base alle necessità per servire le raccolte del gruppo.

È possibile controllare l'allocazione delle risorse impostando limiti OCU minimi e massimi a livello di gruppo di raccolta. Questi limiti si applicano a tutte le raccolte del gruppo e aiutano a gestire i costi garantendo al contempo prestazioni costanti.

Per impostazione predefinita, esiste una quota di servizio (limite) per il numero di raccolte in un gruppo di raccolte, il numero di indici in una raccolta e il numero di raccolte OCUs in un gruppo di raccolte. Per ulteriori informazioni, consulta Quote OpenSearch serverless.

Concetti chiave

Gruppo di raccolta

Una AWS risorsa che funge da contenitore per una o più raccolte. Ogni gruppo di raccolta ha un identificatore univoco e può avere limiti di capacità e impostazioni di configurazione propri.

Elaborazione condivisa

Le raccolte all'interno dello stesso gruppo di raccolte condividono lo stesso set di OCUs, indipendentemente dalle chiavi KMS utilizzate. Questa condivisione riduce i costi eliminando la necessità di risorse di elaborazione separate per ogni chiave KMS.

Limiti di capacità

È possibile impostare limiti OCU minimi e massimi per le operazioni di indicizzazione e ricerca a livello di gruppo di raccolta. Questi limiti aiutano a controllare i costi e a garantire prestazioni costanti.

Limiti di capacità del gruppo di raccolta

I gruppi di raccolta forniscono un controllo granulare sull'allocazione delle risorse attraverso limiti OCU minimi e massimi. Questi limiti si applicano a tutte le raccolte all'interno del gruppo e funzionano indipendentemente dalle impostazioni di capacità a livello di account.

Per impostazione predefinita, esiste una quota di servizio (limite) per il numero di raccolte in un gruppo di raccolte, il numero di indici in una raccolta e il numero di raccolte in un gruppo di OCUs raccolte. Per ulteriori informazioni, consulta Quote OpenSearch serverless.

Comprensione dei limiti di capacità dei gruppi di raccolta

È possibile configurare limiti OCU minimi e massimi per le operazioni di indicizzazione e ricerca a livello di gruppo di raccolta. Questi limiti controllano il modo in cui OpenSearch Serverless ridimensiona le risorse per le raccolte del gruppo:

- OCU minima: il numero minimo di quelle OCUs che OpenSearch Serverless mantiene per il gruppo di raccolta, garantendo prestazioni di base costanti.

  - Se il carico di lavoro richiede un numero inferiore di OCU rispetto al valore minimo specificato, OpenSearch Serverless manterrebbe comunque il valore minimo specificato delle OCU e la fatturazione rifletterebbe lo stesso.
  - Se il carico di lavoro richiede un numero maggiore di OCU rispetto al valore minimo specificato, OpenSearch Serverless manterrebbe il livello di OCU richiesto per il carico di lavoro e la fatturazione rifletterebbe il maggiore utilizzo dell'OCU.
- OCU massimo: è possibile scalare il numero massimo di OCU OpenSearch Serverless per OCUs il gruppo di raccolta, aiutandoti a controllare i costi.

I limiti di capacità del gruppo di raccolta sono disaccoppiati dai limiti a livello di account. Le impostazioni OCU massime a livello di account si applicano solo alle raccolte non associate a nessun gruppo di raccolta, mentre le impostazioni OCU massime del gruppo di raccolta si applicano alle raccolte all'interno di quel gruppo specifico.

Valori limite di capacità validi

Quando si impostano i limiti OCU minimi e massimi per un gruppo di raccolta, è possibile utilizzare solo i valori del seguente set: 1, 2, 4, 8, 16 e multipli di 16 (ad esempio 32, 48, 64, 80, 96) fino a un massimo di 1.696. OCUs

I limiti OCU minimi e massimi sono opzionali quando si crea un gruppo di raccolta. Se non si specifica un limite OCU massimo, OpenSearch Serverless utilizza il valore predefinito di 96. OCUs

Il limite OCU minimo deve essere inferiore o uguale al limite OCU massimo.

Comprensione della relazione tra i limiti OCU a livello di account e del gruppo di raccolta

Quando si pianifica la capacità OpenSearch Serverless, è importante comprendere come interagiscono i limiti OCU a livello di account e i limiti OCU del gruppo di raccolta. La somma delle impostazioni OCU massime in tutti i gruppi di raccolta più l'impostazione OCU massima a livello di account deve essere inferiore o uguale al limite di quota di servizio per account. Per i valori limite correnti, vedere Quote OpenSearch serverless.

> **Note**
>
> Le impostazioni OCU massime a livello di account si applicano solo alle raccolte non associate a nessun gruppo di raccolta. Le raccolte all'interno dei gruppi di raccolta sono regolate dai rispettivi limiti dei gruppi di raccolta, non dai limiti a livello di account.

Questo vincolo si applica sia all'indicizzazione che alla ricerca in modo indipendente. OCUs Ad esempio, se configuri impostazioni a livello di account e gruppi di raccolta, devi assicurarti che il totale non superi il limite della quota di servizio per l'indicizzazione OCUs e separatamente non superi il limite della quota di servizio per la ricerca. OCUs Inoltre, puoi creare un massimo di 300 gruppi di raccolta per account.

Esempio: capacità di pianificazione con limiti a livello di account e gruppo di raccolta

Se si imposta l'OCU di ricerca massima a livello di account su 500 e il limite della quota di servizio è 1.700:

- E create 2 gruppi di raccolta, la somma dell'OCU massimo per i 2 gruppi di raccolta non deve essere superiore a 1.200 (1.700 - 500)
- È possibile lasciare ogni gruppo di raccolta all'OCU massimo predefinito di 96 (96 + 96 + 500 = 692), lasciando spazio per la crescita futura
- Oppure puoi aumentare il massimo di ogni gruppo di raccolta a 600 (600 + 600 + 500 = 1.700), utilizzando la piena capacità consentita dalla quota di servizio

Questa relazione è fondamentale per la pianificazione della capacità. Prima di creare nuovi gruppi di raccolta o aumentare i limiti massimi di OCU, verifica che l'allocazione totale non superi il limite della quota di servizio. Se raggiungi questo limite, devi ridurre le impostazioni OCU massime sui gruppi di raccolta esistenti o ridurre le impostazioni OCU massime a livello di account per fare spazio a nuove allocazioni.

Configurazione dei limiti di capacità

È possibile impostare limiti di capacità quando si crea un gruppo di raccolta o li si aggiorna in un secondo momento. Per configurare i limiti di capacità utilizzando AWS CLI, utilizzate i UpdateCollectionGroupcomandi CreateCollectionGroupo:

```
aws opensearchserverless create-collection-group \
```

```
    --name my-collection-group \
    --capacity-limits
 maxIndexingCapacityInOCU=32,maxSearchCapacityInOCU=32,minIndexingCapacityInOCU=4,minSearchCapa
```

Per aggiornare i limiti di capacità per un gruppo di raccolta esistente:

```
aws opensearchserverless update-collection-group \
    --id abcdef123456 \
    --capacity-limits
 maxIndexingCapacityInOCU=48,maxSearchCapacityInOCU=48,minIndexingCapacityInOCU=8,minSearchCapa
```

Monitoraggio della capacità del gruppo di raccolta

OpenSearch Serverless emette i seguenti parametri di Amazon CloudWatch Logs a intervalli di un minuto per aiutarti a monitorare l'utilizzo dell'OCU e i limiti di capacità a livello di gruppo di raccolta:

- `IndexingOCU`— Il numero di indicizzazione attualmente utilizzato dal gruppo di raccolta OCUs .
- `SearchOCU`— Il numero di ricerche OCUs attualmente utilizzate dal gruppo di raccolta.

OpenSearch Serverless emette anche metriche OCU a livello di account per raccolte non associate a nessun gruppo di raccolta. Puoi aggregare queste metriche CloudWatch per visualizzare la somma di tutti i gruppi di raccolta e le raccolte a OCUs livello di account.

Configura gli allarmi per avvisarti quando il gruppo di raccolta si avvicina ai limiti di capacità in modo da poter modificare le impostazioni secondo necessità. Per ulteriori informazioni sulle metriche OpenSearch Serverless, consulta. the section called “Monitoraggio senza server OpenSearch ”

Come vengono applicati i limiti di capacità

OpenSearch Serverless applica i limiti di capacità dei gruppi di raccolta durante le operazioni di scalabilità. Quando le tue raccolte richiedono risorse aggiuntive, OpenSearch Serverless è scalabile fino al limite massimo di OCU. Quando la domanda diminuisce, OpenSearch Serverless si ridimensiona ma mantiene almeno il limite minimo di OCU per garantire prestazioni costanti.

I limiti di capacità vengono applicati solo quando il gruppo di raccolta contiene almeno una raccolta. I gruppi di raccolta vuoti non consumano OCUs né impongono limiti di capacità.

Se un'operazione di scalabilità dovesse superare il limite massimo di OCU o violare il requisito minimo dell'OCU, OpenSearch Serverless rifiuta l'operazione per mantenere la conformità ai limiti configurati.

Crittografia e chiavi KMS nei gruppi di raccolta

Ogni raccolta OpenSearch Serverless che crei è protetta con la crittografia dei dati inattivi, utilizzata AWS KMS per archiviare e gestire le chiavi di crittografia. Quando lavori con i gruppi di raccolta, hai flessibilità nel modo in cui specifichi la chiave KMS per le tue raccolte.

Puoi fornire la chiave KMS associata a una raccolta in due modi:

- Nella CreateCollection richiesta: specifica la chiave KMS direttamente quando crei la raccolta utilizzando il `encryption-config` parametro.
- Nelle politiche di sicurezza: definisci l'associazione delle chiavi KMS in una politica di sicurezza di crittografia.

Quando si specifica una chiave KMS in entrambe le posizioni, la chiave KMS fornita nella CreateCollection richiesta ha la precedenza sulla configurazione della politica di sicurezza.

Questa flessibilità semplifica la gestione delle collezioni su larga scala, in particolare quando è necessario creare più raccolte con chiavi KMS uniche. Invece di creare e gestire migliaia di politiche di crittografia, puoi specificare la chiave KMS direttamente durante la creazione della raccolta.

Condivisione OCUs tra diverse chiavi KMS

I gruppi di raccolta consentono la condivisione delle risorse di calcolo tra raccolte con chiavi KMS diverse. Le raccolte dello stesso gruppo di raccolta condividono lo spazio di memoria OCU, indipendentemente dalle relative chiavi di crittografia. Questo modello di elaborazione condiviso riduce i costi eliminando la necessità di utilizzare chiavi OCUs KMS separate.

I gruppi di raccolta forniscono l'isolamento per i requisiti di sicurezza e prestazioni. È possibile raggruppare le raccolte con la stessa chiave KMS in un unico gruppo di raccolta per l'isolamento della sicurezza o combinare raccolte con diverse chiavi KMS nello stesso gruppo per ottimizzare i costi. Questa flessibilità consente di bilanciare i requisiti di sicurezza con l'efficienza delle risorse.

Il sistema mantiene la sicurezza crittografando i dati di ogni raccolta con la chiave KMS designata. I controlli di accesso continuano ad essere applicati a livello di raccolta e le risorse di elaborazione condivise accedono a più chiavi KMS in base alle necessità per servire le raccolte del gruppo.

Autorizzazioni KMS richieste

Quando specifichi una chiave KMS nella CreateCollection richiesta, hai bisogno delle seguenti autorizzazioni aggiuntive:

- `kms:DescribeKey`— Consente a OpenSearch Serverless di recuperare informazioni sulla chiave KMS.
- `kms:CreateGrant`— Consente a OpenSearch Serverless di creare una concessione per la chiave KMS per abilitare le operazioni di crittografia.

Queste autorizzazioni non sono richieste quando si utilizzano AWS chiavi di proprietà.

Crea gruppi di raccolta

Questo argomento descrive come creare, configurare e gestire gruppi di raccolta in Amazon OpenSearch Serverless. Usa i gruppi di raccolta per organizzare le raccolte e condividere le risorse di calcolo per ottimizzare i costi. Imposta limiti OCU minimi e massimi a livello di gruppo di raccolta per controllare le prestazioni e la spesa.

Crea un gruppo di raccolta

Utilizzate le seguenti procedure per creare un nuovo gruppo di raccolta e configurarne le impostazioni. Crea un gruppo di raccolta utilizzando la console OpenSearch Serverless o il AWS SDKs . AWS CLI Quando create un gruppo di raccolta, specificate i limiti di capacità e altre opzioni di configurazione.

Console

1. Apri la console Amazon OpenSearch Service all'indirizzo https://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Serverless, quindi scegli Collection groups
3. Scegli Crea gruppo di raccolta.
4. Per Nome del gruppo di raccolta, inserisci un nome per il tuo gruppo di raccolta. Il nome deve avere una lunghezza di 3-32 caratteri, iniziare con una lettera minuscola e contenere solo lettere minuscole, numeri e trattini.
5. (Facoltativo) In Descrizione, inserisci una descrizione per il tuo gruppo di raccolta.
6. Nella sezione Gestione della capacità, configura i limiti dell'OCU:
   - Capacità massima di indicizzazione: il numero massimo di indicizzazione OCUs che le raccolte di questo gruppo possono raggiungere.
   - Capacità di ricerca massima: il numero massimo di ricerche fino a OCUs cui le raccolte di questo gruppo possono aumentare.

- Capacità di indicizzazione minima: il numero minimo di indicizzazione da mantenere OCUs per prestazioni costanti.
- Capacità di ricerca minima: il numero minimo di ricerche da mantenere OCUs per prestazioni costanti.

7. (Facoltativo) Nella sezione Tag, aggiungi tag per organizzare e identificare il tuo gruppo di raccolta.
8. Scegli Crea gruppo di raccolta.

AWS CLI

- Usa il create-collection-groupcomando per creare un nuovo gruppo di raccolta. Nel comando, sostituisci il *example* contenuto con le tue informazioni specifiche.

```
aws opensearchserverless create-collection-group \
    --name my-collection-group \
    --description "Collection group for production workloads" \
    --capacity-limits
 maxIndexingCapacityInOCU=20,maxSearchCapacityInOCU=20,minIndexingCapacityInOCU=2,minSea
 \
    --tags key=Environment,value=Production key=Team,value=DataEngineering
```

Il comando restituisce i dettagli sul gruppo di raccolta creato, inclusi l'ID univoco e l'ARN.

Aggiungi una nuova collezione a un gruppo di raccolte

Quando create una nuova collezione, specificate il nome del gruppo di raccolta esistente a cui associare la raccolta. Utilizzate le seguenti procedure per aggiungere una nuova raccolta a un gruppo di raccolte.

Console

1. Apri la console Amazon OpenSearch Service all'indirizzo https://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Serverless, quindi scegli Collezioni
3. Scegli Create collection (Crea raccolta).

4. Per Nome della raccolta, inserisci un nome per la tua raccolta. Il nome deve avere una lunghezza di 3-28 caratteri, iniziare con una lettera minuscola e contenere solo lettere minuscole, numeri e trattini.
5. (Facoltativo) In Descrizione, inserisci una descrizione per la tua collezione.
6. Nella sezione Gruppo di raccolta, seleziona il gruppo di raccolta a cui desideri assegnare la raccolta. Una raccolta può appartenere a un solo gruppo di raccolte alla volta.

   (Facoltativo) Puoi anche scegliere di creare un nuovo gruppo. In questo modo accederai al flusso di lavoro Crea gruppo di raccolta. Dopo aver completato la creazione del gruppo di raccolta, torna al passaggio 1 di questa procedura per iniziare a creare la nuova raccolta.
7. Continua con il flusso di lavoro per creare la raccolta.

> **Important**
>
> Non uscite dal flusso di lavoro Crea raccolta. In questo modo si interromperà la configurazione della raccolta. La pagina dei dettagli della collezione viene visualizzata al termine della configurazione.

AWS CLI

- Utilizzate il comando create-collection per creare una nuova raccolta e aggiungerla a un gruppo di raccolte esistente. Nel comando, sostituisci il *example* contenuto con le tue informazioni specifiche.

```
aws opensearchserverless create-collection \
    --name my-collection \
    --type SEARCH \
    --collection-group-name my-collection-group \
    --description "Collection for search workloads"
```

Gestisci i gruppi di raccolta Amazon OpenSearch Serverless

Dopo aver creato i gruppi di raccolta Amazon OpenSearch Serverless, puoi modificarne le impostazioni in base alle tue esigenze. Utilizza queste operazioni di gestione per aggiornare i limiti di

capacità e visualizzare i dettagli del gruppo di raccolta. Queste modifiche consentono di ottimizzare l'allocazione delle risorse e mantenere un'organizzazione efficiente delle raccolte.

Visualizza i gruppi di collezioni

Visualizza i tuoi gruppi di raccolte OpenSearch Serverless per esaminarne le configurazioni, le raccolte associate e lo stato attuale.

Console

1. Apri la console Amazon OpenSearch Service all'indirizzo https://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Serverless, quindi scegli Collezioni
3. Scegli la scheda Gruppi di raccolte. Vengono visualizzati i gruppi di raccolta del tuo account.
4. Scegli il nome di un gruppo di raccolte per visualizzarne i dettagli.

AWS CLI

- Usa il list-collection-groupscomando per elencare tutti i gruppi di raccolta nel tuo account. Utilizzate il batch-get-collection-groupcomando per visualizzare i dettagli su gruppi di raccolta specifici. Nei comandi seguenti, sostituite il *example* contenuto con le vostre informazioni specifiche.

  Per elencare tutti i gruppi di raccolta:

  ```
  aws opensearchserverless list-collection-groups
  ```

  Per ottenere dettagli su gruppi di raccolta specifici:

  ```
  aws opensearchserverless batch-get-collection-group \
      --names my-collection-group another-group
  ```

Aggiorna le impostazioni dei gruppi di raccolta

Aggiorna le impostazioni del gruppo di raccolta OpenSearch Serverless per modificare configurazioni come i limiti di capacità e la descrizione.

Console

1. Apri la console Amazon OpenSearch Service all'indirizzo https://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Serverless, quindi scegli Collezioni
3. Scegli la scheda Gruppi di raccolte. Vengono visualizzati i gruppi di raccolta del tuo account.
4. Scegli il nome di un gruppo di raccolte per visualizzarne i dettagli.
5. In Dettagli del gruppo di raccolta, scegli Modifica.
6. Apporta le modifiche, quindi scegli Salva.

AWS CLI

- Utilizzate il update-collection-groupcomando per aggiornare la descrizione e i limiti di capacità di un gruppo di raccolta esistente. Nel comando seguente, sostituite il *example* contenuto con le vostre informazioni.

```
aws opensearchserverless update-collection-group \
    --id abcdef123456 \
    --description "Updated description for production workloads" \
    --capacity-limits
 maxIndexingCapacityInOCU=30,maxSearchCapacityInOCU=30,minIndexingCapacityInOCU=4,minSea
```

Le modifiche ai limiti di capacità hanno effetto immediato e potrebbero influire sul comportamento di ridimensionamento delle raccolte del gruppo.

Eliminare i gruppi di raccolta

Prima di poter eliminare un gruppo di raccolte, è necessario rimuovere tutte le raccolte dal gruppo. Non è possibile eliminare un gruppo di raccolte che contiene raccolte.

Console

1. Apri la console Amazon OpenSearch Service all'indirizzo https://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Serverless, quindi scegli Collezioni
3. Scegli la scheda Gruppi di raccolte. Vengono visualizzati i gruppi di raccolta del tuo account.
4. Scegli il nome del gruppo di raccolta che desideri eliminare.

> **Important**
>
> Rimuovi tutte le raccolte dal gruppo di raccolta aggiornando ogni raccolta per rimuovere l'associazione del gruppo di raccolta o spostandole in altri gruppi di raccolta.

5. Nella parte superiore della pagina, scegli Delete (Elimina).
6. Conferma l'eliminazione, quindi scegli Elimina.

AWS CLI

- Utilizzate il delete-collection-groupcomando per eliminare un gruppo di raccolta.

> **Important**
>
> Rimuovi tutte le raccolte dal gruppo di raccolta aggiornando ogni raccolta per rimuovere l'associazione del gruppo di raccolta o spostandole in altri gruppi di raccolta.

Nel comando seguente, sostituite il *example* contenuto con le vostre informazioni.

Eliminate il gruppo di raccolta vuoto:

```
aws opensearchserverless delete-collection-group \
    --id abcdef123456
```

Gestione dei limiti di capacità per Amazon OpenSearch Serverless

Con Amazon OpenSearch Serverless, non devi gestire la capacità da solo. OpenSearch Serverless ridimensiona automaticamente la capacità di elaborazione del tuo account in base al carico di lavoro corrente. La capacità di elaborazione serverless viene misurata in unità di calcolo (). OpenSearch OCUs Ogni OCU è una combinazione di 6 GiB di memoria e della CPU virtuale (vCPU) corrispondente, oltre al trasferimento dei dati su Amazon S3. Per ulteriori informazioni sull'architettura disaccoppiata in Serverless, vedere. OpenSearch the section called “Come funziona”

Quando crei la tua prima raccolta, OpenSearch Serverless crea istanze OCUs in base alle tue impostazioni di ridondanza. Per impostazione predefinita, sono abilitate le repliche attive ridondanti, che ne creano quattro OCUs (due per l'indicizzazione e due per la ricerca). Ciò garantisce un'elevata disponibilità con nodi di standby in un'altra zona di disponibilità.

Per lo sviluppo e il test, puoi disabilitare l'impostazione Abilita ridondanza per una raccolta. Ciò rimuove le repliche in standby e ne utilizza solo due OCUs (una per l'indicizzazione e una per la ricerca).

Esistono OCUs sempre, anche quando non è prevista alcuna attività di indicizzazione o ricerca. Tutte le raccolte successive possono condividerle OCUs, ad eccezione delle raccolte con AWS KMS chiavi univoche, che creano un'istanza del proprio set di. OCUs Tutte le raccolte associate a un gruppo di raccolte possono condividere lo stesso set di. OCUs È possibile includere un solo tipo di raccolta (ricerca, serie temporale o ricerca vettoriale) in un singolo gruppo di raccolte. Per ulteriori informazioni, consulta the section called “Gruppi di raccolta”.

OpenSearch Serverless si ridimensiona e aggiunge automaticamente OCUs man mano che l'indicizzazione e l'utilizzo della ricerca aumentano. Quando il traffico diminuisce, la capacità si riduce al numero minimo OCUs richiesto per la dimensione dei dati.

Per le raccolte di ricerche e serie temporali, il numero di dati OCUs richiesti in caso di inattività è proporzionale alla dimensione dei dati e al numero di indici. Per le raccolte vettoriali, i requisiti OCU dipendono dalla memoria (RAM) per archiviare i grafici vettoriali e dallo spazio su disco per archiviare gli indici. Quando non è inattiva, i requisiti OCU tengono conto di entrambi i fattori.

Le raccolte vettoriali archiviano i dati dell'indice nella memoria locale OCU. I limiti di RAM OCU vengono raggiunti più rapidamente dei limiti del disco, il che limita le raccolte vettoriali in base allo spazio RAM.

Con la ridondanza abilitata, la capacità dell'OCU si riduce a un minimo di 1 OCU (0,5 OCU x 2) per l'indicizzazione e 1 OCU (0,5 OCU x 2) per la ricerca. Quando si disabilita la ridondanza, la raccolta può essere ridotta a 0,5 OCU per l'indicizzazione e 0,5 OCU per la ricerca.

Il ridimensionamento tiene conto anche del numero di frammenti necessari per la raccolta o l'indice. Ogni OCU supporta un numero specifico di shard e il numero di indici deve essere proporzionale al numero di shard. Il numero totale di base OCUs richiesto è il massimo dei requisiti di dati, memoria e shard. Per ulteriori informazioni, consulta le funzionalità di ricerca economiche di Amazon OpenSearch Serverless, su qualsiasi scala, sul AWS Big Data Blog.

Per le raccolte di ricerca e di ricerca vettoriale, tutti i dati vengono archiviati su indici caldi per garantire tempi di risposta rapidi alle query. Le raccolte di serie temporali utilizzano una combinazione di archiviazione a caldo e a caldo, conservando i dati più recenti in una memoria a caldo per ottimizzare i tempi di risposta alle query per i dati a cui si accede più frequentemente. Per ulteriori informazioni, consulta the section called “Scelta di un tipo di raccolta”.

> **Note**
>
> Una raccolta di ricerca vettoriale non può essere condivisa OCUs con le raccolte di ricerca e le raccolte di serie temporali, anche se la raccolta di ricerca vettoriale utilizza la stessa chiave KMS delle raccolte di ricerca o di serie temporali. OCUs Verrà creato un nuovo set di immagini per la tua prima raccolta vettoriale. Le raccolte OCUs vettoriali sono condivise tra le stesse raccolte di chiavi KMS.

Per gestire la capacità delle raccolte e controllare i costi, è possibile specificare la capacità massima complessiva di indicizzazione e ricerca per l'account corrente e la regione, e OpenSearch Serverless ridimensiona automaticamente le risorse di raccolta in base a queste specifiche.

Poiché la capacità di indicizzazione e di ricerca è scalabile separatamente, è necessario specificare i limiti a livello di account per ciascuna di esse:

- Capacità di indicizzazione massima: OpenSearch Serverless può aumentare la capacità di indicizzazione fino a questo numero di. OCUs
- Capacità di ricerca massima: OpenSearch Serverless può aumentare la capacità di ricerca fino a questo numero di. OCUs

> **Note**
>
> Al momento, le impostazioni relative alla capacità si applicano solo a livello di account. Non è possibile configurare limiti di capacità per ciascuna raccolta.

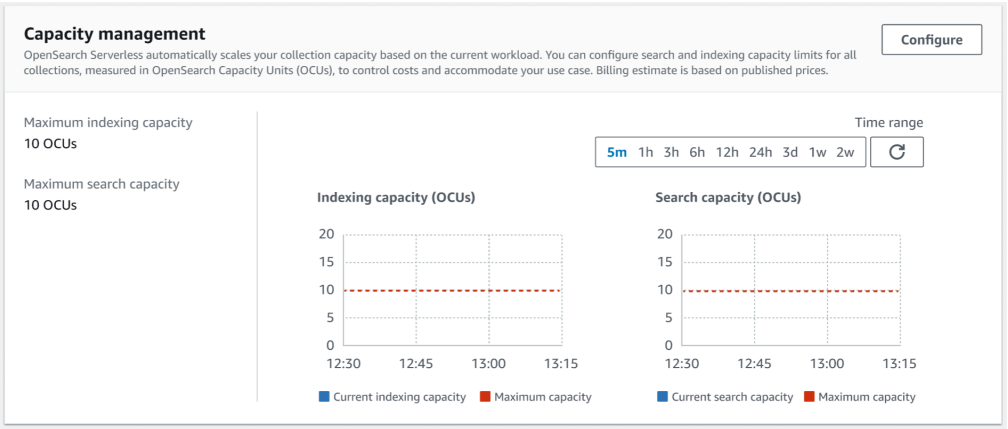
L'obiettivo è quello di garantire che la capacità massima sia sufficientemente elevata da gestire i picchi del carico di lavoro. In base alle impostazioni, OpenSearch Serverless ridimensiona automaticamente il numero di raccolte OCUs per elaborare il carico di lavoro di indicizzazione e ricerca.

Argomenti

- Configurazione delle impostazioni di capacità
- Limiti di capacità massima
- Monitoraggio dell'utilizzo della capacità

Configurazione delle impostazioni di capacità

Per configurare le impostazioni di capacità nella console Serverless, espandi OpenSearch Serverless nel riquadro di navigazione a sinistra e seleziona Dashboard. In Capacity management (Gestione della capacità) specifica la capacità massima di indicizzazione e ricerca:



Per configurare la capacità utilizzando AWS CLI, invia una UpdateAccountSettingsrichiesta:

```
aws opensearchserverless update-account-settings \
    --capacity-limits '{ "maxIndexingCapacityInOCU": 8,"maxSearchCapacityInOCU": 9 }'
```

Limiti di capacità massima

Il totale massimo di indici che una raccolta può contenere è 1000. Per tutti e tre i tipi di raccolte, la capacità OCU massima predefinita è 10 OCUs per l'indicizzazione e 10 per la ricerca. OCUs La capacità OCU minima consentita per un account è di 1 OCU [0,5 OCU x 2] per l'indicizzazione e 1 OCU [0,5 OCU x 2] per la ricerca. Per tutte le raccolte, la capacità massima consentita è di 1.700 per l'indicizzazione e 1.700 per la ricerca OCUs . OCUs È possibile configurare il conteggio OCU in modo che sia qualsiasi numero compreso tra 2 e la capacità massima consentita, in multipli di 2.

Ogni OCU include una quantità di storage temporaneo a caldo sufficiente per 120 GiB di dati di indice. OpenSearch Serverless supporta fino a 1 TiB di dati per indice nelle raccolte di ricerca e di ricerca vettoriale e 100 TiB di hot data per indice in una raccolta di serie temporali. Per le raccolte di serie temporali, puoi comunque inserire più dati, che possono essere archiviati come dati «warm data» in S3.

Per un elenco di tutte le quote, vedi OpenSearch Quote serverless.

Monitoraggio dell'utilizzo della capacità

Puoi monitorare le `SearchOCU` CloudWatch metriche a `IndexingOCU` livello di account per capire in che modo le tue collezioni stanno scalando. Ti consigliamo di configurare allarmi che possono avvisarti se il tuo account si sta avvicinando a una soglia per i parametri relativi alla capacità, in modo da poter adattare di conseguenza le impostazioni di capacità.

Puoi anche utilizzare questi parametri per determinare se le impostazioni di capacità massima sono appropriate o se è necessario adeguarle. Analizza questi parametri per concentrarti sull'ottimizzazione dell'efficienza delle tue raccolte. Per ulteriori informazioni sulle metriche a cui OpenSearch Serverless invia, consulta. CloudWatch the section called “Monitoraggio senza server OpenSearch ”

Inserimento di dati in raccolte Amazon Serverless OpenSearch

Queste sezioni forniscono dettagli sulle pipeline di importazione supportate per l'inserimento di dati nelle raccolte Amazon OpenSearch Serverless. Esse riguardano anche alcuni client che puoi utilizzare per interagire con le operazioni dell'API. OpenSearch I tuoi client devono essere compatibili con OpenSearch 2.x per integrarsi con OpenSearch Serverless.

Argomenti

- Autorizzazioni minime richieste
- OpenSearch Ingestione
- Fluent Bit
- Amazon Data Firehose
- Go
- Java
- JavaScript
- Logstash
- Python
- Ruby
- Firma delle richieste HTTP con altri client

Autorizzazioni minime richieste

Per inserire dati in una raccolta OpenSearch Serverless, il responsabile che scrive i dati deve disporre delle seguenti autorizzazioni minime assegnate in una politica di accesso ai dati:

```
[
   {
      "Rules":[
         {
            "ResourceType":"index",
            "Resource":[
               "index/target-collection/logs"
            ],
            "Permission":[
               "aoss:CreateIndex",
               "aoss:WriteDocument",
               "aoss:UpdateIndex"
            ]
         }
      ],
      "Principal":[
         "arn:aws:iam::123456789012:user/my-user"
      ]
   }
]
```

Le autorizzazioni possono essere più ampie se prevedi di scrivere su indici aggiuntivi. Ad esempio, anziché specificare un singolo indice di destinazione, è possibile concedere l'autorizzazione a tutti gli indici (index/ *target-collection* /*) o a un sottoinsieme di indici (index//). *target-collection logs**

Per un riferimento a tutte le operazioni API disponibili e alle relative autorizzazioni, consulta. OpenSearch the section called “Operazioni e plug-in supportati”

OpenSearch Ingestione

Invece di utilizzare un client di terze parti per inviare dati direttamente a una raccolta OpenSearch Serverless, puoi utilizzare Amazon OpenSearch Ingestion. Puoi configurare i tuoi produttori di dati per inviare dati a OpenSearch Ingestion, che li consegna automaticamente alla raccolta specificata. Puoi anche configurare OpenSearch Ingestion per trasformare i dati prima di consegnarli. Per ulteriori informazioni, consulta OpenSearch Ingestione di Amazon.

Una pipeline OpenSearch di Ingestion necessita dell'autorizzazione per scrivere su una raccolta OpenSearch Serverless configurata come sink. Queste autorizzazioni includono la possibilità di descrivere la raccolta e inviarle richieste HTTP. Per istruzioni su come utilizzare OpenSearch Ingestion per aggiungere dati a una raccolta, consulta. the section called “Concedere alle pipeline l'accesso alle raccolte”

Per iniziare a usare OpenSearch Ingestion, consulta. the section called “Tutorial: inserisci dati in una raccolta”

Fluent Bit

È possibile utilizzare Fluent Bit image e il plug-in di OpenSearch output AWS per inserire dati in raccolte Serverless. OpenSearch

> **Note**
>
> È necessario disporre della versione 2.30.0 o successiva dell'immagine AWS for Fluent Bit per l'integrazione con Serverless. OpenSearch

Configurazione di esempio:

Questa sezione di output di esempio del file di configurazione mostra come utilizzare una raccolta OpenSearch Serverless come destinazione. L'aggiunta importante è il parametro `AWS_Service_Name`, che è `aoss`. `Host` è l'endpoint della raccolta.

```
[OUTPUT]
    Name  opensearch
    Match *
    Host  collection-endpoint.us-west-2.aoss.amazonaws.com
    Port  443
    Index  my_index
    Trace_Error On
    Trace_Output On
    AWS_Auth On
    AWS_Region <region>
    AWS_Service_Name aoss
    tls     On
    Suppress_Type_Name On
```

Amazon Data Firehose

Firehose supporta OpenSearch Serverless come destinazione di consegna. Per istruzioni su come inviare dati in OpenSearch Serverless, consulta Creating a Kinesis Data Firehose Delivery Stream e OpenSearch Choose Serverless for Your Destination nella Amazon Data Firehose Developer Guide.

Il ruolo IAM che fornisci a Firehose per la consegna deve essere specificato all'interno di una policy di accesso ai dati con l'autorizzazione `aoss:WriteDocument` minima per la raccolta di destinazione e devi disporre di un indice preesistente a cui inviare i dati. Per ulteriori informazioni, consulta the section called “Autorizzazioni minime richieste”.

Prima di inviare dati a OpenSearch Serverless, potrebbe essere necessario eseguire delle trasformazioni sui dati. Per ulteriori informazioni su come usare le funzioni Lambda per completare questa attività, consultare Trasformazione dei dati di Amazon Kinesis Data Firehose nella stessa guida.

Go

Il codice di esempio seguente utilizza il client opensearch-go per Go per stabilire una connessione sicura alla raccolta OpenSearch Serverless specificata e creare un singolo indice. È necessario fornire valori per `region` e `host`.

```
package main
```

```
import (
  "context"
  "log"
  "strings"
  "github.com/aws/aws-sdk-go-v2/aws"
  "github.com/aws/aws-sdk-go-v2/config"
  opensearch "github.com/opensearch-project/opensearch-go/v2"
  opensearchapi "github.com/opensearch-project/opensearch-go/v2/opensearchapi"
  requestsigner "github.com/opensearch-project/opensearch-go/v2/signer/awsv2"
)

const endpoint = "" // serverless collection endpoint

func main() {
 ctx := context.Background()

 awsCfg, err := config.LoadDefaultConfig(ctx,
  config.WithRegion("<AWS_REGION>"),
  config.WithCredentialsProvider(
   getCredentialProvider("<AWS_ACCESS_KEY>", "<AWS_SECRET_ACCESS_KEY>",
 "<AWS_SESSION_TOKEN>"),
  ),
 )
 if err != nil {
  log.Fatal(err) // don't log.fatal in a production-ready app
 }

 // create an AWS request Signer and load AWS configuration using default config folder
 or env vars.
 signer, err := requestsigner.NewSignerWithService(awsCfg, "aoss") // "aoss" for Amazon
 OpenSearch Serverless
 if err != nil {
  log.Fatal(err) // don't log.fatal in a production-ready app
 }

 // create an opensearch client and use the request-signer
 client, err := opensearch.NewClient(opensearch.Config{
  Addresses: []string{endpoint},
  Signer:    signer,
 })
 if err != nil {
  log.Fatal("client creation err", err)
 }
```

```
 indexName := "go-test-index"

  // define index mapping
 mapping := strings.NewReader(`{
  "settings": {
    "index": {
         "number_of_shards": 4
         }
       }
  }`)

 // create an index
 createIndex := opensearchapi.IndicesCreateRequest{
  Index: indexName,
    Body: mapping,
 }
 createIndexResponse, err := createIndex.Do(context.Background(), client)
 if err != nil {
  log.Println("Error ", err.Error())
  log.Println("failed to create index ", err)
  log.Fatal("create response body read err", err)
 }
 log.Println(createIndexResponse)

 // delete the index
 deleteIndex := opensearchapi.IndicesDeleteRequest{
  Index: []string{indexName},
 }

 deleteIndexResponse, err := deleteIndex.Do(context.Background(), client)
 if err != nil {
  log.Println("failed to delete index ", err)
  log.Fatal("delete index response body read err", err)
 }
 log.Println("deleting index", deleteIndexResponse)
}

func getCredentialProvider(accessKey, secretAccessKey, token string)
 aws.CredentialsProviderFunc {
 return func(ctx context.Context) (aws.Credentials, error) {
  c := &aws.Credentials{
   AccessKeyID:     accessKey,
   SecretAccessKey: secretAccessKey,
```

```
   SessionToken:    token,
  }
  return *c, nil
 }
}
```

Java

Il codice di esempio seguente utilizza il client opensearch-java per Java per stabilire una connessione sicura alla raccolta OpenSearch Serverless specificata e creare un singolo indice. È necessario fornire valori per `region` e `host`.

La differenza importante rispetto ai domini di OpenSearch servizio è il nome del servizio (anziché). `aoss` `es`

```
// import OpenSearchClient to establish connection to OpenSearch Serverless collection
import org.opensearch.client.opensearch.OpenSearchClient;

SdkHttpClient httpClient = ApacheHttpClient.builder().build();
// create an opensearch client and use the request-signer
OpenSearchClient client = new OpenSearchClient(
    new AwsSdk2Transport(
        httpClient,
        "...us-west-2.aoss.amazonaws.com", // serverless collection endpoint
        "aoss" // signing service name
        Region.US_WEST_2, // signing service region
        AwsSdk2TransportOptions.builder().build()
    )
);

String index = "sample-index";

// create an index
CreateIndexRequest createIndexRequest = new
 CreateIndexRequest.Builder().index(index).build();
CreateIndexResponse createIndexResponse = client.indices().create(createIndexRequest);
System.out.println("Create index reponse: " + createIndexResponse);

// delete the index
DeleteIndexRequest deleteIndexRequest = new
 DeleteIndexRequest.Builder().index(index).build();
DeleteIndexResponse deleteIndexResponse = client.indices().delete(deleteIndexRequest);
System.out.println("Delete index reponse: " + deleteIndexResponse);
```

```
httpClient.close();
```

Il codice di esempio seguente stabilisce nuovamente una connessione sicura e quindi esegue la ricerca in un indice.

```
import org.opensearch.client.opensearch.OpenSearchClient;

SdkHttpClient httpClient = ApacheHttpClient.builder().build();

OpenSearchClient client = new OpenSearchClient(
    new AwsSdk2Transport(
        httpClient,
        "...us-west-2.aoss.amazonaws.com", // serverless collection endpoint
        "aoss" // signing service name
        Region.US_WEST_2, // signing service region
        AwsSdk2TransportOptions.builder().build()
    )
);

Response response = client.generic()
    .execute(
        Requests.builder()
            .endpoint("/" + "users" + "/_search?typed_keys=true")
            .method("GET")
            .json("{"
                + "    \"query\": {"
                + "        \"match_all\": {}"
                + "    }"
                + "}")
            .build());

httpClient.close();
```

JavaScript

Il codice di esempio seguente utilizza il client opensearch-js per JavaScript stabilire una connessione sicura alla raccolta OpenSearch Serverless specificata, creare un singolo indice, aggiungere un documento ed eliminare l'indice. È necessario fornire valori per `node` e `region`.

La differenza importante rispetto ai domini di OpenSearch servizio è il nome del servizio (anziché). `aoss` `es`

Version 3

Questo esempio utilizza la versione 3 dell'SDK for JavaScript in Node.js.

```
const { defaultProvider } = require('@aws-sdk/credential-provider-node');
const { Client } = require('@opensearch-project/opensearch');
const { AwsSigv4Signer } = require('@opensearch-project/opensearch/aws');

async function main() {
    // create an opensearch client and use the request-signer
    const client = new Client({
        ...AwsSigv4Signer({
            region: 'us-west-2',
            service: 'aoss',
            getCredentials: () => {
                const credentialsProvider = defaultProvider();
                return credentialsProvider();
            },
        }),
        node: '' # // serverless collection endpoint
    });

    const index = 'movies';

    // create index if it doesn't already exist
    if (!(await client.indices.exists({ index })).body) {
        console.log((await client.indices.create({ index })).body);
    }

    // add a document to the index
    const document = { foo: 'bar' };
    const response = await client.index({
        id: '1',
        index: index,
        body: document,
    });
    console.log(response.body);

    // delete the index
    console.log((await client.indices.delete({ index })).body);
}

main();
```

Version 2

Questo esempio utilizza la versione 2 dell'SDK for JavaScript in Node.js.

```
const AWS = require('aws-sdk');
const { Client } = require('@opensearch-project/opensearch');
const { AwsSigv4Signer } = require('@opensearch-project/opensearch/aws');

async function main() {
    // create an opensearch client and use the request-signer
    const client = new Client({
        ...AwsSigv4Signer({
            region: 'us-west-2',
            service: 'aoss',
            getCredentials: () =>
                new Promise((resolve, reject) => {
                    AWS.config.getCredentials((err, credentials) => {
                        if (err) {
                            reject(err);
                        } else {
                            resolve(credentials);
                        }
                    });
                }),
        }),
        node: '' # // serverless collection endpoint
    });

    const index = 'movies';

    // create index if it doesn't already exist
    if (!(await client.indices.exists({ index })).body) {
        console.log((await client.indices.create({
            index
        })).body);
    }

    // add a document to the index
    const document = {
        foo: 'bar'
    };
    const response = await client.index({
        id: '1',
        index: index,
```

```
        body: document,
    });
    console.log(response.body);

    // delete the index
    console.log((await client.indices.delete({ index })).body);
}

main();
```

Logstash

È possibile utilizzare il OpenSearch plug-in Logstash per pubblicare i log nelle raccolte Serverless. OpenSearch

Per utilizzare Logstash per inviare dati a Serverless OpenSearch

1. Installa la versione 2.0.0 o successiva del logstash-output-opensearchplug-in utilizzando Docker o Linux.

   Docker

   Docker ospita il software Logstash OSS con il plug-in di OpenSearch output preinstallato: opensearchproject/ -output-plugin. logstash-oss-with-opensearch Puoi estrarre l'immagine come qualsiasi altra immagine:

   ```
   docker pull opensearchproject/logstash-oss-with-opensearch-output-plugin:latest
   ```

   Linux

   Per prima cosa, se ancora non è stato fatto, installa la versione più recente di Logstash. Quindi, installa la versione 2.0.0 del plug-in di output:

   ```
   cd logstash-8.5.0/
   bin/logstash-plugin install --version 2.0.0 logstash-output-opensearch
   ```

   Se il plug-in è già installato, aggiornalo alla versione più recente:

   ```
   bin/logstash-plugin update logstash-output-opensearch
   ```

A partire dalla versione 2.0.0 del plugin, l'SDK utilizza la versione 3. AWS Se utilizzi una versione di Logstash precedente alla 8.4.0, devi rimuovere tutti i plugin AWS preinstallati e installare il plug-in: `logstash-integration-aws`

```
/usr/share/logstash/bin/logstash-plugin remove logstash-input-s3
/usr/share/logstash/bin/logstash-plugin remove logstash-input-sqs
/usr/share/logstash/bin/logstash-plugin remove logstash-output-s3
/usr/share/logstash/bin/logstash-plugin remove logstash-output-sns
/usr/share/logstash/bin/logstash-plugin remove logstash-output-sqs
/usr/share/logstash/bin/logstash-plugin remove logstash-output-cloudwatch

/usr/share/logstash/bin/logstash-plugin install --version 0.1.0.pre logstash-
integration-aws
```

2. Affinché il plugin OpenSearch di output funzioni con OpenSearch Serverless, devi apportare le seguenti modifiche alla sezione di output di logstash.conf: `opensearch`

   - Specifica `aoss` come `service_name` in `auth_type`.
   - Specifica l'endpoint di raccolta per `hosts`.
   - Aggiungi i parametri `default_server_major_version` e `legacy_template`. Questi parametri sono necessari per il funzionamento del plugin con Serverless. OpenSearch

```
output {
  opensearch {
    hosts => "collection-endpoint:443"
    auth_type => {
      ...
      service_name => 'aoss'
    }
    default_server_major_version => 2
    legacy_template => false
  }
}
```

Questo file di configurazione di esempio prende l'input dai file in un bucket S3 e li invia a una OpenSearch raccolta Serverless:

```
input {
  s3  {
```

```
    bucket => "my-s3-bucket"
    region => "us-east-1"
  }
}

output {
  opensearch {
    ecs_compatibility => disabled
    hosts => "https://my-collection-endpoint.us-east-1.aoss.amazonaws.com:443"
    index => my-index
    auth_type => {
      type => 'aws_iam'
      aws_access_key_id => 'your-access-key'
      aws_secret_access_key => 'your-secret-key'
      region => 'us-east-1'
      service_name => 'aoss'
    }
    default_server_major_version => 2
    legacy_template => false
  }
}
```

3. Quindi, esegui Logstash con la nuova configurazione per testare il plug-in:

```
bin/logstash -f config/test-plugin.conf
```

Python

Il codice di esempio seguente utilizza il client opensearch-py per Python per stabilire una connessione sicura alla raccolta OpenSearch Serverless specificata, creare un singolo indice e cercare nell'indice. È necessario fornire valori per `region` e `host`.

La differenza importante rispetto ai domini di OpenSearch servizio è il nome del servizio (anziché). `aoss` `es`

```
from opensearchpy import OpenSearch, RequestsHttpConnection, AWSV4SignerAuth
import boto3

host = ''  # serverless collection endpoint, without https://
region = ''  # e.g. us-east-1

service = 'aoss'
```

```
credentials = boto3.Session().get_credentials()
auth = AWSV4SignerAuth(credentials, region, service)

# create an opensearch client and use the request-signer
client = OpenSearch(
    hosts=[{'host': host, 'port': 443}],
    http_auth=auth,
    use_ssl=True,
    verify_certs=True,
    connection_class=RequestsHttpConnection,
    pool_maxsize=20,
)

# create an index
index_name = 'books-index'
create_response = client.indices.create(
    index_name
)

print('\nCreating index:')
print(create_response)

# index a document
document = {
  'title': 'The Green Mile',
  'director': 'Stephen King',
  'year': '1996'
}

response = client.index(
    index = 'books-index',
    body = document,
    id = '1'
)


# delete the index
delete_response = client.indices.delete(
    index_name
)

print('\nDeleting index:')
print(delete_response)
```

Ruby

La `opensearch-aws-sigv4` gemma fornisce l'accesso OpenSearch immediato a Serverless, insieme a OpenSearch Service. Ha tutte le funzionalità del client opensearch-ruby perché è una dipendenza di questo pacchetto gem.

Quando si crea un'istanza del firmatario Sigv4, specifica `aoss` come nome del servizio:

```
require 'opensearch-aws-sigv4'
require 'aws-sigv4'

signer = Aws::Sigv4::Signer.new(service: 'aoss',
                                region: 'us-west-2',
                                access_key_id: 'key_id',
                                secret_access_key: 'secret')

# create an opensearch client and use the request-signer
client = OpenSearch::Aws::Sigv4Client.new(
  { host: 'https://your.amz-opensearch-serverless.endpoint',
    log: true },
  signer)

# create an index
index = 'prime'
client.indices.create(index: index)

# insert data
client.index(index: index, id: '1', body: { name: 'Amazon Echo',
                                            msrp: '5999',
                                            year: 2011 })

# query the index
client.search(body: { query: { match: { name: 'Echo' } } })

# delete index entry
client.delete(index: index, id: '1')

# delete the index
client.indices.delete(index: index)
```

Firma delle richieste HTTP con altri client

I seguenti requisiti si applicano quando si firmano le richieste nelle raccolte OpenSearch Serverless quando si creano richieste HTTP con altri client.

- È necessario specificare il nome del servizio come `aoss`.
- L'intestazione `x-amz-content-sha256` è obbligatoria per tutte le richieste di AWS Signature Version 4. Fornisce un hash del payload di richiesta. Se è presente un payload di richiesta, impostate il valore sull'hash crittografico Secure Hash Algorithm (SHA) (). SHA256 Se non c'è alcun payload di richiesta, imposta il valore su `e3b0c44298fc1c149afbf4c8996fb92427ae41e4649b934ca495991b7852b855`, che è l'hash di una stringa vuota.

Argomenti

- Indicizzazione con cURL
- Indicizzazione con Postman

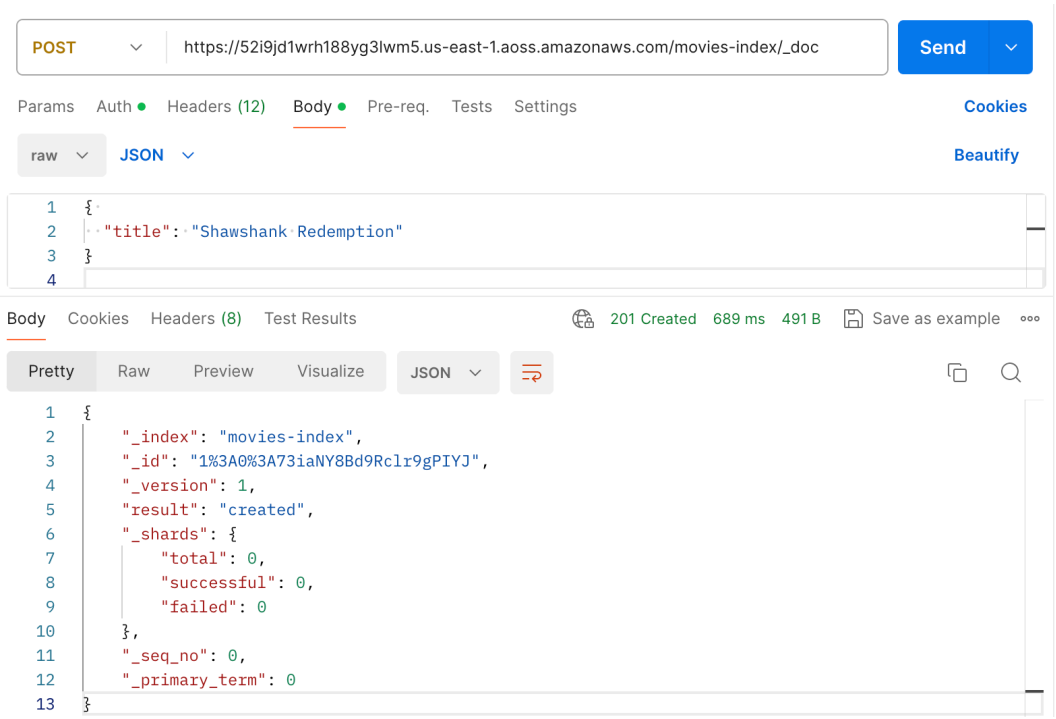
Indicizzazione con cURL

La seguente richiesta di esempio utilizza la Client URL Request Library (cURL) per inviare un singolo documento a un indice denominato `movies-index` all'interno di una raccolta:

```
curl -XPOST \
    --user "$AWS_ACCESS_KEY_ID":"$AWS_SECRET_ACCESS_KEY" \
    --aws-sigv4 "aws:amz:us-east-1:aoss" \
    --header "x-amz-content-sha256: $REQUEST_PAYLOAD_SHA_HASH" \
    --header "x-amz-security-token: $AWS_SESSION_TOKEN" \
    "https://my-collection-endpoint.us-east-1.aoss.amazonaws.com/movies-index/_doc" \
    -H "Content-Type: application/json" -d '{"title": "Shawshank Redemption"}'
```

Indicizzazione con Postman

L'immagine seguente mostra come inviare una richiesta a una raccolta utilizzando Postman. Per istruzioni sull'autenticazione, consulta il flusso di lavoro di autenticazione Authenticate with AWS Signature in Postman.



Configurazione del Machine Learning su Amazon OpenSearch Serverless

Machine Learning

Machine Learning (ML) fornisce funzionalità ML sotto forma di algoritmi ML e modelli remoti. Con l'accesso a questi modelli, puoi eseguire diversi flussi di lavoro AI come RAG o la ricerca semantica. Il machine learning supporta la sperimentazione e l'implementazione in produzione di casi d'uso generativi di intelligenza artificiale utilizzando i più recenti modelli ospitati esternamente che è possibile configurare con connettori. Dopo aver configurato un connettore, è necessario configurarlo su un modello e quindi distribuirlo per eseguire la previsione.

Connettori

I connettori facilitano l'accesso ai modelli ospitati su piattaforme ML di terze parti. Servono da gateway tra il OpenSearch cluster e un modello remoto. Per ulteriori informazioni, consulta la seguente documentazione:

- Creazione di connettori per piattaforme ML di terze parti sul sito Web di OpenSearch documentazione
- Connettori per piattaforme esterne

- Connettori per Servizi AWS

> **Important**
>
> - Quando crei una politica di fiducia, **`ml.opensearchservice.amazonaws.com`** aggiungila come principio del OpenSearch servizio.
> - Salta i passaggi nella pagina Connettori che mostrano come configurare un dominio nella policy.
> - Aggiungi l'`iam:PassRole`istruzione nel passaggio Configura le autorizzazioni.
> - Salta il passaggio Maps the ML role in OpenSearch Dashboards. La configurazione del ruolo di backend non è richiesta. Questo vale per i connettori per Servizi AWS e per i connettori per piattaforme esterne.
> - Nella richiesta SigV4 all'endpoint di raccolta, imposta il nome del servizio su invece di**`aoss`**. **`es`**

Modelli

Un modello è la funzionalità principale utilizzata in vari flussi di lavoro di intelligenza artificiale. In genere, si associa il connettore a un modello per eseguire previsioni utilizzando il connettore. Dopo che un modello è nello stato di distribuzione, è possibile eseguire la previsione. Per ulteriori informazioni, consulta Registrare un modello ospitato su una piattaforma di terze parti nel sito Web della OpenSearch documentazione.

> **Note**
>
> Non tutte le funzionalità del modello sono supportate su OpenSearch Serverless, ad esempio i modelli locali. Per ulteriori informazioni, consulta Machine Learning APIs e funzionalità non supportati.

Configurare le autorizzazioni per il Machine Learning

La sezione seguente descrive le politiche di accesso ai dati di raccolta richieste per il Machine Learning (ML). *`placeholder values`*Sostituiscili con le tue informazioni specifiche. Per ulteriori informazioni, consulta Autorizzazioni delle policy supportate.

```
{
    "Rules": [
        {
            "Resource": [
                "model/collection_name/*"
            ],
            "Permission": [
                "aoss:DescribeMLResource",
                "aoss:CreateMLResource",
                "aoss:UpdateMLResource",
                "aoss:DeleteMLResource",
                "aoss:ExecuteMLResource"
            ],
            "ResourceType": "model"
        }
    ],
    "Principal": [
        "arn:aws:iam::account_id:role/role_name"
    ],
    "Description": "ML full access policy for collection_name"
}
```

- aoss:Describe MLResource — Concede il permesso di cercare e interrogare connettori, modelli e gruppi di modelli.
- aOSS:Create MLResource — Concede l'autorizzazione a creare connettori, modelli e gruppi di modelli.
- aOSS:Update MLResource — Concede l'autorizzazione ad aggiornare connettori, modelli e gruppi di modelli.
- aOSS:Delete MLResource — Concede il permesso di eliminare connettori, modelli e gruppi di modelli.
- aOSS:Execute MLResource — Concede l'autorizzazione a eseguire previsioni sui modelli.

Machine Learning APIs e funzionalità non supportati

Non supportata APIs

I seguenti sistemi di Machine Learning (ML) non APIs sono supportati su Amazon OpenSearch Serverless:

- Funzionalità del modello locale
- API Model Train
- API dell'algoritmo Model Predict
- Modello Batch Predict API
- Agents API e relativi strumenti
- Server MCP APIs
- Memoria APIs
- Controller APIs
- Esegui l'API dell'algoritmo
- Profilo ML AP
- API di statistiche ML

Per ulteriori informazioni su ML APIs, consulta ML APIs sul sito Web della OpenSearch documentazione.

Caratteristiche non supportate

Le seguenti funzionalità ML non sono supportate su Amazon OpenSearch Serverless:

- Agenti e strumenti
- Modelli locali
- Il processore di inferenza ML all'interno di Search and Ingest Pipelines
  - Processore ML Inference Ingest
  - Processore di risposta alla ricerca di inferenza ML
  - Processore di richiesta di ricerca inferenziale ML

Per ulteriori informazioni su queste funzionalità, vedere la seguente documentazione sul sito Web della OpenSearch documentazione:

- Machine learning
- Processore di inferenza ML
- Pipeline di ricerca

Configura la ricerca neurale e la ricerca ibrida su Serverless OpenSearch

Ricerca neurale

Amazon OpenSearch Serverless supporta la funzionalità di ricerca neurale per operazioni di ricerca semantica sui tuoi dati. Neural Search utilizza modelli di apprendimento automatico per comprendere il significato semantico e il contesto delle tue query, fornendo risultati di ricerca più pertinenti rispetto alle tradizionali ricerche basate su parole chiave. Questa sezione spiega come configurare Neural Search in OpenSearch Serverless, incluse le autorizzazioni richieste, i processori supportati e le differenze principali rispetto all'implementazione standard. OpenSearch

Con Neural Search puoi eseguire ricerche semantiche sui tuoi dati, che prendono in considerazione il significato semantico per comprendere l'intento delle tue query di ricerca. Questa funzionalità è alimentata dai seguenti componenti:

- Processore di pipeline di inserimento del testo
- Interrogazione neurale
- Interrogazione neurale sparsa

Ricerca ibrida

Con la ricerca ibrida, puoi migliorare la pertinenza della ricerca combinando funzionalità di ricerca semantica e parole chiave. Per utilizzare la ricerca ibrida, crea una pipeline di ricerca che elabori i risultati della ricerca e combini i punteggi dei documenti. Per ulteriori informazioni, consulta Search pipeline sul sito Web della OpenSearch documentazione. Utilizzate i seguenti componenti per implementare la ricerca ibrida:

- Processore di pipeline di ricerca per la normalizzazione

  Tecniche di normalizzazione supportate

  - `min_max`

- `l2`

  Tecniche di combinazione supportate

  - `arithmetic_mean`
  - `geometric_mean`
  - `harmonic_mean`

  Per ulteriori informazioni sulle tecniche di normalizzazione e combinazione, consulta Request body fields sul sito Web della OpenSearchdocumentazione.

- Query ibrida

Interrogazioni neurali e ibride

Per impostazione predefinita, OpenSearch calcola i punteggi dei documenti utilizzando l' BM25algoritmo Okapi basato su parole chiave, che funziona bene per le query di ricerca che contengono parole chiave. Neural Search offre nuovi tipi di query per le query in linguaggio naturale e la possibilità di combinare la ricerca semantica con quella per parole chiave.

Example: neural

```
"neural": {
  "vector_field": {
    "query_text": "query_text",
    "query_image": "image_binary",
    "model_id": "model_id",
    "k": 100
  }
}
```

Per ulteriori informazioni, consulta Neural query sul sito Web della documentazione. OpenSearch

Example: hybrid

```
"hybrid": {
      "queries": [
        array of lexical, neural, or combined queries
      ]
    }
```

Per ulteriori informazioni, vedere Hybrid query sul sito Web della OpenSearch documentazione.

Per configurare i componenti di ricerca semantica in Amazon OpenSearch Serverless, segui i passaggi del tutorial Neural Search sul sito Web della OpenSearch documentazione. Tieni a mente queste importanti differenze:

- OpenSearch Serverless supporta solo modelli remoti. È necessario configurare i connettori per i modelli ospitati in remoto. Non è necessario distribuire o rimuovere modelli remoti. Per ulteriori informazioni, consulta Guida introduttiva alla ricerca semantica e ibrida sul sito Web della OpenSearch documentazione.
- Aspettati fino a 15 secondi di latenza quando esegui una ricerca nell'indice vettoriale o cerchi pipeline di ricerca e inserimento create di recente.

Configurazione delle autorizzazioni

La ricerca neurale in OpenSearch Serverless richiede le seguenti autorizzazioni. Per ulteriori informazioni, consulta Autorizzazioni delle policy supportate.

Example: Politica di ricerca neurale

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "NeuralSearch",
            "Effect": "Allow",
            "Action": [
              "aoss:CreateIndex",
              "aoss:CreateCollection",
              "aoss:UpdateCollection",
              "aoss:DeleteIndex",
              "aoss:DeleteCollection"
            ],
            "Resource": "arn:aws:aoss:us-east-1:111122223333:collection/your-collection-name"
        }
    ]
}
```

- aoss: *Index — Crea un indice vettoriale in cui vengono memorizzati gli incorporamenti di testo.
- aoss: * CollectionItems — Crea pipeline di importazione e ricerca.
- aoss: * MLResource — Crea e registra modelli di incorporamento del testo.
- aoss: APIAccess All — Fornisce l'accesso a OpenSearch APIs per le operazioni di ricerca e inserimento.

Di seguito vengono descritte le politiche di accesso ai dati di raccolta necessarie per la ricerca neurale. Sostituiscili *placeholder values* con le tue informazioni specifiche.

```
            {
                "ResourceType": "index",
                "Resource": ["index/collection_name/*"],
                "Permission": [
                  "aoss:CreateIndex",
                  "aoss:DescribeIndex",
                  "aoss:UpdateIndex",
                  "aoss:DeleteIndex"
                ]
            }
        ],
        "Principal": [
            "arn:aws:iam::account_id:role/role_name"
        ]
    },
    {
        "Description": "Create pipeline permission",
        "Rules": [
            {
                "ResourceType": "collection",
                "Resource": ["collection/collection_name"],
                "Permission": [
                  "aoss:CreateCollectionItems",
                  "aoss:DescribeCollectionItems",
                  "aoss:UpdateCollectionItems",
                  "aoss:DeleteCollectionItems"
                ]
            }
        ],
        "Principal": [
            "arn:aws:iam::account_id:role/role_name"
        ]
    },
    {
        "Description": "Create model permission",
        "Rules": [
            {
                "ResourceType": "model",
                "Resource": ["model/collection_name/*"],
                "Permission": ["aoss:CreateMLResources"]
            }
        ],
        "Principal": [
            "arn:aws:iam::account_id:role/role_name"
        ]
    }
]
```

Configurazione dei flussi di lavoro su Amazon Serverless OpenSearch

Flussi di lavoro

I flussi di lavoro supportano i costruttori nell'innovazione delle applicazioni di intelligenza artificiale su. OpenSearch L'attuale processo di utilizzo delle offerte di machine learning (ML) OpenSearch, come la ricerca semantica, richiede attività complesse di configurazione e pre-elaborazione, oltre a domande dettagliate degli utenti, entrambe operazioni che possono richiedere molto tempo e sono soggette a errori. I flussi di lavoro sono un framework di semplificazione per cui concatenare più chiamate API. OpenSearch

Per la configurazione e l'utilizzo, consulta Automazione delle configurazioni sul sito Web. OpenSearch Quando utilizzi Workflows in OpenSearch Serverless, considera queste importanti differenze:

- OpenSearch Serverless utilizza solo modelli remoti nelle fasi del flusso di lavoro. Non è necessario implementare questi modelli.
- OpenSearch Serverless non supporta la fase Re-index Workflow.
- Quando cerchi gli stati di Workflows e Workflow dopo altre chiamate API, aspettati fino a 15 secondi di latenza per la visualizzazione degli aggiornamenti.

OpenSearch Le raccolte Serverless supportano i flussi di lavoro quando vengono utilizzate come fonte di dati nell'applicazione dell'interfaccia utente. OpenSearch Per ulteriori informazioni, consulta Gestione delle associazioni delle fonti di dati.

Configurazione delle autorizzazioni

Prima di creare e fornire un modello, verifica di disporre delle autorizzazioni necessarie. Se hai bisogno di assistenza, contatta l'amministratore del tuo account. OpenSearch I flussi di lavoro serverless richiedono le seguenti autorizzazioni. Puoi definire l'ambito delle autorizzazioni per una raccolta specifica definendo l'ARN della risorsa di raccolta nella tua policy IAM.

Example: politica sui flussi di lavoro

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Sid": "NeuralSearch",
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "arn:aws:iam::111122223333:role/Cognito_identitypoolname/Auth_Role"
        ]
      },
      "Action": [
        "aoss:CreateIndex",
        "aoss:CreateCollection",
        "aoss:UpdateCollection",
        "aoss:DeleteIndex",
        "aoss:DeleteCollection"
      ],
      "Resource": "arn:aws:aoss:us-east-1:111122223333:collection/your-
collection-name"
    }
  ]
}
```

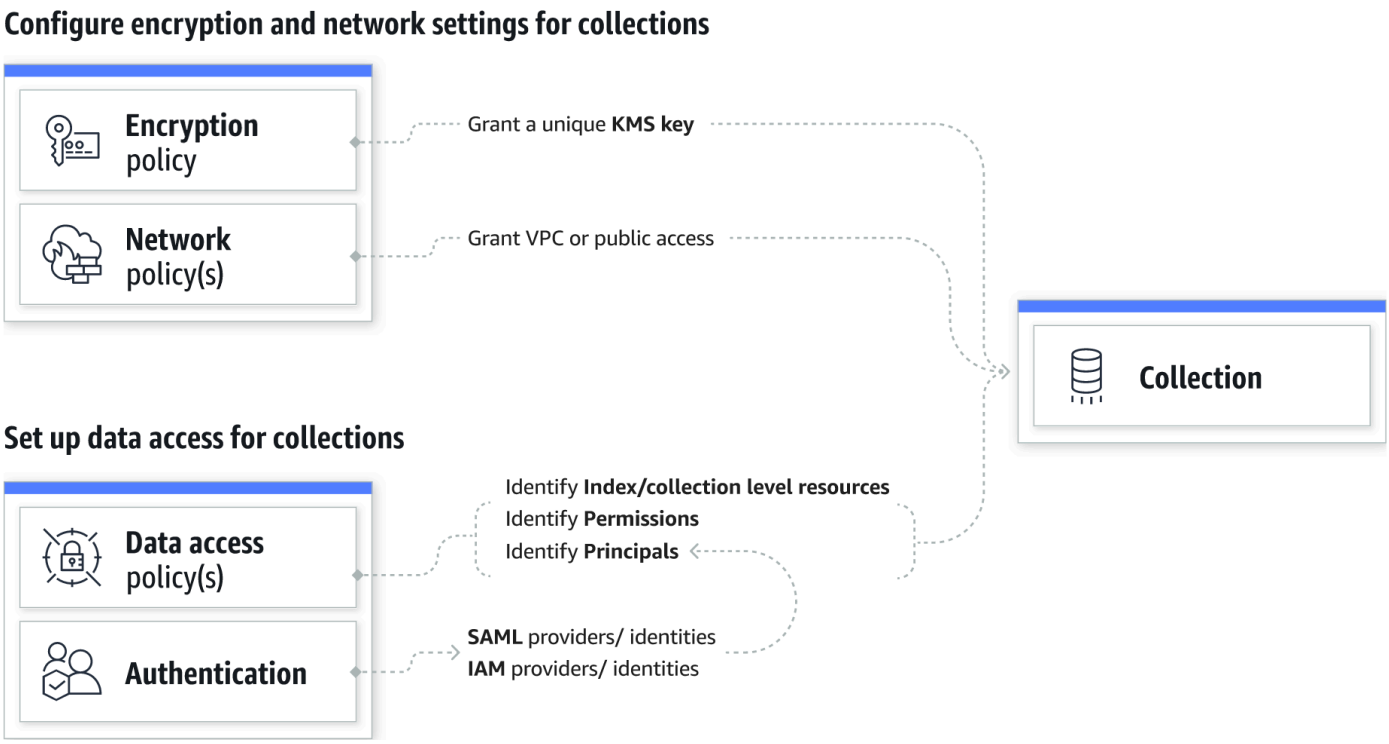
- aoss: * CollectionItems — Concede l'autorizzazione a creare e gestire modelli e fornisce pipeline di ricerca e importazione.
- aoss: *Index — Concede l'autorizzazione a creare ed eliminare indici utilizzando le operazioni API. OpenSearch
- aoss: * MLResource — Concede l'autorizzazione per eseguire il provisioning delle fasi del flusso di lavoro che utilizzano Configure Machine Learning.

Panoramica della sicurezza in Amazon OpenSearch Serverless

La sicurezza in Amazon OpenSearch Serverless si differenzia fondamentalmente dalla sicurezza in Amazon OpenSearch Service nei seguenti modi:

Funzionalità	OpenSearch Servizio	OpenSearch Senza server
Controllo dell'accesso ai dati	L'accesso ai dati è determinato dalle policy IAM e dal controllo granulare degli accessi.	L'accesso ai dati è determinato dalle policy di accesso ai dati.
Crittografia dei dati inattivi	La crittografia dei dati inattivi è facoltativa per i domini.	La crittografia dei dati inattivi è obbligatoria per le raccolte.
Configurazione e amministr azione della sicurezza	È necessario configurare la rete, la crittografia e l'accesso ai dati singolarm ente per ogni dominio.	Puoi utilizzare le policy di sicurezza per gestire le impostazioni di sicurezza per più raccolte su larga scala.

Il diagramma seguente illustra i componenti di sicurezza che costituiscono una raccolta funzionale. Una raccolta deve avere una chiave di crittografia assegnata, impostazioni di accesso alla rete e una policy di accesso ai dati corrispondente che garantisca l'autorizzazione alle relative risorse.



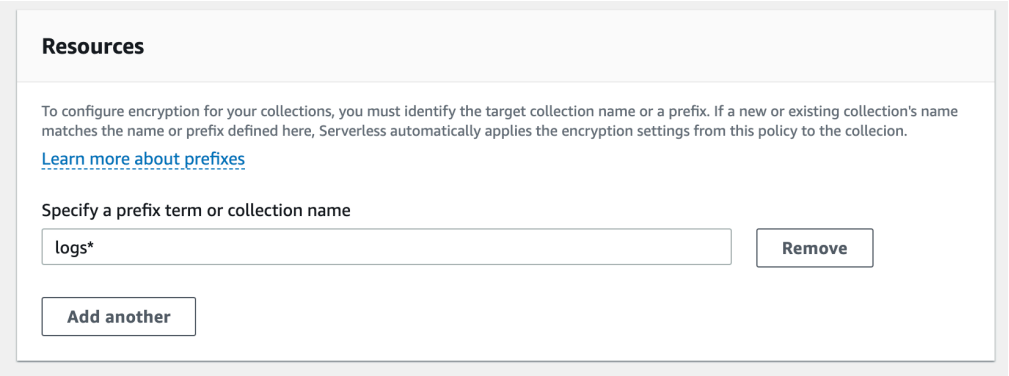
Argomenti

- Policy di crittografia
- Policy di rete
- Policy di accesso ai dati
- Autenticazione IAM e SAML
- Sicurezza dell'infrastruttura
- Guida introduttiva alla sicurezza in Amazon OpenSearch Serverless
- Identity and Access Management per Amazon OpenSearch Serverless
- Crittografia in Amazon OpenSearch Serverless
- Accesso alla rete per Amazon OpenSearch Serverless
- Conformità FIPS in Amazon Serverless OpenSearch
- Controllo dell'accesso ai dati per Amazon OpenSearch Serverless
- Accesso al piano dati tramite AWS PrivateLink
- Controlla l'accesso al piano tramite AWS PrivateLink
- Autenticazione SAML per Amazon Serverless OpenSearch
- Convalida della conformità per Amazon Serverless OpenSearch

Policy di crittografia

Le politiche di crittografia definiscono se le raccolte sono crittografate con una chiave Chiave di proprietà di AWS o con una chiave gestita dal cliente. Le policy di crittografia sono costituite da due componenti: un modello di risorse e una chiave di crittografia. Il modello di risorse definisce a quale raccolta o raccolte si applica la policy. La chiave di crittografia determina il modo in cui verranno protette le raccolte associate.

Per applicare una policy a più raccolte, includi un carattere jolly (*) nella regola della policy. Ad esempio, la seguente policy si applica a tutte le raccolte i cui nomi iniziano con "log".

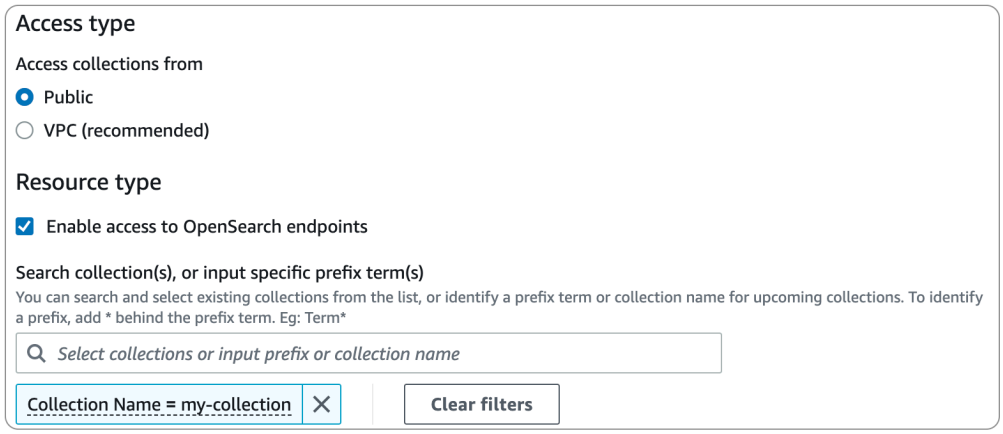


Le policy di crittografia semplificano il processo di creazione e gestione delle raccolte, soprattutto quando lo si fa a livello di programmazione. È possibile creare una raccolta specificando un nome e una chiave di crittografia viene assegnata automaticamente alla raccolta al momento della creazione.

Policy di rete

I criteri di rete definiscono se le raccolte sono accessibili privatamente o tramite Internet da reti pubbliche. È possibile accedere alle raccolte private tramite endpoint OpenSearch VPC gestiti senza server o tramite dispositivi specifici Servizi AWS come Amazon Bedrock che utilizzano l'accesso privato.Servizio AWS Proprio come le policy di crittografia, le policy di rete possono essere applicate a più raccolte e questo consente di gestire l'accesso alla rete per molte raccolte su larga scala.

Le policy di rete sono costituite da due componenti: un tipo di accesso e un tipo di risorsa. Il tipo di accesso può essere pubblico o privato. Il tipo di risorsa determina se l'accesso scelto si applica all'endpoint di raccolta, all'endpoint OpenSearch Dashboards o a entrambi.



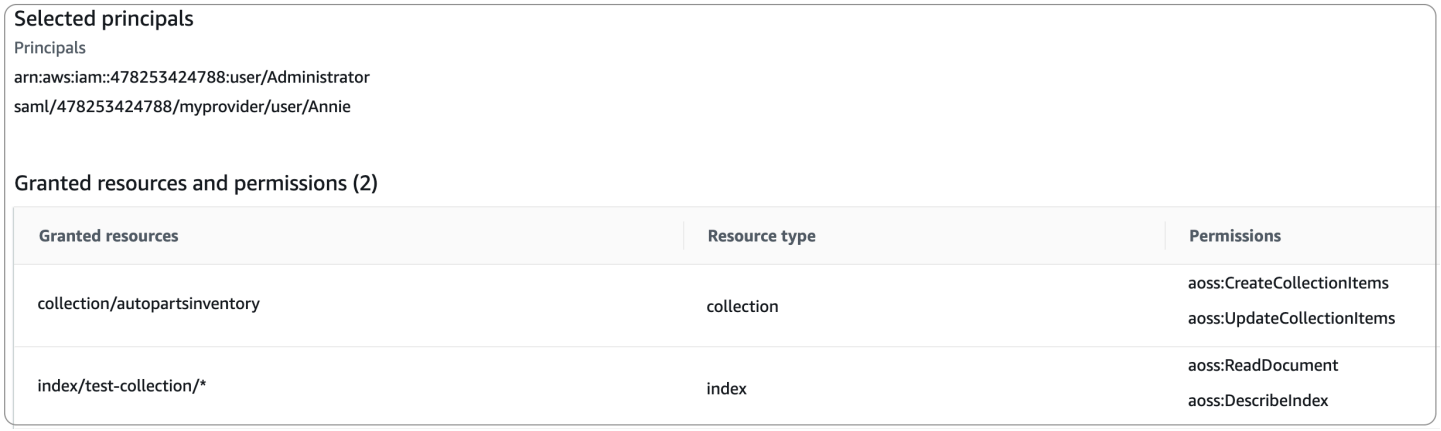
Se prevedi di configurare l'accesso VPC all'interno di una policy di rete, devi prima creare uno o più endpoint VPC gestiti OpenSearch senza server. Questi endpoint ti consentono di accedere a OpenSearch Serverless come se fosse nel tuo VPC, senza l'uso di un gateway Internet, un dispositivo NAT, una connessione VPN o una connessione. Direct Connect

L'accesso privato a Servizi AWS può essere applicato solo all'endpoint della raccolta, non all' OpenSearchendpoint Dashboards. OpenSearch Servizi AWS non può essere concesso l'accesso alle dashboard. OpenSearch

Policy di accesso ai dati

Le policy di accesso ai dati definiscono il modo in cui gli utenti accedono ai dati all'interno delle raccolte. Le policy di accesso ai dati consentono di gestire le raccolte su larga scala assegnando automaticamente autorizzazioni di accesso a raccolte e indici che corrispondono a uno schema specifico. È possibile applicare più policy a una singola risorsa.

Le policy di accesso ai dati sono costituite da un insieme di regole, ciascuna con tre componenti: un tipo di risorsa, le risorse concesse e un elenco delle autorizzazioni. Il tipo di risorsa può essere una raccolta o un indice. Le risorse concesse possono essere collection/index nomi o pattern con un carattere jolly (*). L'elenco delle autorizzazioni specifica a quali operazioni OpenSearch API la policy concede l'accesso. Inoltre, la policy contiene un elenco di principali, che specificano i ruoli IAM, gli utenti e le identità SAML a cui concedere l'accesso.



Per ulteriori informazioni sul formato di una policy di accesso ai dati, consulta la sintassi della policy.

Prima di creare una policy di accesso ai dati, è necessario disporre di uno o più utenti o ruoli IAM, o identità SAML, a cui fornire l'accesso nella policy. Per ulteriori informazioni, consulta la prossima sezione.

> **Note**
>
> Il passaggio dall'accesso pubblico a quello privato per la raccolta rimuoverà la scheda Indici nella console Serverless Collection. OpenSearch

Autenticazione IAM e SAML

I principali IAM e le identità SAML sono uno degli elementi costitutivi di una policy di accesso ai dati. All'interno dell'istruzione `principal` di una policy di accesso, è possibile includere ruoli IAM, utenti e identità SAML. A questi principali vengono quindi concesse le autorizzazioni specificate nelle regole delle policy associate.

```
[
   {
      "Rules":[
         {
            "ResourceType":"index",
            "Resource":[
               "index/marketing/orders*"
            ],
            "Permission":[
               "aoss:*"
            ]
```

```
         }
      ],
      "Principal":[
         "arn:aws:iam::123456789012:user/Dale",
         "arn:aws:iam::123456789012:role/RegulatoryCompliance",
         "saml/123456789012/myprovider/user/Annie"
      ]
   }
]
```

L'autenticazione SAML viene configurata direttamente all'interno di Serverless. OpenSearch Per ulteriori informazioni, consulta the section called “Autenticazione SAML”.

Sicurezza dell’infrastruttura

Amazon OpenSearch Serverless è protetto dalla sicurezza di rete AWS globale. Per informazioni sui servizi di AWS sicurezza e su come AWS protegge l'infrastruttura, consulta AWS Cloud Security. Per progettare il tuo AWS ambiente utilizzando le migliori pratiche per la sicurezza dell'infrastruttura, vedi Infrastructure Protection in Security Pillar AWS Well-Architected Framework.

Utilizzi chiamate API AWS pubblicate per accedere ad Amazon OpenSearch Serverless attraverso la rete. I client devono supportare Transport Layer Security (TLS). È richiesto TLS 1.2 ed è consigliato TLS 1.3. Per un elenco dei codici supportati per TLS 1.3, consulta i protocolli e le cifrari TLS nella documentazione di Elastic Load Balancing.

Inoltre, è necessario firmare le richieste utilizzando un ID di chiave di accesso e una chiave di accesso segreta associata a un'entità IAM. In alternativa, è possibile utilizzare AWS Security Token Service (AWS STS) per generare le credenziali di sicurezza temporanee per sottoscrivere le richieste.

Guida introduttiva alla sicurezza in Amazon OpenSearch Serverless

I seguenti tutorial ti aiutano a iniziare a usare Amazon OpenSearch Serverless. Entrambi i tutorial completano le stesse fasi di base, ma uno utilizza la console mentre l'altro utilizza la AWS CLI.

Tieni presente che i casi d'uso in questi tutorial sono semplificati. Le policy di rete e di sicurezza sono relativamente aperte. Nei carichi di lavoro di produzione, si consiglia di configurare funzionalità di sicurezza più affidabili come l'autenticazione SAML, l'accesso al VPC e le policy restrittive di accesso ai dati.

Argomenti

- Tutorial: Guida introduttiva alla sicurezza in Amazon OpenSearch Serverless (console)
- Tutorial: Guida introduttiva alla sicurezza in Amazon OpenSearch Serverless (CLI)

Tutorial: Guida introduttiva alla sicurezza in Amazon OpenSearch Serverless (console)

Questo tutorial illustra i passaggi di base per creare e gestire le policy di sicurezza utilizzando la console Amazon OpenSearch Serverless.

In questo tutorial completerai le seguenti fasi:

1. Configurazione delle autorizzazioni
2. Creazione di una policy di crittografia
3. Creazione di una policy di rete
4. Configurazione di una policy di accesso ai dati
5. Creazione di una raccolta
6. Caricamento e ricerca dei dati

Questo tutorial illustra come configurare una raccolta utilizzando. Console di gestione AWS Per gli stessi passaggi utilizzando il AWS CLI, vederethe section called “Tutorial: nozioni di base sulla sicurezza (CLI)”.

Fase 1: configurazione delle autorizzazioni

> **Note**
>
> Puoi saltare questa fase se stai già utilizzando una policy più ampia basata sull'identità, ad esempio `Action":"aoss:*"` o `Action":"*"`. Negli ambienti di produzione, tuttavia, si consiglia di seguire il principio del privilegio minimo e di assegnare solo le autorizzazioni minime necessarie per completare un'attività.

Per completare questo tutorial è necessario disporre delle autorizzazioni IAM corrette. L'utente o il ruolo devono avere una policy basata sull'identità allegata con le seguenti autorizzazioni minime:

JSON

```
{
```

```
  "Version":"2012-10-17",
  "Statement": [
    {
      "Action": [
        "aoss:ListCollections",
        "aoss:BatchGetCollection",
        "aoss:CreateCollection",
        "aoss:CreateSecurityPolicy",
        "aoss:GetSecurityPolicy",
        "aoss:ListSecurityPolicies",
        "aoss:CreateAccessPolicy",
        "aoss:GetAccessPolicy",
        "aoss:ListAccessPolicies"
      ],
      "Effect": "Allow",
      "Resource": "*"
    }
  ]
}
```

Per un elenco completo delle autorizzazioni OpenSearch Serverless, consulta. the section called "Identity and Access Management"

Fase 2: creazione di una policy di crittografia

Le politiche di crittografia specificano la AWS KMS chiave che OpenSearch Serverless utilizzerà per crittografare la raccolta. È possibile crittografare le raccolte con una Chiave gestita da AWS o un'altra chiave. Per semplicità, in questo tutorial crittograferemo la nostra raccolta con una Chiave gestita da AWS.

Creazione di una policy di crittografia

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Espandi Serverless nel pannello di navigazione a sinistra e scegli Encryption policies (Policy di crittografia).
3. Scegli Create encryption policy (Crea policy di crittografia).
4. Nomina la policy books-policy. Nella descrizione, inserisci Crittografia per una raccolta di libri.
5. In Resources (Risorse), inserisci books (libri), che è il nome che darai alla tua raccolta. Se si desidera essere più generici, è possibile includere un asterisco (`books*`) per applicare la policy a tutte le raccolte che iniziano con la parola "books".

6. Per la crittografia, mantieni selezionata l'opzione Usa chiave AWS di proprietà.
7. Scegli Create (Crea).

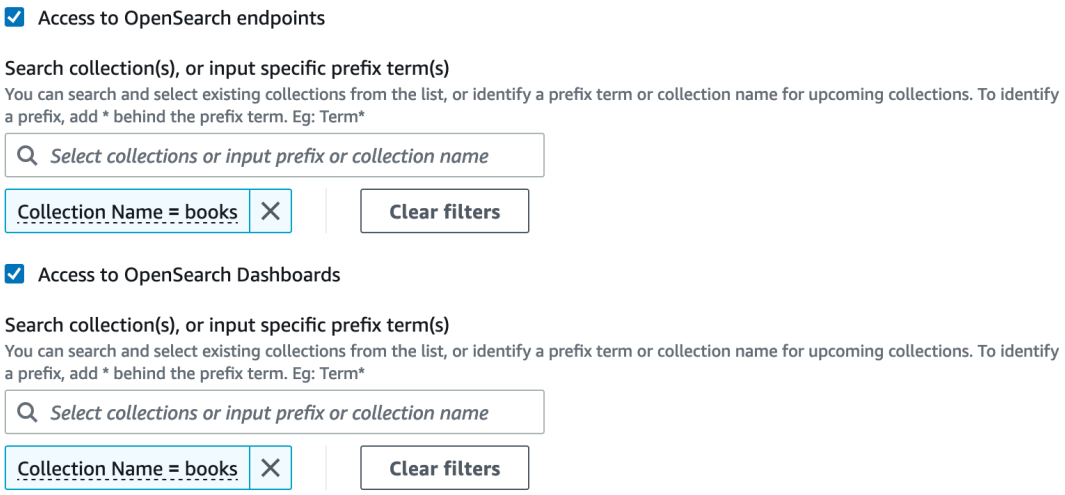
Fase 3: Creare una politica di rete

Le politiche di rete determinano se la raccolta è accessibile su Internet dalle reti pubbliche o se è necessario accedervi tramite endpoint VPC OpenSearch gestiti senza server. In questo tutorial, configureremo l'accesso pubblico.

Creazione di una policy di rete

1. Nel pannello di navigazione a sinistra, scegli Network policies (Policy di rete) e Create network policy (Crea policy di rete).
2. Nomina la policy books-policy. Nella descrizione, inserisci Policy di rete per una raccolta di libri.
3. Nella Regola 1, nomina la regola Accesso pubblico per una raccolta di libri.
4. Per semplicità, in questo tutorial configureremo l'accesso pubblico alla raccolta books. Per il tipo di accesso, seleziona Public (Pubblico).
5. Accederemo alla raccolta da Dashboards. OpenSearch Per fare ciò, devi configurare l'accesso alla rete per le dashboard e l' OpenSearch endpoint, altrimenti le dashboard non funzioneranno.

   Per il tipo di risorsa, abilita sia Access to OpenSearch endpoint che Access to Dashboards. OpenSearch
6. In entrambe le caselle di input, inserisci Collection Name = books (Nome raccolta = libri). Questa impostazione riduce l'ambito della policy in modo che si applichi solo a una singola raccolta (`books`). La tua regola dovrebbe assomigliare a questa:



7. Scegli Create (Crea).

Passaggio 4: Creare una politica di accesso ai dati

I dati della raccolta non saranno accessibili finché non configuri l'accesso ai dati. Le policy di accesso ai dati sono separate dalla policy basata sull'identità IAM configurata nella fase 1. Consentono agli utenti di accedere ai dati effettivi all'interno di una raccolta.

In questo tutorial, forniremo a un singolo utente le autorizzazioni necessarie per indicizzare i dati nella raccolta denominata books.

Creazione di una policy di accesso ai dati

1. Nel pannello di navigazione a sinistra, scegli Data access policies (Policy di accesso ai dati) e Create access policy (Crea policy di accesso).
2. Nomina la policy books-policy. Nella descrizione, inserisci policy di accesso ai dati per la raccolta di libri.
3. Per il metodo di definizione della policy seleziona JSON e incolla la seguente policy nell'editor JSON.

   Sostituisci l'ARN principale con l'ARN dell'account che utilizzerai per accedere alle OpenSearch dashboard e ai dati degli indici.

```
[
   {
      "Rules":[
         {
            "ResourceType":"index",
            "Resource":[
               "index/books/*"
            ],
            "Permission":[
               "aoss:CreateIndex",
               "aoss:DescribeIndex",
               "aoss:ReadDocument",
               "aoss:WriteDocument",
               "aoss:UpdateIndex",
               "aoss:DeleteIndex"
            ]
         }
      ],
```

```
      "Principal":[
         "arn:aws:iam::123456789012:user/my-user"
      ]
   }
]
```

Questa policy fornisce a un singolo utente le autorizzazioni minime necessarie per creare un indice nella raccolta books, indicizzare alcuni dati e cercarli.

4. Scegli Create (Crea).

Passaggio 5: Crea una raccolta

Ora che hai configurato la crittografia e le policy di rete, puoi creare una raccolta corrispondente e le impostazioni di sicurezza verranno applicate automaticamente ad essa.

Per creare una raccolta OpenSearch Serverless

1. Scegli Collections (Raccolte) nel pannello di navigazione a sinistra e scegli Create collection (Crea raccolta).
2. Nomina la raccolta books.
3. Per il tipo di raccolta, scegli Search (Cerca).
4. In Encryption, OpenSearch Serverless informa che il nome della raccolta corrisponde alla `books-policy` politica di crittografia.
5. In Impostazioni di accesso alla rete, OpenSearch Serverless informa che il nome della raccolta corrisponde alla politica di rete. `books-policy`
6. Scegli Next (Successivo).
7. In Opzioni della politica di accesso ai dati, OpenSearch Serverless ti informa che il nome della raccolta corrisponde alla politica di accesso ai `books-policy` dati.
8. Scegli Next (Successivo).
9. Rivedi la configurazione della raccolta e scegli Submit (Invia). Generalmente, l'inizializzazione delle raccolte richiede meno di un minuto.

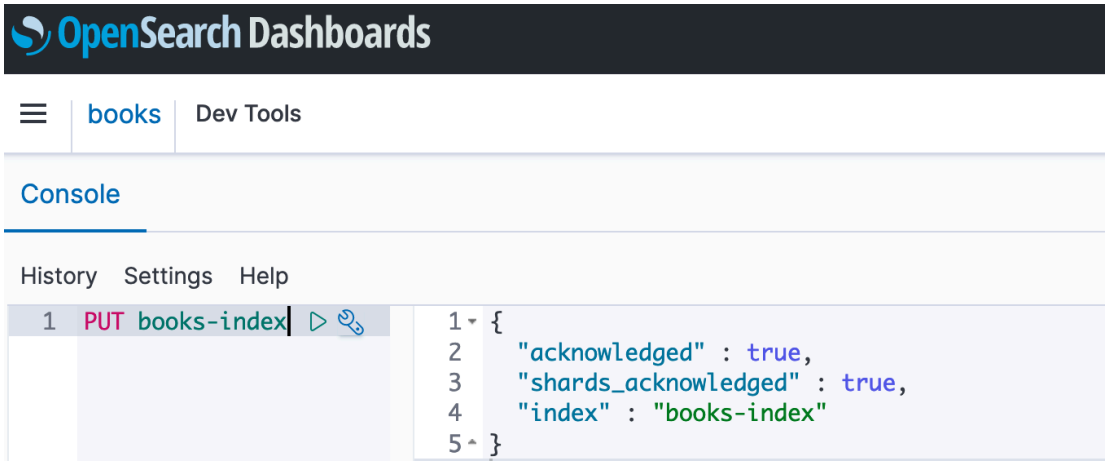
Fase 6: caricamento e ricerca dei dati

Puoi caricare dati in una raccolta OpenSearch Serverless utilizzando Postman o curl. Per brevità, questi esempi utilizzano Dev Tools all'interno della console Dashboards. OpenSearch

Indicizzazione e ricerca di dati in una raccolta

1. Scegli Collections (Raccolte) nel pannello di navigazione a sinistra e scegli la raccolta books per aprirne la pagina dei dettagli.
2. Scegli l'URL delle OpenSearch dashboard per la raccolta. L'URL assume il formato `https://collection-id.us-east-1.aoss.amazonaws.com/_dashboards`.
3. Accedi alle OpenSearch dashboard utilizzando le chiavi di AWS accesso e segrete per il principale che hai specificato nella politica di accesso ai dati.
4. All'interno di OpenSearch Dashboards, apri il menu di navigazione a sinistra e scegli Dev Tools.
5. Per creare un singolo indice chiamato books-index, esegui il seguente comando:

```
PUT books-index
```



6. Per indicizzare un singolo documento in books-index, esegui il seguente comando:

```
PUT books-index/_doc/1
{
  "title": "The Shining",
  "author": "Stephen King",
  "year": 1977
}
```

7. Per cercare dati nelle OpenSearch dashboard, devi configurare almeno un modello di indice. OpenSearch utilizza questi modelli per identificare gli indici da analizzare. Apri il menu principale di Dashboards, scegliere Gestione stack, scegliere Modelli di indice, quindi scegliere Crea modello di indice. Per questo tutorial, inserisci books-index.

8. Scegliere Fase successiva quindi selezionare Crea modello di indice. Dopo aver creato il modello, è possibile visualizzare i vari campi del documento, ad esempio `author` e `title`.
9. Per iniziare a cercare i dati, apri di nuovo il menu principale e scegli Discover (Rileva) o utilizza l'API di ricerca.

Tutorial: Guida introduttiva alla sicurezza in Amazon OpenSearch Serverless (CLI)

Questo tutorial illustra i passaggi descritti nel tutorial introduttivo per la sicurezza della console, ma utilizza la console AWS CLI anziché la console di OpenSearch servizio.

In questo tutorial completerai le seguenti fasi:

1. Crea una politica di autorizzazioni IAM
2. Associa la policy IAM a un ruolo IAM
3. Creare una policy di crittografia
4. Creazione di una policy di rete
5. Creare una raccolta
6. Configurazione di una policy di accesso ai dati
7. Recupera l'endpoint di raccolta
8. Carica i dati sulla tua connessione
9. Cerca i dati nella tua raccolta

L'obiettivo di questo tutorial è configurare un'unica raccolta OpenSearch Serverless con impostazioni di crittografia, rete e accesso ai dati abbastanza semplici. Ad esempio, configureremo l'accesso alla rete pubblica, una Chiave gestita da AWS per la crittografia e una politica di accesso ai dati semplificata che conceda autorizzazioni minime a un singolo utente.

In uno scenario di produzione, ti consigliamo di implementare una configurazione più affidabile, che includa l'autenticazione SAML, una chiave di crittografia personalizzata e l'accesso al VPC.

Per iniziare con le politiche di sicurezza in Serverless OpenSearch

1.

> **Note**
>
> Puoi saltare questa fase se stai già utilizzando una policy più ampia basata sull'identità, ad esempio `Action":"aoss:*"` o `Action":"*"`. Negli ambienti di produzione,

tuttavia, si consiglia di seguire il principio del privilegio minimo e di assegnare solo le autorizzazioni minime necessarie per completare un'attività.

Per iniziare, crea una AWS Identity and Access Management politica con le autorizzazioni minime richieste per eseguire i passaggi di questo tutorial. Denomineremo la policy `TutorialPolicy`:

```
aws iam create-policy \
  --policy-name TutorialPolicy \
  --policy-document "{\"Version\": \"2012-10-17\",\"Statement\":
 [{\"Action\": [\"aoss:ListCollections\",\"aoss:BatchGetCollection\",
\"aoss:CreateCollection\",\"aoss:CreateSecurityPolicy\",\"aoss:GetSecurityPolicy\",
\"aoss:ListSecurityPolicies\",\"aoss:CreateAccessPolicy\",\"aoss:GetAccessPolicy\",
\"aoss:ListAccessPolicies\"],\"Effect\": \"Allow\",\"Resource\": \"*\"}]}"
```

Risposta di esempio

```
{
    "Policy": {
        "PolicyName": "TutorialPolicy",
        "PolicyId": "ANPAW6WRAECKG6QJWUV7U",
        "Arn": "arn:aws:iam::123456789012:policy/TutorialPolicy",
        "Path": "/",
        "DefaultVersionId": "v1",
        "AttachmentCount": 0,
        "PermissionsBoundaryUsageCount": 0,
        "IsAttachable": true,
        "CreateDate": "2022-10-16T20:57:18+00:00",
        "UpdateDate": "2022-10-16T20:57:18+00:00"
    }
}
```

2. Allega la `TutorialPolicy` al ruolo IAM che indicizzerà e cercherà i dati nella raccolta. Denomineremo l'utente `TutorialRole`:

```
aws iam attach-role-policy \
  --role-name TutorialRole \
  --policy-arn arn:aws:iam::123456789012:policy/TutorialPolicy
```

3. Prima di creare una raccolta, è necessario creare una policy di crittografia che assegni una Chiave di proprietà di AWS alla raccolta books che creerai in una fase successiva.

   Invia la seguente richiesta per creare una policy di crittografia per la raccolta books:

```
aws opensearchserverless create-security-policy \
  --name books-policy \
  --type encryption --policy "{\"Rules\":[{\"ResourceType\":\"collection\",
\"Resource\":[\"collection\/books\"]}],\"AWSOwnedKey\":true}"
```

   Risposta di esempio

```
{
    "securityPolicyDetail": {
        "type": "encryption",
        "name": "books-policy",
        "policyVersion": "MTY2OTI0MDAwNTk5MF8x",
        "policy": {
            "Rules": [
                {
                    "Resource": [
                        "collection/books"
                    ],
                    "ResourceType": "collection"
                }
            ],
            "AWSOwnedKey": true
        },
        "createdDate": 1669240005990,
        "lastModifiedDate": 1669240005990
    }
}
```

4. Crea una policy di rete che fornisca l'accesso pubblico alla raccolta books:

```
aws opensearchserverless create-security-policy --name books-policy --type network
 \
  --policy "[{\"Description\":\"Public access for books collection\",\"Rules
\":[{\"ResourceType\":\"dashboard\",\"Resource\":[\"collection\/books\"]},
{\"ResourceType\":\"collection\",\"Resource\":[\"collection\/books\"]}],
\"AllowFromPublic\":true}]"
```

Risposta di esempio

```
{
    "securityPolicyDetail": {
        "type": "network",
        "name": "books-policy",
        "policyVersion": "MTY2OTI0MDI1Njk1NV8x",
        "policy": [
            {
                "Rules": [
                    {
                        "Resource": [
                            "collection/books"
                        ],
                        "ResourceType": "dashboard"
                    },
                    {
                        "Resource": [
                            "collection/books"
                        ],
                        "ResourceType": "collection"
                    }
                ],
                "AllowFromPublic": true,
                "Description": "Public access for books collection"
            }
        ],
        "createdDate": 1669240256955,
        "lastModifiedDate": 1669240256955
    }
}
```

5. Crea la raccolta books:

```
aws opensearchserverless create-collection --name books --type SEARCH
```

Risposta di esempio

```
{
    "createCollectionDetail": {
        "id": "8kw362bpwg4gx9b2f6e0",
        "name": "books",
```

```
        "status": "CREATING",
        "type": "SEARCH",
        "arn": "arn:aws:aoss:us-
east-1:123456789012:collection/8kw362bpwg4gx9b2f6e0",
        "kmsKeyArn": "auto",
        "createdDate": 1669240325037,
        "lastModifiedDate": 1669240325037
    }
}
```

6. Crea una policy di accesso ai dati che fornisca le autorizzazioni minime per indicizzare e cercare i dati nella raccolta books. Sostituisci l'ARN principale con l'ARN del `TutorialRole` dalla fase 1:

```
aws opensearchserverless create-access-policy \
  --name books-policy \
  --type data \
  --policy "[{\"Rules\":[{\"ResourceType\":\"index\",\"Resource\":
[\"index\/books\/books-index\"],\"Permission\":[\"aoss:CreateIndex
\",\"aoss:DescribeIndex\",\"aoss:ReadDocument\",\"aoss:WriteDocument
\",\"aoss:UpdateIndex\",\"aoss:DeleteIndex\"]}],\"Principal\":
[\"arn:aws:iam::123456789012:role\/TutorialRole\"]}]"
```

Risposta di esempio

```
{
    "accessPolicyDetail": {
        "type": "data",
        "name": "books-policy",
        "policyVersion": "MTY2OTI0MDM5NDY1M18x",
        "policy": [
            {
                "Rules": [
                    {
                        "Resource": [
                            "index/books/books-index"
                        ],
                        "Permission": [
                            "aoss:CreateIndex",
                            "aoss:DescribeIndex",
                            "aoss:ReadDocument",
                            "aoss:WriteDocument",
```

```
                            "aoss:UpdateDocument",
                            "aoss:DeleteDocument"
                        ],
                        "ResourceType": "index"
                    }
                ],
                "Principal": [
                    "arn:aws:iam::123456789012:role/TutorialRole"
                ]
            }
        ],
        "createdDate": 1669240394653,
        "lastModifiedDate": 1669240394653
    }
}
```

Ora `TutorialRole` dovrebbe essere in grado di indicizzare e cercare documenti nella raccolta books.

7. Per effettuare chiamate all' OpenSearch API, è necessario l'endpoint di raccolta. Invia la seguente richiesta per recuperare il parametro `collectionEndpoint`:

```
aws opensearchserverless batch-get-collection --names books
```

Risposta di esempio

```
{
    "collectionDetails": [
        {
            "id": "8kw362bpwg4gx9b2f6e0",
            "name": "books",
            "status": "ACTIVE",
            "type": "SEARCH",
            "description": "",
            "arn": "arn:aws:aoss:us-
east-1:123456789012:collection/8kw362bpwg4gx9b2f6e0",
            "createdDate": 1665765327107,
            "collectionEndpoint": "https://8kw362bpwg4gx9b2f6e0.us-
east-1.aoss.amazonaws.com",
            "dashboardEndpoint": "https://8kw362bpwg4gx9b2f6e0.us-
east-1.aoss.amazonaws.com/_dashboards"
        }
```

```
    ],
    "collectionErrorDetails": []
}
```

> **Note**
>
> Non sarà possibile visualizzare l'endpoint della nuova raccolta fino a quando lo stato della nuova raccolta non sarà `ACTIVE`. Potrebbe essere necessario effettuare più chiamate per verificare lo stato fino a quando la raccolta non viene creata correttamente.

8. Usa uno strumento HTTP come Postman o curl per indicizzare i dati nella raccolta books. Creeremo un indice denominato books-index e aggiungeremo un singolo documento.

   Invia la richiesta seguente all'endpoint di raccolta recuperato nella fase precedente, utilizzando le credenziali del `TutorialRole`.

```
PUT https://8kw362bpwg4gx9b2f6e0.us-east-1.aoss.amazonaws.com/books-index/_doc/1
{
  "title": "The Shining",
  "author": "Stephen King",
  "year": 1977
}
```

   Risposta di esempio

```
{
  "_index" : "books-index",
  "_id" : "1",
  "_version" : 1,
  "result" : "created",
  "_shards" : {
    "total" : 0,
    "successful" : 0,
    "failed" : 0
  },
  "_seq_no" : 0,
  "_primary_term" : 0
}
```

9. Per iniziare a cercare i dati nella raccolta, usa l'API di ricerca. La seguente query esegue una ricerca di base:

```
GET https://8kw362bpwg4gx9b2f6e0.us-east-1.aoss.amazonaws.com/books-index/_search
```

Risposta di esempio


```
{
    "took": 405,
    "timed_out": false,
    "_shards": {
        "total": 6,
        "successful": 6,
        "skipped": 0,
        "failed": 0
    },
    "hits": {
        "total": {
            "value": 2,
            "relation": "eq"
        },
        "max_score": 1.0,
        "hits": [
            {
                "_index": "books-index:0::3xJq14MBUaOS0wL26UU9:0",
                "_id": "F_bt4oMBLle5pYmm5q4T",
                "_score": 1.0,
                "_source": {
                    "title": "The Shining",
                    "author": "Stephen King",
                    "year": 1977
                }
            }
        ]
    }
}
```


Identity and Access Management per Amazon OpenSearch Serverless

AWS Identity and Access Management (IAM) è uno strumento Servizio AWS che aiuta un amministratore a controllare in modo sicuro l'accesso alle risorse. AWS Gli amministratori IAM controllano chi può essere autenticato (effettuato l'accesso) e autorizzato (disporre delle

autorizzazioni) a utilizzare le risorse Serverless. OpenSearch IAM è un software Servizio AWS che puoi utilizzare senza costi aggiuntivi.

Argomenti

- Policy basate sull'identità per Serverless OpenSearch
- Azioni politiche per Serverless OpenSearch
- Risorse politiche per Serverless OpenSearch
- Chiavi delle condizioni delle policy per Amazon OpenSearch Serverless
- ABAC con Serverless OpenSearch
- Utilizzo di credenziali temporanee con Serverless OpenSearch
- Ruoli collegati ai servizi per Serverless OpenSearch
- Altri tipi di policy
- Esempi di policy basate sull'identità per Serverless OpenSearch
- Supporto IAM Identity Center per Amazon OpenSearch Serverless

Policy basate sull'identità per Serverless OpenSearch

Supporta le policy basate sull’identità: sì

Le policy basate sull'identità sono documenti di policy di autorizzazione JSON che è possibile allegare a un'identità (utente, gruppo di utenti o ruolo IAM). Tali policy definiscono le operazioni che utenti e ruoli possono eseguire, su quali risorse e in quali condizioni. Per informazioni su come creare una policy basata su identità, consulta Definizione di autorizzazioni personalizzate IAM con policy gestite dal cliente nella Guida per l’utente di IAM.

Con le policy basate sull’identità di IAM, è possibile specificare quali operazioni e risorse sono consentite o respinte, nonché le condizioni in base alle quali le operazioni sono consentite o respinte. Per informazioni su tutti gli elementi utilizzabili in una policy JSON, consulta Guida di riferimento agli elementi delle policy JSON IAM nella Guida per l’utente IAM.

Esempi di policy basate sull'identità per Serverless OpenSearch

Per visualizzare esempi di policy basate sull'identità OpenSearch serverless, vedere. the section called “Esempi di policy basate su identità”

Azioni politiche per Serverless OpenSearch

Supporta le operazioni di policy: si

L'elemento `Action` di una policy JSON descrive le operazioni che è possibile utilizzare per consentire o negare l'accesso in una policy. Le azioni politiche in genere hanno lo stesso nome dell'operazione AWS API associata. Ci sono alcune eccezioni, ad esempio le operazioni di sola autorizzazione che non hanno un'operazione API corrispondente. Esistono anche alcune operazioni che richiedono più operazioni in una policy. Queste operazioni aggiuntive sono denominate operazioni dipendenti.

Includere le operazioni in una policy per concedere le autorizzazioni a eseguire l'operazione associata.

Le azioni politiche in OpenSearch Serverless utilizzano il seguente prefisso prima dell'azione:

```
aoss
```

Per specificare più operazioni in una sola istruzione, occorre separarle con la virgola.

```
"Action": [
      "aoss:action1",
      "aoss:action2"
         ]
```

Puoi specificare più operazioni tramite caratteri jolly (*). Ad esempio, per specificare tutte le azioni che iniziano con la parola `Describe`, includi la seguente azione:

```
"Action": "aoss:List*"
```

Per visualizzare esempi di politiche basate sull'identità OpenSearch Serverless, vedere. Esempi di policy basate sull'identità per Serverless OpenSearch

Risorse politiche per Serverless OpenSearch

Supporta le risorse relative alle policy: sì

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. In altre parole, quale entità principale può eseguire operazioni su quali risorse e in quali condizioni.

L'elemento JSON `Resource` della policy specifica l'oggetto o gli oggetti ai quali si applica l'operazione. Come best practice, specifica una risorsa utilizzando il suo nome della risorsa Amazon (ARN). Per le azioni che non supportano le autorizzazioni a livello di risorsa, si utilizza un carattere jolly (*) per indicare che l'istruzione si applica a tutte le risorse.

```
"Resource": "*"
```

Chiavi delle condizioni delle policy per Amazon OpenSearch Serverless

Supporta le chiavi di condizione delle policy specifiche del servizio: sì

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. In altre parole, quale entità principale può eseguire operazioni su quali risorse e in quali condizioni.

L'elemento `Condition` specifica quando le istruzioni vengono eseguite in base a criteri definiti. È possibile compilare espressioni condizionali che utilizzano operatori di condizione, ad esempio uguale a o minore di, per soddisfare la condizione nella policy con i valori nella richiesta. Per visualizzare tutte le chiavi di condizione AWS globali, consulta le chiavi di contesto delle condizioni AWS globali nella Guida per l'utente IAM.

Oltre al controllo degli accessi basato sugli attributi (ABAC), OpenSearch Serverless supporta le seguenti chiavi di condizione:

- `aoss:collection`
- `aoss:CollectionId`
- `aoss:index`

È possibile utilizzare le chiavi di condizione anche quando si forniscono le autorizzazioni per le policy di accesso e le policy di sicurezza. Esempio:

```
[
   {
      "Effect":"Allow",
      "Action":[
         "aoss:CreateAccessPolicy",
         "aoss:CreateSecurityPolicy"
      ],
      "Resource":"*",
      "Condition":{
         "StringLike":{
```

```
            "aoss:collection":"log"
         }
      }
   }
]
```

In questo esempio, la condizione si applica alle policy che contengono regole che corrispondono al nome o al modello di una raccolta. Le condizioni hanno il seguente comportamento:

- `StringEquals` - Si applica alle policy con regole che contengono la stringa di risorsa esatta "log" (ad esempio, `collection/log`).
- `StringLike` - Si applica alle policy con regole che contengono una stringa di risorsa che include "log" (ad esempio, `collection/log` ma anche `collection/logs-application` o `collection/applogs123`).

> **Note**
>
> Le chiavi di condizione raccolta non si applicano a livello di indice. Ad esempio, nella policy precedente, la condizione non si applicherebbe ad una policy di accesso o di sicurezza contenente la stringa di risorsa `index/logs-application/*`.

Per visualizzare un elenco di chiavi di condizione OpenSearch Serverless, consulta Condition keys for Amazon OpenSearch Serverless nel Service Authorization Reference. Per sapere con quali azioni e risorse puoi utilizzare una chiave di condizione, consulta Azioni definite da Amazon OpenSearch Serverless.

ABAC con Serverless OpenSearch

Supporta ABAC (tag nelle policy): sì

Il controllo degli accessi basato su attributi (ABAC) è una strategia di autorizzazione che definisce le autorizzazioni in base ad attributi chiamati tag. Puoi allegare tag a entità e AWS risorse IAM, quindi progettare politiche ABAC per consentire le operazioni quando il tag del principale corrisponde al tag sulla risorsa.

Per controllare l'accesso basato su tag, fornire informazioni sui tag nell'elemento condizione di una policy utilizzando le chiavi di condizione `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` o `aws:TagKeys`.

Se un servizio supporta tutte e tre le chiavi di condizione per ogni tipo di risorsa, il valore per il servizio è Sì. Se un servizio supporta tutte e tre le chiavi di condizione solo per alcuni tipi di risorsa, allora il valore sarà Parziale.

Per maggiori informazioni su ABAC, consulta Definizione delle autorizzazioni con autorizzazione ABAC nella Guida per l'utente di IAM. Per visualizzare un tutorial con i passaggi per l'impostazione di ABAC, consulta Utilizzo del controllo degli accessi basato su attributi (ABAC) nella Guida per l'utente di IAM.

Per ulteriori informazioni sull'etichettatura delle risorse OpenSearch Serverless, consulta. the section called "Assegnazione di tag alle raccolte"

Utilizzo di credenziali temporanee con Serverless OpenSearch

Supporta le credenziali temporanee: sì

Le credenziali temporanee forniscono l'accesso a breve termine alle AWS risorse e vengono create automaticamente quando si utilizza la federazione o si cambia ruolo. AWS consiglia di generare dinamicamente credenziali temporanee anziché utilizzare chiavi di accesso a lungo termine. Per ulteriori informazioni, consulta Credenziali di sicurezza temporanee in IAM e Servizi AWS compatibili con IAM nella Guida per l'utente IAM.

Ruoli collegati ai servizi per Serverless OpenSearch

Supporta i ruoli collegati ai servizi: sì

Un ruolo collegato al servizio è un tipo di ruolo di servizio collegato a un. Servizio AWS Il servizio può assumere il ruolo per eseguire un'operazione per tuo conto. I ruoli collegati al servizio vengono visualizzati nel tuo account Account AWS e sono di proprietà del servizio. Un amministratore IAM può visualizzare le autorizzazioni per i ruoli collegati al servizio, ma non modificarle.

Per informazioni dettagliate sulla creazione e la gestione di ruoli OpenSearch Serverless collegati ai servizi, consulta. the section called "Ruolo di creazione della raccolta"

Altri tipi di policy

AWS supporta tipi di policy aggiuntivi e meno comuni. Questi tipi di policy possono impostare il numero massimo di autorizzazioni concesse dai più tipi di policy comuni.

- Politiche di controllo del servizio (SCPs): SCPs sono politiche JSON che specificano le autorizzazioni massime per un'organizzazione o un'unità organizzativa (OU) in. AWS Organizations

AWS Organizations è un servizio per il raggruppamento e la gestione centralizzata di più AWS account di proprietà dell'azienda. Se abiliti tutte le funzionalità di un'organizzazione, puoi applicare le politiche di controllo del servizio (SCPs) a uno o tutti i tuoi account. L'SCP limita le autorizzazioni per le entità negli account dei membri, incluso ogni AWS utente root dell'account. Per ulteriori informazioni su Organizations and SCPs, consulta le politiche di controllo dei servizi nella Guida AWS Organizations per l'utente.

- Politiche di controllo delle risorse (RCPs): RCPs sono politiche JSON che puoi utilizzare per impostare le autorizzazioni massime disponibili per le risorse nei tuoi account senza aggiornare le politiche IAM allegate a ciascuna risorsa di tua proprietà. L'RCP limita le autorizzazioni per le risorse negli account dei membri e può influire sulle autorizzazioni effettive per le identità, incluso l'utente root dell' AWS account, indipendentemente dal fatto che appartengano all'organizzazione. Per ulteriori informazioni su Organizations e RCPs, incluso un elenco di AWS servizi che supportano RCPs, vedere Resource control policies (RCPs) nella Guida per l'AWS Organizations utente.

Esempi di policy basate sull'identità per Serverless OpenSearch

Per impostazione predefinita, gli utenti e i ruoli non dispongono dell'autorizzazione per creare o modificare risorse Serverless. OpenSearch Per concedere agli utenti l'autorizzazione a eseguire azioni sulle risorse di cui hanno bisogno, un amministratore IAM può creare policy IAM.

Per informazioni su come creare una policy basata su identità IAM utilizzando questi documenti di policy JSON di esempio, consulta Creazione di policy IAM (console) nella Guida per l'utente di IAM.

Per dettagli sulle azioni e sui tipi di risorse definiti da Amazon OpenSearch Serverless, incluso il formato di ARNs per ogni tipo di risorsa, consulta Azioni, risorse e chiavi di condizione per Amazon OpenSearch Serverless nel Service Authorization Reference.

Argomenti

- Best practice per le policy
- Utilizzo di OpenSearch Serverless nella console
- Amministrazione delle raccolte Serverless OpenSearch
- Visualizzazione delle raccolte Serverless OpenSearch
- Utilizzo delle operazioni OpenSearch API
- ABAC per le operazioni OpenSearch API

Best practice per le policy

Le policy basate su identità sono molto efficaci. Determinano se qualcuno può creare, accedere o eliminare risorse OpenSearch Serverless nel tuo account. Queste azioni possono comportare costi aggiuntivi per l' Account AWS. Quando si creano o modificano policy basate sull'identità, seguire queste linee guida e raccomandazioni:

Le politiche basate sull'identità determinano se qualcuno può creare, accedere o eliminare risorse OpenSearch Serverless nel tuo account. Queste azioni possono comportare costi aggiuntivi per l' Account AWS. Quando si creano o modificano policy basate sull'identità, seguire queste linee guida e raccomandazioni:

- Inizia con le policy AWS gestite e passa alle autorizzazioni con privilegi minimi: per iniziare a concedere autorizzazioni a utenti e carichi di lavoro, utilizza le politiche gestite che concedono le autorizzazioni per molti casi d'uso comuni.AWS Sono disponibili nel tuo. Account AWS Ti consigliamo di ridurre ulteriormente le autorizzazioni definendo politiche gestite dai AWS clienti specifiche per i tuoi casi d'uso. Per maggiori informazioni, consulta Policy gestite da AWS o Policy gestite da AWS per le funzioni dei processi nella Guida per l'utente di IAM.
- Applicazione delle autorizzazioni con privilegio minimo - Quando si impostano le autorizzazioni con le policy IAM, concedere solo le autorizzazioni richieste per eseguire un'attività. È possibile farlo definendo le azioni che possono essere intraprese su risorse specifiche in condizioni specifiche, note anche come autorizzazioni con privilegio minimo. Per maggiori informazioni sull'utilizzo di IAM per applicare le autorizzazioni, consulta Policy e autorizzazioni in IAM nella Guida per l'utente di IAM.
- Condizioni d'uso nelle policy IAM per limitare ulteriormente l'accesso - Per limitare l'accesso ad azioni e risorse è possibile aggiungere una condizione alle policy. Ad esempio, è possibile scrivere una condizione di policy per specificare che tutte le richieste devono essere inviate utilizzando SSL. Puoi anche utilizzare le condizioni per concedere l'accesso alle azioni del servizio se vengono utilizzate tramite uno specifico Servizio AWS, ad esempio CloudFormation. Per maggiori informazioni, consultare la sezione Elementi delle policy JSON di IAM: condizione nella Guida per l'utente di IAM.
- Utilizzo dello strumento di analisi degli accessi IAM per convalidare le policy IAM e garantire autorizzazioni sicure e funzionali - Lo strumento di analisi degli accessi IAM convalida le policy nuove ed esistenti in modo che aderiscano al linguaggio (JSON) della policy IAM e alle best practice di IAM. Lo strumento di analisi degli accessi IAM offre oltre 100 controlli delle policy e consigli utili per creare policy sicure e funzionali. Per maggiori informazioni, consultare Convalida delle policy per il Sistema di analisi degli accessi IAM nella Guida per l'utente di IAM.

- Richiedi l'autenticazione a più fattori (MFA): se hai uno scenario che richiede utenti IAM o un utente root nel Account AWS tuo, attiva l'MFA per una maggiore sicurezza. Per richiedere la MFA quando vengono chiamate le operazioni API, aggiungere le condizioni MFA alle policy. Per maggiori informazioni, consultare Protezione dell'accesso API con MFA nella Guida per l'utente di IAM.

Per maggiori informazioni sulle best practice in IAM, consulta Best practice di sicurezza in IAM nella Guida per l'utente di IAM.

Utilizzo di OpenSearch Serverless nella console

Per accedere a OpenSearch Serverless all'interno della console OpenSearch di servizio, è necessario disporre di un set minimo di autorizzazioni. Queste autorizzazioni devono consentirti di elencare e visualizzare i dettagli sulle risorse OpenSearch Serverless presenti nel tuo account. AWS Se crei una policy basata su identità più restrittiva rispetto alle autorizzazioni minime richieste, la console non funzionerà nel modo previsto per le entità (come i ruoli IAM) associate a tale policy.

Non è necessario consentire autorizzazioni minime per la console per gli utenti che effettuano chiamate solo verso AWS CLI o l'API. AWS Al contrario, è possibile accedere solo alle operazioni che soddisfano l'operazione API che stai cercando di eseguire.

La seguente politica consente a un utente di accedere a OpenSearch Serverless dalla console di OpenSearch servizio:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Resource": "*",
            "Effect": "Allow",
            "Action": [
                "aoss:ListCollections",
                "aoss:BatchGetCollection",
                "aoss:ListAccessPolicies",
                "aoss:ListSecurityConfigs",
                "aoss:ListSecurityPolicies",
                "aoss:ListTagsForResource",
                "aoss:ListVpcEndpoints",
                "aoss:GetAccessPolicy",
```

```
                "aoss:GetAccountSettings",
                "aoss:GetSecurityConfig",
                "aoss:GetSecurityPolicy"
            ]
        }
    ]
}
```

Amministrazione delle raccolte Serverless OpenSearch

Questa policy è un esempio di policy di «amministrazione della raccolta» che consente a un utente di gestire e amministrare le raccolte Amazon OpenSearch Serverless. L'utente può creare, visualizzare ed eliminare le raccolte.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Resource": "arn:aws:aoss:us-east-1:111122223333:collection/*",
            "Action": [
                "aoss:CreateCollection",
                "aoss:DeleteCollection",
                "aoss:UpdateCollection"
            ],
            "Effect": "Allow"
        },
        {
            "Resource": "*",
            "Action": [
                "aoss:BatchGetCollection",
                "aoss:ListCollections",
                "aoss:CreateAccessPolicy",
                "aoss:CreateSecurityPolicy"
            ],
            "Effect": "Allow"
        }
    ]
}
```

Visualizzazione delle raccolte Serverless OpenSearch

Questa policy di esempio consente a un utente di visualizzare i dettagli di tutte le raccolte Amazon OpenSearch Serverless nel proprio account. L'utente non può modificare le raccolte o le policy di sicurezza associate.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Resource": "*",
            "Action": [
                "aoss:ListAccessPolicies",
                "aoss:ListCollections",
                "aoss:ListSecurityPolicies",
                "aoss:ListTagsForResource",
                "aoss:BatchGetCollection"
            ],
            "Effect": "Allow"
        }
    ]
}
```

Utilizzo delle operazioni OpenSearch API

Le operazioni API del piano dati sono costituite dalle funzioni utilizzate in OpenSearch Serverless per ricavare valore in tempo reale dal servizio. Le operazioni API del piano di controllo sono costituite dalle funzioni utilizzate per configurare l'ambiente.

Per accedere al piano dati APIs e alle OpenSearch dashboard di Amazon OpenSearch Serverless dal browser, devi aggiungere due autorizzazioni IAM per le risorse di raccolta. Queste autorizzazioni sono e. `aoss:APIAccessAll` `aoss:DashboardsAccessAll`

> **Note**
>
> A partire dal 10 maggio 2023, OpenSearch Serverless richiede queste due nuove autorizzazioni IAM per le risorse di raccolta. L'`aoss:APIAccessAll`autorizzazione consente l'accesso al piano dati e l'`aoss:DashboardsAccessAll`autorizzazione consente

l'accesso alle OpenSearch dashboard dal browser. La mancata aggiunta delle due nuove autorizzazioni IAM genera un errore 403.

Questa policy di esempio consente a un utente di accedere al piano dati APIs per una raccolta specifica nel proprio account e di accedere alle OpenSearch dashboard per tutte le raccolte nel proprio account.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
         {
            "Effect": "Allow",
            "Action": "aoss:APIAccessAll",
            "Resource": "arn:aws:aoss:us-
east-1:111122223333:collection/collection-id"
        },
        {
            "Effect": "Allow",
            "Action": "aoss:DashboardsAccessAll",
            "Resource": "arn:aws:aoss:us-east-1:111122223333:dashboards/default"
        }
    ]
}
```

Entrambi `aoss:APIAccessAll` `aoss:DashboardsAccessAll` concedi l'autorizzazione IAM completa alle risorse di raccolta, mentre l'autorizzazione Dashboards fornisce anche l'accesso alle OpenSearch dashboard. Ogni autorizzazione funziona in modo indipendente, quindi una negazione esplicita `aoss:APIAccessAll` non blocca `aoss:DashboardsAccessAll` l'accesso alle risorse, inclusi Dev Tools. Lo stesso vale per una negazione dell'accesso. `aoss:DashboardsAccessAll` OpenSearch Serverless supporta le seguenti chiavi di condizione globali:

- `aws:CalledVia`
- `aws:CalledViaAWSService`
- `aws:CalledViaFirst`
- `aws:CalledViaLast`

- aws:CurrentTime
- aws:EpochTime
- aws:PrincipalAccount
- aws:PrincipalArn
- aws:PrincipallsAWSService
- aws:PrincipalOrgID
- aws:PrincipalOrgPaths
- aws:PrincipalType
- aws:PrincipalServiceName
- aws:PrincipalServiceNamesList
- aws:ResourceAccount
- aws:ResourceOrgID
- aws:ResourceOrgPaths
- aws:RequestedRegion
- aws:ResourceTag
- aws:SourceIp
- aws:SourceVpce
- aws:SourceVpc
- aws:userid
- aws:username
- aws:VpcSourceIp

Di seguito è riportato un esempio di utilizzo aws:SourceIp del blocco condition nella politica IAM del principale per le chiamate sul piano dati:

```
"Condition": {
    "IpAddress": {
         "aws:SourceIp": "203.0.113.0"
    }
}
```

Di seguito è riportato un esempio di utilizzo aws:SourceVpc del blocco condition nella politica IAM del principale per le chiamate sul piano dati:

```
"Condition": {
    "StringEquals": {
        "aws:SourceVpc": "vpc-0fdd2445d8EXAMPLE"
    }
}
```

Inoltre, viene offerto supporto per le seguenti chiavi specifiche OpenSearch Serverless:

- `aoss:CollectionId`
- `aoss:collection`

Di seguito è riportato un esempio di utilizzo del blocco condition `aoss:collection` nella politica IAM del principale per le chiamate sul piano dati:

```
"Condition": {
    "StringLike": {
         "aoss:collection": "log-*"
    }
}
```

ABAC per le operazioni OpenSearch API

Le policy basate sull'identità consentono di utilizzare i tag per controllare l'accesso al piano dati Amazon OpenSearch Serverless. APIs La seguente politica è un esempio per consentire ai principali collegati di accedere al piano dati APIs se la raccolta ha il tag: `team:devops`

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "aoss:APIAccessAll",
            "Resource": "arn:aws:aoss:us-
east-1:111122223333:collection/collection-id",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/team": "devops"
```

```
                }
            }
        }
    ]
}
```

La seguente politica è un esempio per negare ai principali collegati l'accesso al piano dati APIs e alle dashboard se la raccolta ha il tag: `environment:production`

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Deny",
            "Action": [
                "aoss:APIAccessAll",
                "aoss:DashboardsAccessAll"
            ],
            "Resource": "arn:aws:aoss:us-
east-1:111122223333:collection/collection-id"
        }
    ]
}
```

Amazon OpenSearch Serverless non supporta chiavi `RequestTag` di condizione `TagKeys` globali per il piano APIs dati.

Supporto IAM Identity Center per Amazon OpenSearch Serverless

Supporto IAM Identity Center per Amazon OpenSearch Serverless

Puoi utilizzare i principali (utenti e gruppi) di IAM Identity Center per accedere ai dati Amazon OpenSearch Serverless tramite Amazon OpenSearch Applications. Per abilitare il supporto di IAM Identity Center per Amazon OpenSearch Serverless, dovrai abilitare l'uso di IAM Identity Center. Per ulteriori informazioni su come eseguire questa operazione, consulta Cos'è IAM Identity Center?

> **Note**
>
> Per accedere alle raccolte Amazon OpenSearch Serverless utilizzando utenti o gruppi IAM Identity Center, devi utilizzare la funzionalità OpenSearch UI (Applications). L'accesso diretto ai dashboard OpenSearch Serverless utilizzando le credenziali IAM Identity Center non è supportato. Per ulteriori informazioni, consulta Guida introduttiva all' OpenSearch interfaccia utente.

Dopo aver creato l'istanza IAM Identity Center, l'amministratore dell'account del cliente deve creare un'applicazione IAM Identity Center per il servizio Amazon OpenSearch Serverless. Questo può essere fatto chiamando CreateSecurityConfig:. L'amministratore dell'account cliente può specificare quali attributi verranno utilizzati per autorizzare la richiesta. Gli attributi predefiniti utilizzati sono e `UserId` `GroupId.`

L'integrazione di IAM Identity Center per Amazon OpenSearch Serverless utilizza le seguenti autorizzazioni AWS IAM Identity Center (IAM):

- `aoss:CreateSecurityConfig`— Creare un provider IAM Identity Center
- `aoss:ListSecurityConfig`— Elenca tutti i provider IAM Identity Center presenti nell'account corrente.
- `aoss:GetSecurityConfig`— Visualizza le informazioni sui provider di IAM Identity Center.
- `aoss:UpdateSecurityConfig`— Modifica una determinata configurazione di IAM Identity Center
- `aoss:DeleteSecurityConfig`— Eliminare un provider IAM Identity Center.

La seguente policy di accesso basata sull'identità può essere utilizzata per gestire tutte le configurazioni di IAM Identity Center:

JSON

```
{
"Version": "2012-10-17",
    "Statement": [
        {
"Action": [
                "aoss:CreateSecurityConfig",
```

```
                "aoss:DeleteSecurityConfig",
                "aoss:GetSecurityConfig",
                "aoss:UpdateSecurityConfig",
                "aoss:ListSecurityConfigs"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

> **Note**
>
> L'`Resource`elemento deve essere un jolly.

Creazione di un provider IAM Identity Center (console)

Puoi creare un provider IAM Identity Center per abilitare l'autenticazione con OpenSearch l'applicazione. Per abilitare l'autenticazione IAM Identity Center for OpenSearch Dashboards, procedi nel seguente modo:

1. Accedi alla console di Amazon OpenSearch Service.
2. Nel pannello di navigazione a sinistra, espandi Serverless e scegli Autenticazione.
3. Scegli l'autenticazione IAM Identity Center.
4. Seleziona Modifica
5. Seleziona la casella accanto a Autentica con IAM Identity Center.
6. Seleziona la chiave degli attributi utente e gruppo dal menu a discesa. Gli attributi utente verranno utilizzati per autorizzare gli utenti in base a `UserName``UserId`, e. `Email` Gli attributi di gruppo verranno utilizzati per autenticare gli utenti in base `GroupName` a e. `GroupId`
7. Seleziona l'istanza IAM Identity Center.
8. Seleziona Salva

Creazione del provider IAM Identity Center (AWS CLI)

Per creare un provider IAM Identity Center utilizzando AWS Command Line Interface (AWS CLI), utilizza il seguente comando:

```
aws opensearchserverless create-security-config \
--region us-east-2 \
--name "iamidentitycenter-config" \
--description "description" \
--type "iamidentitycenter" \
--iam-identity-center-options '{
    "instanceArn": "arn:aws:sso:::instance/ssoins-99199c99e99ee999",
    "userAttribute": "UserName",
    "groupAttribute": "GroupId"
}'
```

Dopo aver abilitato un IAM Identity Center, i clienti possono solo modificare gli attributi di utenti e gruppi.

```
aws opensearchserverless update-security-config \
--region us-east-1 \
--id <id_from_list_security_configs> \
--config-version <config_version_from_get_security_config> \
--iam-identity-center-options-updates '{
    "userAttribute": "UserId",
    "groupAttribute": "GroupId"
}'
```

Per visualizzare il provider IAM Identity Center utilizzando il AWS Command Line Interface, utilizza il seguente comando:

```
aws opensearchserverless list-security-configs --type iamidentitycenter
```

Eliminazione di un provider IAM Identity Center

IAM Identity Center offre due istanze di provider, una per l'account dell'organizzazione e una per l'account membro. Se è necessario modificare l'istanza IAM Identity Center, è necessario eliminare la configurazione di sicurezza tramite l'`DeleteSecurityConfig`API e creare una nuova configurazione di sicurezza utilizzando la nuova istanza IAM Identity Center. Il seguente comando può essere utilizzato per eliminare un provider IAM Identity Center:

```
aws opensearchserverless delete-security-config \
--region us-east-1 \
--id <id_from_list_security_configs>
```

Concessione dell'accesso a IAM Identity Center ai dati di raccolta

Dopo aver abilitato il provider IAM Identity Center, puoi aggiornare la politica di accesso ai dati di raccolta per includere i principali di IAM Identity Center. I principali di IAM Identity Center devono essere aggiornati nel seguente formato:

```
[
   {
"Rules":[
       ...
      ],
      "Principal":[
         "iamidentitycenter/<iamidentitycenter-instance-id>/user/<UserName>",
         "iamidentitycenter/<iamidentitycenter-instance-id>/group/<GroupId>"
      ]
   }
]
```

> **Note**
>
> Amazon OpenSearch Serverless supporta solo un'istanza IAM Identity Center per tutte le raccolte di clienti e può supportare fino a 100 gruppi per un singolo utente. Se tenti di utilizzare un numero di istanze superiore a quello consentito, potresti riscontrare un'incoerenza nell'elaborazione delle autorizzazioni relative alla politica di accesso ai dati e riceverai un `403` messaggio di errore.

Puoi concedere l'accesso a raccolte, a indici o a entrambi. Se desideri che utenti diversi abbiano autorizzazioni diverse, dovrai creare più regole. Per un elenco delle autorizzazioni disponibili, consulta Identity and Access Management in Amazon OpenSearch Service. Per informazioni su come formattare una politica di accesso, consulta Concedere alle identità SAML l'accesso ai dati di raccolta.

Crittografia in Amazon OpenSearch Serverless

Crittografia dei dati a riposo

Ogni raccolta Amazon OpenSearch Serverless che crei è protetta con la crittografia dei dati inattivi, una funzionalità di sicurezza che aiuta a prevenire l'accesso non autorizzato ai tuoi dati. Encryption

at rest utilizza AWS Key Management Service (AWS KMS) per archiviare e gestire le chiavi di crittografia. Per eseguire la crittografia, utilizza l'algoritmo Advanced Encryption Standard con chiavi a 256 bit (AES-256).

Argomenti

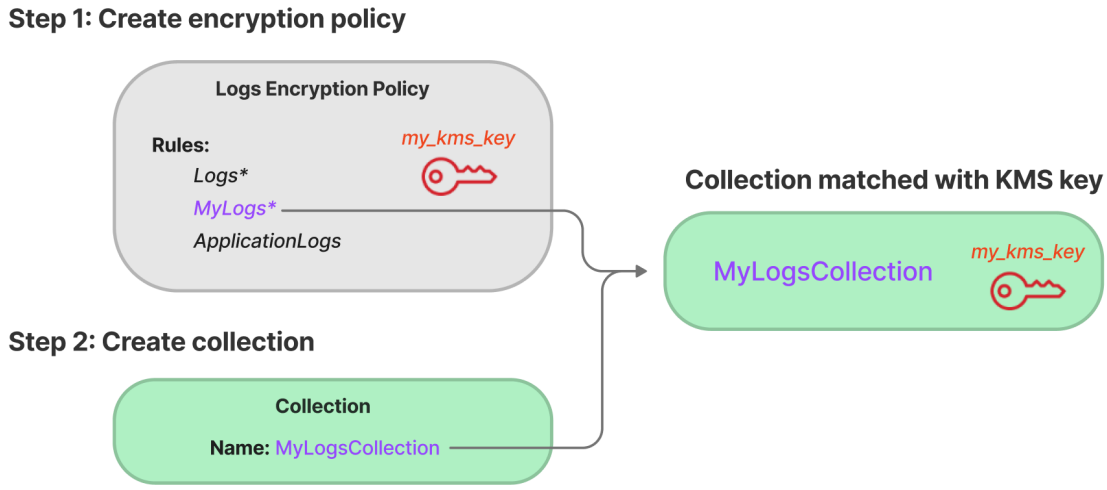
- Policy di crittografia
- Considerazioni
- Autorizzazioni richieste
- Policy della chiave per una chiave gestita dal cliente
- In OpenSearch che modo Serverless utilizza le sovvenzioni in AWS KMS
- Creazione di policy di crittografia (console)
- Creazione di policy di crittografia (AWS CLI)
- Visualizzazione delle policy di crittografia
- Aggiornamento di policy di crittografia
- Eliminazione delle policy di crittografia

Policy di crittografia

Con le policy di crittografia è possibile gestire molte raccolte su larga scala assegnando automaticamente una chiave di crittografia alle raccolte appena create che corrispondono a un nome o a un modello specifico.

Quando si crea una policy di crittografia, è possibile specificare un prefisso, ossia una regola di corrispondenza basata su caratteri jolly come `MyCollection*`, oppure inserire un unico nome di raccolta. Quindi, quando si crea una raccolta che corrisponde a tale modello di prefisso o nome, ad essa vengono assegnate automaticamente la policy e la chiave KMS corrispondenti.

Quando si crea una raccolta, è possibile specificare una AWS KMS chiave in due modi: tramite politiche di sicurezza o direttamente nella `CreateCollection` richiesta. Se fornite una AWS KMS chiave come parte della `CreateCollection` richiesta, questa ha la precedenza su qualsiasi politica di sicurezza corrispondente. Con questo approccio, hai la flessibilità di ignorare le impostazioni di crittografia basate su policy per raccolte specifiche quando necessario.



Le policy di crittografia contengono i seguenti elementi:

- `Rules`: una o più regole di abbinamento delle raccolte, ciascuna con i seguenti sottoelementi:
  - `ResourceType`: attualmente l'unica opzione è "raccolta". Le policy di crittografia si applicano solo alle risorse di raccolta.
  - `Resource`: uno o più nomi o modelli di raccolta a cui si applicherà la policy, nel formato `collection/<collection name|pattern>`.
- `AWSOwnedKey`: se usare o meno una Chiave di proprietà di AWS.
- `KmsARN`: se impostati `AWSOwnedKey` su falso, specifica il nome della risorsa Amazon (ARN) della chiave KMS con cui crittografare le raccolte associate. Se si include questo parametro, OpenSearch Serverless lo ignora. `AWSOwnedKey`

La seguente policy di esempio assegnerà una chiave gestita dal cliente a qualsiasi raccolta futura denominata `autopartsinventory`, nonché alle raccolte che iniziano con il termine "vendite":

```
{
   "Rules":[
      {
         "ResourceType":"collection",
         "Resource":[
            "collection/autopartsinventory",
            "collection/sales*"
         ]
      }
   ],
   "AWSOwnedKey":false,
```

```
   "KmsARN":"arn:aws:kms:us-east-1:123456789012:key/93fd6da4-a317-4c17-
bfe9-382b5d988b36"
}
```

Anche se una policy corrisponde al nome di una raccolta, è possibile scegliere di sovrascrivere tale assegnazione automatica durante la creazione della raccolta se il modello di risorsa contiene un carattere jolly (*). Se scegli di ignorare l'assegnazione automatica delle chiavi, OpenSearch Serverless crea per te una politica di crittografia denominata auto-< ***collection-name*** > e la allega alla raccolta. La policy inizialmente si applica solo a una singola raccolta, ma è possibile modificarla per includere raccolte aggiuntive.

Se modifichi le regole delle policy in modo che non corrispondano più a una raccolta, la chiave KMS associata non verrà annullata da tale raccolta. La raccolta rimane sempre crittografata con la sua chiave di crittografia iniziale. Se desideri modificare la chiave di crittografia per una raccolta, è necessario creare nuovamente la raccolta.

Se a una raccolta corrispondono le regole di più policy, viene utilizzata la regola più specifica. Ad esempio, se una policy contiene una regola per `collection/log*` e un'altra per `collection/logSpecial`, la chiave di crittografia per la seconda policy viene utilizzata in quanto è più specifica.

Non è possibile utilizzare un nome o un prefisso in una politica se esiste già in un'altra politica. OpenSearch Serverless visualizza un errore se si tenta di configurare modelli di risorse identici in diverse politiche di crittografia.

Considerazioni

Quando configuri la crittografia per le tue raccolte considera quanto segue:

- La crittografia dei dati inattivi è obbligatoria per tutte le raccolte serverless.
- Hai la possibilità di utilizzare una chiave gestita dal cliente o una Chiave di proprietà di AWS. Se scegli una chiave gestita dal cliente, ti consigliamo di abilitare la rotazione automatica delle chiavi.
- Dopo la creazione di una raccolta non è possibile modificarne la chiave di crittografia. Scegliete con AWS KMS attenzione quale utilizzare la prima volta che configurate una raccolta.
- Una raccolta può corrispondere solo a una singola policy di crittografia.
- Le raccolte con chiavi KMS uniche non possono condividere le unità di OpenSearch calcolo (OCUs) con altre raccolte. Ogni raccolta con una chiave unica richiede le sue 4. OCUs
- Se aggiorni la chiave KMS in una policy di crittografia, la modifica non influirà sulle corrispondenti raccolte esistenti a cui sono state già assegnate chiavi KMS.

- OpenSearch Serverless non controlla esplicitamente le autorizzazioni degli utenti sulle chiavi gestite dal cliente. Se un utente dispone delle autorizzazioni per accedere a una raccolta tramite una policy di accesso ai dati, sarà in grado di inserire e interrogare i dati crittografati con la chiave associata.

Autorizzazioni richieste

Encryption at rest for OpenSearch Serverless utilizza le seguenti autorizzazioni AWS Identity and Access Management (IAM). È possibile specificare le condizioni IAM per limitare gli utenti a raccolte specifiche.

- `aoss:CreateSecurityPolicy`: crea una policy di crittografia.
- `aoss:ListSecurityPolicies`: elenca tutte le policy e le raccolte di crittografia a cui sono allegate.
- `aoss:GetSecurityPolicy`: visualizza i dettagli di una policy di crittografia specifica.
- `aoss:UpdateSecurityPolicy`: modifica una policy di crittografia.
- `aoss:DeleteSecurityPolicy`: elimina una policy di crittografia.

Il seguente esempio di policy di accesso basata sull'identità fornisce le autorizzazioni minime necessarie all'utente per gestire le policy di crittografia con il modello di risorsa `collection/application-logs`.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Action":[
            "aoss:CreateSecurityPolicy",
            "aoss:UpdateSecurityPolicy",
            "aoss:DeleteSecurityPolicy",
            "aoss:GetSecurityPolicy"
         ],
         "Resource":"*",
         "Condition":{
            "StringEquals":{
```

```
                "aoss:collection":"application-logs"
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":[
            "aoss:ListSecurityPolicies"
         ],
         "Resource":"*"
      }
   ]
}
```

Policy della chiave per una chiave gestita dal cliente

Se si seleziona una chiave gestita dal cliente per proteggere una raccolta, OpenSearch Serverless ottiene l'autorizzazione a utilizzare la chiave KMS per conto del principale che effettua la selezione. Tale principale, un utente o un ruolo, deve disporre delle autorizzazioni sulla chiave KMS richieste da Serverless. OpenSearch Puoi fornire queste autorizzazioni in una policy delle chiavi o in una policy IAM.

OpenSearch Serverless effettua chiamate API `Decrypt` KMS durante le operazioni di manutenzione, come la scalabilità automatica `GenerateDataKey` e gli aggiornamenti software. Potresti osservare queste chiamate al di fuori dei tuoi schemi di traffico tipici. Queste chiamate fanno parte delle normali operazioni di servizio e non indicano traffico utente attivo.

OpenSearch Serverless genera un messaggio `KMSKeyInaccessibleException` quando non può accedere alla chiave KMS che crittografa i dati inattivi. Ciò si verifica quando si disabilita o si elimina la chiave KMS o si revocano le concessioni che consentono a Serverless di utilizzare la chiave. OpenSearch

OpenSearch Serverless richiede almeno le seguenti autorizzazioni su una chiave gestita dal cliente:

- km: DescribeKey
- km: CreateGrant

Esempio:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
        "Action": "kms:DescribeKey",
        "Effect": "Allow",
        "Principal": {
            "AWS": "arn:aws:iam::123456789012:user/Dale"
        },
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "kms:ViaService": "aoss.us-east-1.amazonaws.com"
            }
        }
    },
    {
        "Action": "kms:CreateGrant",
        "Effect": "Allow",
        "Principal": {
            "AWS": "arn:aws:iam::123456789012:user/Dale"
        },
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "kms:ViaService": "aoss.us-east-1.amazonaws.com"
            },
            "ForAllValues:StringEquals": {
                "kms:GrantOperations": [
                    "Decrypt",
                    "GenerateDataKey"
                ]
            },
            "Bool": {
                "kms:GrantIsForAWSResource": "true"
            }
        }
    }
  ]
}
```

OpenSearch Serverless crea una concessione con le autorizzazioni kms: GenerateDataKey e kms:decrypt.

Per ulteriori informazioni, consulta Utilizzo delle policy delle chiavi in AWS KMS nella Guida per gli sviluppatori di AWS Key Management Service .

In OpenSearch che modo Serverless utilizza le sovvenzioni in AWS KMS

OpenSearch Serverless richiede una concessione per utilizzare una chiave gestita dal cliente.

Quando crei una politica di crittografia nel tuo account con una nuova chiave, OpenSearch Serverless crea una concessione per tuo conto inviando una CreateGrantrichiesta a. AWS KMS Le concessioni AWS KMS vengono utilizzate per fornire a OpenSearch Serverless l'accesso a una chiave KMS in un account cliente.

OpenSearch Serverless richiede la concessione per utilizzare la chiave gestita dal cliente per le seguenti operazioni interne:

- Invia DescribeKeyrichieste AWS KMS a per verificare che l'ID della chiave gestita dal cliente simmetrico fornito sia valido.
- Invia GenerateDataKeyle richieste alla chiave KMS per creare chiavi di dati con cui crittografare gli oggetti.
- Invia richieste Decrypt a per AWS KMS decrittografare le chiavi di dati crittografate in modo che possano essere utilizzate per crittografare i dati.

Puoi revocare l'accesso alla concessione o rimuovere l'accesso del servizio alla chiave gestita dal cliente in qualsiasi momento. In tal caso, OpenSearch Serverless non sarà in grado di accedere a nessuno dei dati crittografati dalla chiave gestita dal cliente, il che influirà su tutte le operazioni che dipendono da tali dati, causando `AccessDeniedException` errori e guasti nei flussi di lavoro asincroni.

OpenSearch Serverless annulla le sovvenzioni in un flusso di lavoro asincrono quando una determinata chiave gestita dal cliente non è associata a politiche o raccolte di sicurezza.

Creazione di policy di crittografia (console)

In una policy di crittografia, si specifica una chiave KMS e una serie di modelli di raccolta a cui verrà applicata la policy. A qualsiasi nuova raccolta che corrisponde a uno dei modelli definiti nella policy verrà assegnata la chiave KMS corrispondente al momento della creazione della raccolta. Consigliamo di creare policy di crittografia prima di iniziare a creare raccolte.

Per creare una politica di crittografia Serverless OpenSearch

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Nel pannello di navigazione a sinistra, espandi Serverless e scegli Encryption policies (Policy di crittografia).
3. Scegli Create encryption policy (Crea policy di crittografia).
4. Fornisci un nome e una descrizione per la policy.
5. In Resources (Risorse), inserisci uno o più modelli di risorse per questa policy di crittografia. Tutte le nuove raccolte create nella Regione e nell' Account AWS attuali che corrispondono a uno dei modelli vengono automaticamente assegnate a questa policy. Ad esempio, se inserisci `ApplicationLogs` (senza caratteri jolly) e successivamente crei una raccolta con quel nome, la policy e la chiave KMS corrispondente vengono assegnate a tale raccolta.

   Puoi anche fornire un prefisso, ad esempio `Logs*`, che assegni la policy a tutte le nuove raccolte i cui nomi iniziano con `Logs`. Utilizzando i caratteri jolly, puoi gestire le impostazioni di crittografia per varie raccolte su larga scala.
6. In Encryption (Crittografia), scegli una chiave KMS da usare.
7. Scegli Create (Crea).

Fase successiva: creazione di raccolte

Dopo aver configurato una o più policy di crittografia, puoi iniziare a creare raccolte che corrispondono alle regole definite in tali policy. Per istruzioni, consulta the section called “Creazione di raccolte”.

Nella fase di crittografia della creazione della raccolta, OpenSearch Serverless ti informa che il nome che hai inserito corrisponde al modello definito in una politica di crittografia e assegna automaticamente la chiave KMS corrispondente alla raccolta. Se il modello di risorsa contiene un carattere jolly (*), puoi scegliere di ignorare la corrispondenza e selezionare la tua chiave.

Creazione di policy di crittografia (AWS CLI)

Per creare una politica di crittografia utilizzando le operazioni dell'API OpenSearch Serverless, si specificano modelli di risorse e una chiave di crittografia in formato JSON. La CreateSecurityPolicyrichiesta accetta sia le politiche in linea che i file.json.

Le policy di crittografia assumono il formato seguente. Questo file `my-policy.json` di esempio corrisponde a qualsiasi raccolta futura denominata `autopartsinventory`, nonché a qualsiasi raccolta con nomi che iniziano con `sales`.

```
{
   "Rules":[
      {
         "ResourceType":"collection",
         "Resource":[
            "collection/autopartsinventory",
            "collection/sales*"
         ]
      }
   ],
   "AWSOwnedKey":false,
   "KmsARN":"arn:aws:kms:us-east-1:123456789012:key/93fd6da4-a317-4c17-
bfe9-382b5d988b36"
}
```

Per utilizzare una chiave di proprietà del servizio, imposta `AWSOwnedKey` su `true`:

```
{
   "Rules":[
      {
         "ResourceType":"collection",
         "Resource":[
            "collection/autopartsinventory",
            "collection/sales*"
         ]
      }
   ],
   "AWSOwnedKey":true
}
```

La seguente richiesta crea la policy di crittografia:

```
aws opensearchserverless create-security-policy \
    --name sales-inventory \
    --type encryption \
    --policy file://my-policy.json
```

Quindi, utilizzate l'operazione CreateCollectionAPI per creare una o più raccolte che corrispondono a uno dei modelli di risorse.

Visualizzazione delle policy di crittografia

Prima di creare una raccolta, hai la possibilità di visualizzare in anteprima le policy di crittografia esistenti nel tuo account per vedere quali hanno uno schema di risorse che corrisponde al nome della raccolta. La seguente ListSecurityPoliciesrichiesta elenca tutte le politiche di crittografia del tuo account:

```
aws opensearchserverless list-security-policies --type encryption
```

La richiesta restituisce informazioni su tutte le policy di crittografia configurate. Utilizza i contenuti dell'elemento `policy` per visualizzare le regole del modello definite nella policy:

```
{
   "securityPolicyDetails": [
      {
         "createdDate": 1663693217826,
         "description": "Sample encryption policy",
         "lastModifiedDate": 1663693217826,
         "name": "my-policy",
         "policy": "{\"Rules\":[{\"ResourceType\":\"collection\",\"Resource\":
[\"collection/autopartsinventory\",\"collection/sales*\"]}],\"AWSOwnedKey\":true}",
         "policyVersion": "MTY2MzY5MzIxNzgyNl8x",
         "type": "encryption"
      }
   ]
}
```

Per visualizzare informazioni dettagliate su una politica specifica, inclusa la chiave KMS, usa il GetSecurityPolicycomando.

Aggiornamento di policy di crittografia

Se aggiorni la chiave KMS in una policy di crittografia, la modifica verrà applicata solo alle raccolte appena create che corrispondono al nome o al modello configurato. Non influisce sulle raccolte esistenti a cui sono già assegnate chiavi KMS.

Lo stesso vale per le regole di abbinamento delle policy. Se aggiungi, modifichi o elimini una regola, la modifica si applica solo alle raccolte appena create. Se modifichi le regole di una policy in modo

che non corrisponda più al nome di una raccolta, le raccolte esistenti non perdono la chiave KMS assegnata.

Per aggiornare una politica di crittografia nella console OpenSearch Serverless, scegli Politiche di crittografia, seleziona la politica da modificare e scegli Modifica. Apporta le modifiche necessarie, quindi scegli Save (Salva).

Per aggiornare una politica di crittografia utilizzando l'API OpenSearch Serverless, utilizza l'UpdateSecurityPolicyoperazione. La seguente richiesta aggiorna una policy di crittografia con un nuovo documento JSON della policy:

```
aws opensearchserverless update-security-policy \
    --name sales-inventory \
    --type encryption \
    --policy-version 2 \
    --policy file://my-new-policy.json
```

Eliminazione delle policy di crittografia

Quando elimini una policy di crittografia, tutte le raccolte che attualmente utilizzano la chiave KMS definita nella policy non saranno modificate. Per eliminare una policy nella console OpenSearch Serverless, seleziona la policy e scegli Elimina.

Puoi anche utilizzare l'DeleteSecurityPolicyoperazione:

```
aws opensearchserverless delete-security-policy --name my-policy --type encryption
```

Crittografia dei dati in transito

In OpenSearch Serverless, tutti i percorsi di una raccolta sono crittografati in transito utilizzando Transport Layer Security 1.2 (TLS) con un codice AES-256 standard di settore. L'accesso a tutti APIs e alle dashboard di Opensearch avviene anche tramite TLS 1.2. TLS è un insieme di protocolli crittografici standard del settore utilizzati per crittografare le informazioni scambiate sulla rete.

Accesso alla rete per Amazon OpenSearch Serverless

Le impostazioni di rete per una raccolta Amazon OpenSearch Serverless determinano se la raccolta è accessibile su Internet da reti pubbliche o se è necessario accedervi privatamente.

L'accesso privato può applicarsi a uno o entrambi i seguenti elementi:

- OpenSearch Endpoint VPC gestiti senza server
- Supportato Servizi AWS come Amazon Bedrock

Puoi configurare l'accesso alla rete separatamente per l'endpoint di una raccolta e l'OpenSearchendpoint OpenSearch Dashboards corrispondente.

L'accesso alla rete è il meccanismo di isolamento che consente l'accesso da diverse reti di origine. Ad esempio, se l'endpoint OpenSearch Dashboards di una raccolta è accessibile pubblicamente ma l'endpoint OpenSearch API no, un utente può accedere ai dati della raccolta solo tramite Dashboards quando si connette da una rete pubblica. Se provano a chiamarli OpenSearch APIs direttamente da una rete pubblica, verranno bloccati. Le impostazioni di rete possono essere utilizzate per tali permutazioni dall'origine al tipo di risorsa. Amazon OpenSearch Serverless supporta sia la connettività che IPv4 la IPv6 connettività.

Argomenti

- Policy di rete
- Considerazioni
- Autorizzazioni necessarie per configurare le politiche di rete
- Priorità delle policy
- Creazione di policy di rete (console)
- Creazione di policy di rete (AWS CLI)
- Visualizzazione delle policy di rete
- Aggiornamento delle policy di rete
- Eliminazione delle policy di rete

Policy di rete

Le policy di rete consentono di gestire molte raccolte su larga scala assegnando automaticamente le impostazioni di accesso alla rete alle raccolte che corrispondono alle regole definite nella policy.

In una policy di rete, si specifica una serie di regole. Queste regole definiscono le autorizzazioni di accesso agli endpoint di raccolta e OpenSearch agli endpoint di Dashboards. Ogni regola è composta da un tipo di accesso (pubblico o privato) e da un tipo di risorsa (raccolta e/o OpenSearch endpoint Dashboards). Per ogni tipo di risorsa (`collection` e `dashboard`), si specifica una serie di regole che definiscono a quali raccolte si applicherà la policy.

In questa policy di esempio, la prima regola specifica l'accesso degli endpoint VPC sia all'endpoint di raccolta che all'endpoint Dashboards per tutte le raccolte che iniziano con il termine. `marketing*` Specifica inoltre l'accesso ad Amazon Bedrock.

> **Note**
>
> L'accesso privato Servizi AWS ad Amazon Bedrock si applica solo all'endpoint della raccolta, non all' OpenSearch endpoint OpenSearch Dashboards. Anche se lo `ResourceType` è`dashboard`, Servizi AWS non può essere concesso l'accesso alle dashboard. OpenSearch

La seconda regola specifica l'accesso pubblico alla raccolta `finance`, ma solo per l'endpoint di raccolta (nessun accesso a Dashboards).

```
[
   {
      "Description":"Marketing access",
      "Rules":[
         {
            "ResourceType":"collection",
            "Resource":[
               "collection/marketing*"
            ]
         },
         {
            "ResourceType":"dashboard",
            "Resource":[
               "collection/marketing*"
            ]
         }
      ],
      "AllowFromPublic":false,
      "SourceVPCEs":[
         "vpce-050f79086ee71ac05"
      ],
      "SourceServices":[
         "bedrock.amazonaws.com"
      ],
   },
   {
      "Description":"Sales access",
      "Rules":[
```

```
         {
            "ResourceType":"collection",
            "Resource":[
               "collection/finance"
            ]
         }
      ],
      "AllowFromPublic":true
   }
]
```

Questa politica fornisce l'accesso pubblico solo alle OpenSearch dashboard per le raccolte che iniziano con «finanza». Qualsiasi tentativo di accedere direttamente all' OpenSearch API fallirà.

```
[
  {
    "Description": "Dashboards access",
    "Rules": [
      {
        "ResourceType": "dashboard",
        "Resource": [
          "collection/finance*"
        ]
      }
    ],
    "AllowFromPublic": true
  }
]
```

Le policy di rete possono essere applicate alle raccolte esistenti e alle raccolte future. Ad esempio, è possibile creare una raccolta e quindi creare una policy di rete con una regola che corrisponda al nome della raccolta. Non è necessario creare policy di rete prima di creare le raccolte.

Considerazioni

Quando configuri l'accesso di rete per le tue raccolte considera quanto segue:

- Se prevedi di configurare l'accesso agli endpoint VPC per una raccolta, devi prima creare almeno un endpoint VPC gestito senza server. OpenSearch

- L'accesso privato a si applica Servizi AWS solo all'endpoint della raccolta, non all' OpenSearchendpoint Dashboards. OpenSearch Anche se lo `ResourceType` è`dashboard`, Servizi AWS non può essere concesso l'accesso alle dashboard. OpenSearch
- Se una raccolta è accessibile dalle reti pubbliche, è accessibile anche da tutti gli endpoint VPC OpenSearch gestiti senza server e tutto il resto. Servizi AWS
- A una singola raccolta possono essere applicate più policy di rete. Per ulteriori informazioni, consulta the section called “Priorità delle policy”.

Autorizzazioni necessarie per configurare le politiche di rete

L'accesso alla rete per OpenSearch Serverless utilizza le seguenti autorizzazioni AWS Identity and Access Management (IAM). È possibile specificare le condizioni IAM per limitare gli utenti alle policy di rete associate a raccolte specifiche.

- `aoss:CreateSecurityPolicy`: crea una policy di accesso alla rete.
- `aoss:ListSecurityPolicies`: elenca tutte le policy di rete nell'account corrente.
- `aoss:GetSecurityPolicy`: visualizza una specifica della policy di accesso alla rete.
- `aoss:UpdateSecurityPolicy`: modifica una determinata policy di accesso alla rete e modifica l'ID VPC o la designazione di accesso pubblico.
- `aoss:DeleteSecurityPolicy`: elimina una policy di accesso alla rete (dopo che è stata scollegata da tutte le raccolte).

La seguente policy di accesso basata sull'identità consente a un utente di visualizzare tutte le policy di rete e di aggiornarle in base al modello delle risorse `collection/application-logs`:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "aoss:UpdateSecurityPolicy"
            ],
            "Resource": "*",
            "Condition": {
```

```
                "StringEquals": {
                    "aoss:collection": "application-logs"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "aoss:ListSecurityPolicies",
                "aoss:GetSecurityPolicy"
            ],
            "Resource": "*"
        }
    ]
}
```

> **Note**
>
> Inoltre, OpenSearch Serverless richiede le autorizzazioni `aoss:APIAccessAll` e le `aoss:DashboardsAccessAll` autorizzazioni per le risorse di raccolta. Per ulteriori informazioni, consulta the section called “Utilizzo delle operazioni OpenSearch API”.

Priorità delle policy

Possono verificarsi situazioni in cui le regole delle policy di rete si sovrappongono, all'interno delle policy o tra di esse. Quando ciò accade, una regola che specifica l'accesso pubblico ha la precedenza su una regola che specifica l'accesso privato per tutte le raccolte comuni a entrambe le regole.

Ad esempio, nella policy seguente, entrambe le regole assegnano l'accesso di rete alla raccolta `finance`, ma una regola specifica l'accesso VPC mentre l'altra specifica l'accesso pubblico. In questa situazione, l'accesso pubblico prevale sull'accesso al VPC solo per la raccolta "finance" (poiché esiste in entrambe le regole), quindi la raccolta "finance" sarà accessibile dalle reti pubbliche. La raccolta delle vendite avrà accesso al VPC dall'endpoint specificato.

```
[
   {
      "Description":"Rule 1",
      "Rules":[
```

```
        {
           "ResourceType":"collection",
           "Resource":[
              "collection/sales",
              "collection/finance"
           ]
        }
     ],
     "AllowFromPublic":false,
     "SourceVPCEs":[
        "vpce-050f79086ee71ac05"
     ]
  },
  {
     "Description":"Rule 2",
     "Rules":[
        {
           "ResourceType":"collection",
           "Resource":[
              "collection/finance"
           ]
        }
     ],
     "AllowFromPublic":true
  }
]
```

Se a una raccolta si applicano più endpoint VPC di regole diverse, la raccolta sarà accessibile da tutti gli endpoint specificati in quanto le regole sono additive. Se lo `AllowFromPublic` imposti `true` ma ne fornisci anche uno o più `SourceVPCEs` oppure`SourceServices`, OpenSearch Serverless ignora gli endpoint VPC e gli identificatori del servizio e le raccolte associate avranno accesso pubblico.

Creazione di policy di rete (console)

Le policy di rete possono essere applicate alle raccolte esistenti e alle raccolte future. Consigliamo di creare le policy di rete prima di iniziare a creare raccolte.

Per creare una policy di rete Serverless OpenSearch

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Nel pannello di navigazione a sinistra, espandi Serverless e scegli Network policies (Policy di rete).

3. Scegli Create network policy (Crea policy di rete).
4. Fornisci un nome e una descrizione per la policy.
5. Fornisci una o più regole. Queste regole definiscono le autorizzazioni di accesso per le tue raccolte OpenSearch Serverless e i relativi endpoint OpenSearch Dashboard.

   Ogni regola contiene i seguenti elementi:

Elemento	Description
Nome regola	Un nome che descrive i contenuti della regola. Ad esempio, "Accesso al VPC per il team di marketing".
Tipo di accesso	Scegli l'accesso pubblico o privato. Quindi, seleziona una o entrambe le seguenti opzioni: • Endpoint VPC per l'accesso: specifica uno o più endpoint VPC gestiti senza server, endpoint VPC gestitiOpenSearch . • Servizio AWS accesso privato: seleziona uno o Servizi AWS più supporti.
Tipo di risorsa	Seleziona se fornire l'accesso agli OpenSearch endpoint (che consente di effettuare chiamate all' OpenSearch API), alle OpenSearch dashboard (che consente l'accesso alle visualizzazioni e all'interfaccia utente per i OpenSearch plug-in) o entrambi. **Note** Servizio AWS l'accesso privato si applica solo all'endpoint della raccolta, non all' OpenSearch endpoint Dashboards. OpenSearc h Anche se selezioni OpenSearc

Elemento	Description
	h Dashboards, Servizi AWS può essere concesso solo l'accesso all'endpoint.

Per ogni tipo di risorsa selezionato, puoi scegliere raccolte esistenti a cui applicare le impostazioni dei criteri, and/or creare uno o più modelli di risorse. I modelli di risorse sono costituiti da un prefisso e da un carattere jolly (*), e definiscono a quali raccolte si applicheranno le impostazioni delle policy.

Ad esempio, se includi un modello denominato `Marketing*`, a qualsiasi raccolta nuova o esistente il cui nome inizia con "Marketing" verranno applicate automaticamente le impostazioni di rete di questa policy. Una singolo carattere jolly (`*`) applica la policy a tutte le raccolte attuali e future.

Inoltre, potete specificare il nome di una collezione futura senza caratteri jolly, ad esempio`Finance`. OpenSearch Serverless applicherà le impostazioni dei criteri a qualsiasi raccolta appena creata con quel nome esatto.

6. Quando sei soddisfatto della configurazione della policy, scegli Create (Crea).

Creazione di policy di rete (AWS CLI)

Per creare una politica di rete utilizzando le operazioni dell'API OpenSearch Serverless, specificate le regole in formato JSON. La CreateSecurityPolicyrichiesta accetta sia le politiche in linea che i file.json. Tutte le raccolte e i modelli devono assumere la forma `collection/<collection name|pattern>`.

> **Note**
>
> Il tipo di risorsa consente `dashboards` solo l'autorizzazione alle OpenSearch dashboard, ma per far funzionare le OpenSearch dashboard è necessario consentire anche l'accesso alla raccolta dalle stesse fonti. Osserva la seconda policy di seguito per un esempio.

Per specificare l'accesso privato, includi uno o entrambi i seguenti elementi:

- `SourceVPCEs`— Specificare uno o più endpoint OpenSearch VPC gestiti senza server.
- `SourceServices`— Specificare l'identificatore di uno o più supportati. Servizi AWS Attualmente sono supportati i seguenti identificatori di servizio:
  - `bedrock.amazonaws.com`— Amazon Bedrock

La seguente policy di rete di esempio fornisce l'accesso privato, a un endpoint VPC e Amazon Bedrock, agli endpoint di raccolta solo per le raccolte che iniziano con il prefisso. `log*` Gli utenti autenticati non possono accedere alle OpenSearch dashboard; possono solo accedere all'endpoint di raccolta a livello di codice.

```
[
   {
      "Description":"Private access for log collections",
      "Rules":[
         {
            "ResourceType":"collection",
            "Resource":[
               "collection/log*"
            ]
         }
      ],
      "AllowFromPublic":false,
      "SourceVPCEs":[
         "vpce-050f79086ee71ac05"
      ],
      "SourceServices":[
         "bedrock.amazonaws.com"
      ],
   }
]
```

La seguente politica fornisce l'accesso pubblico all' OpenSearch endpoint e alle OpenSearch dashboard per una singola raccolta denominata. `finance` Se la raccolta non esiste, le impostazioni di rete verranno applicate alla raccolta se e quando viene creata.

```
[
   {
      "Description":"Public access for finance collection",
      "Rules":[
         {
```

```
            "ResourceType":"dashboard",
            "Resource":[
               "collection/finance"
            ]
         },
         {
            "ResourceType":"collection",
            "Resource":[
               "collection/finance"
            ]
         }
      ],
      "AllowFromPublic":true
   }
]
```

La seguente richiesta crea la policy di rete di cui sopra:

```
aws opensearchserverless create-security-policy \
    --name sales-inventory \
    --type network \
    --policy "[{\"Description\":\"Public access for finance collection\",\"Rules
\":[{\"ResourceType\":\"dashboard\",\"Resource\":[\"collection\/finance\"]},
{\"ResourceType\":\"collection\",\"Resource\":[\"collection\/finance\"]}],
\"AllowFromPublic\":true}]"
```

Per fornire la policy in un file JSON, utilizzare il formato `--policy file://my-policy.json`

Visualizzazione delle policy di rete

Prima di creare una raccolta, hai la possibilità di visualizzare in anteprima le policy di rete esistenti nel tuo account per vedere quali hanno uno schema di risorse che corrisponde al nome della raccolta. La ListSecurityPoliciesrichiesta seguente elenca tutte le politiche di rete del tuo account:

```
aws opensearchserverless list-security-policies --type network
```

La richiesta restituisce informazioni su tutte le policy di rete configurate. Per visualizzare le regole del modello definite in una politica specifica, trova le informazioni sulla politica nel contenuto dell'`securityPolicySummaries`elemento nella risposta. Prendi nota della `name` fine `type` di questa politica e utilizza queste proprietà in una GetSecurityPolicyrichiesta per ricevere una risposta con i seguenti dettagli della politica:

```
{
    "securityPolicyDetail": [
        {
            "type": "network",
            "name": "my-policy",
            "policyVersion": "MTY2MzY5MTY1MDA3Ml8x",
            "policy": "[{\"Description\":\"My network policy rule\",\"Rules\":
[{\"ResourceType\":\"dashboard\",\"Resource\":[\"collection/*\"]}],\"AllowFromPublic
\":true}]",
            "createdDate": 1663691650072,
            "lastModifiedDate": 1663691650072
        }
    ]
}
```

Per visualizzare informazioni dettagliate su una politica specifica, utilizzare il GetSecurityPolicycomando.

Aggiornamento delle policy di rete

Quando si modificano gli endpoint VPC o la designazione di accesso pubblico per una rete, vengono interessate anche tutte le raccolte associate. Per aggiornare una politica di rete nella console OpenSearch Serverless, espandi Criteri di rete, seleziona la politica da modificare e scegli Modifica. Apporta le modifiche necessarie, quindi scegli Save (Salva).

Per aggiornare una politica di rete utilizzando l'API OpenSearch Serverless, usa il UpdateSecurityPolicycomando. È necessario includere una versione della policy nella richiesta. È possibile recuperare la versione della policy utilizzando i comandi `ListSecurityPolicies` o `GetSecurityPolicy`. L'inclusione della versione più recente delle policy garantisce di non sovrascrivere inavvertitamente una modifica apportata da qualcun altro.

La seguente richiesta aggiorna una policy di rete con un nuovo documento JSON della policy:

```
aws opensearchserverless update-security-policy \
    --name sales-inventory \
    --type network \
    --policy-version MTY2MzY5MTY1MDA3Ml8x \
    --policy file://my-new-policy.json
```

Eliminazione delle policy di rete

Prima di eliminare una policy di rete, è necessario scollegarla da tutte le raccolte. Per eliminare una policy nella console OpenSearch Serverless, seleziona la policy e scegli Elimina.

Puoi anche usare il DeleteSecurityPolicycomando:

```
aws opensearchserverless delete-security-policy --name my-policy --type network
```

Conformità FIPS in Amazon Serverless OpenSearch

Amazon OpenSearch Serverless supporta il Federal Information Processing Standards (FIPS) 140-2, uno standard governativo statunitense e canadese che specifica i requisiti di sicurezza per i moduli crittografici che proteggono le informazioni sensibili. Quando ti connetti a endpoint compatibili con FIPS con OpenSearch Serverless, le operazioni crittografiche avvengono utilizzando librerie crittografiche convalidate FIPS.

OpenSearch Gli endpoint FIPS serverless sono disponibili laddove è supportato FIPS. Regioni AWS Questi endpoint utilizzano TLS 1.2 o versioni successive e algoritmi crittografici convalidati da FIPS per tutte le comunicazioni. Per ulteriori informazioni, consulta la conformità FIPS nella Guida per l'utente ad accesso verificato.AWS

Argomenti

- Utilizzo degli endpoint FIPS con Serverless OpenSearch
- Utilizza gli endpoint FIPS con AWS SDKs
- Configurazione dei gruppi di sicurezza per gli endpoint VPC
- Usa l'endpoint FIPS VPC
- Verifica la conformità FIPS
- Risolvi i problemi di connettività degli endpoint FIPS nelle zone ospitate private

Utilizzo degli endpoint FIPS con Serverless OpenSearch

Regioni AWS Laddove è supportato FIPS, le raccolte OpenSearch Serverless sono accessibili tramite endpoint standard e conformi a FIPS. Per ulteriori informazioni, consulta la conformità FIPS nella Guida per l'utente ad accesso verificato.AWS

Negli esempi seguenti, sostituisci *collection_id* e *Regione AWS* con il tuo ID di raccolta e il relativo Regione AWS.

- Endpoint standard:**https://*collection_id*.*Regione AWS*.aoss.amazonaws.com**.
- Endpoint conforme allo standard FIPS:. **https://*collection_id*.*Regione AWS*.aoss-fips.amazonaws.com**

Allo stesso modo, le OpenSearch dashboard sono accessibili tramite endpoint standard e conformi a FIPS:

- Endpoint di dashboard standard:. **https://*collection_id*.*Regione AWS*.aoss.amazonaws.com/_dashboards**
- Endpoint di dashboard conformi allo standard FIPS:. **https://*collection_id*.*Regione AWS*.aoss-fips.amazonaws.com/_dashboards**

> **Note**
>
> Nelle regioni abilitate alla FIPS, gli endpoint standard e conformi allo standard FIPS forniscono una crittografia conforme allo standard FIPS. Gli endpoint specifici per FIPS consentono di soddisfare i requisiti di conformità che impongono specificamente l'uso di endpoint con FIPS nel nome.

Utilizza gli endpoint FIPS con AWS SDKs

Quando si utilizza AWS SDKs, è possibile specificare l'endpoint FIPS durante la creazione del client. Nell'esempio seguente, sostituisci *collection_id* e *Regione AWS* con il tuo ID di raccolta e il relativo. Regione AWS

```
# Python SDK example
from opensearchpy import OpenSearch, RequestsHttpConnection, AWSV4SignerAuth
import boto3
host = '"https://collection_id.Regione AWS.aoss-fips.amazonaws.com"
region = 'us-west-2'
service = 'aoss'
credentials = boto3.Session().get_credentials()
auth = AWSV4SignerAuth(credentials, region, service)
client = OpenSearch(
    hosts = [{'host': host, 'port': 443}],
    http_auth = auth,
    use_ssl = True,
```

```
    verify_certs = True,
    connection_class = RequestsHttpConnection,
    pool_maxsize = 20
)
```

Configurazione dei gruppi di sicurezza per gli endpoint VPC

Per garantire una comunicazione corretta con il tuo endpoint Amazon VPC (VPC) conforme a FIPS, crea o modifica un gruppo di sicurezza per consentire il traffico HTTPS in entrata (porta TCP 443) dalle risorse del tuo VPC che devono accedere a Serverless. OpenSearch Quindi associa questo gruppo di sicurezza al tuo endpoint VPC durante la creazione o modificando l'endpoint dopo la creazione. Per ulteriori informazioni, consulta Creazione di un gruppo di sicurezza nella Guida per l'utente di Amazon VPC.

Usa l'endpoint FIPS VPC

Dopo aver creato l'endpoint VPC conforme a FIPS, puoi utilizzarlo per accedere OpenSearch a Serverless dalle risorse all'interno del tuo VPC. Per utilizzare l'endpoint per le operazioni API, configura il tuo SDK per utilizzare l'endpoint FIPS regionale come descritto nella sezione. Utilizzo degli endpoint FIPS con Serverless OpenSearch Per accedere alle OpenSearch dashboard, utilizza l'URL Dashboards specifico della raccolta, che verrà indirizzato automaticamente attraverso l'endpoint VPC conforme a FIPS quando vi si accede dall'interno del VPC. Per ulteriori informazioni, consulta Utilizzo dei OpenSearch pannelli di controllo con Amazon Service OpenSearch.

Verifica la conformità FIPS

Per verificare che le connessioni a OpenSearch Serverless utilizzino la crittografia conforme a FIPS, utilizza per monitorare le chiamate API effettuate a Serverless. AWS CloudTrail OpenSearch Verifica che il campo nei log venga visualizzato per le chiamate API. `eventSource` CloudTrail `aoss-fips.amazonaws.com`

Per accedere alle OpenSearch dashboard, puoi utilizzare gli strumenti di sviluppo del browser per controllare i dettagli della connessione TLS e verificare che vengano utilizzate suite di crittografia conformi a FIPS.

Risolvi i problemi di connettività degli endpoint FIPS nelle zone ospitate private

Gli endpoint FIPS funzionano con raccolte Amazon OpenSearch Serverless ad accesso pubblico. Per le raccolte VPC appena create che utilizzano endpoint VPC appena creati, gli endpoint

FIPS funzionano come previsto. Per altre raccolte VPC, potrebbe essere necessario eseguire la configurazione manuale per garantire il corretto funzionamento degli endpoint FIPS.

Per configurare zone ospitate private FIPS in Amazon Route 53

1. Apri la console Route 53 all'indirizzo https://console.aws.amazon.com/route53/.
2. Controlla le tue zone ospitate:
   a. Individua le zone ospitate in cui si trovano le Regioni AWS tue raccolte.
   b. Verifica i modelli di denominazione delle zone ospitate:
      - Formato non FIPS:. *region*`.aoss.amazonaws.com`
      - Formato FIPS:. *region*`.aoss-fips.amazonaws.com`
   c. Conferma che il Tipo per tutte le tue zone ospitate sia impostato su Zona ospitata privata.
3. Se manca la zona ospitata privata FIPS:
   a. Seleziona la zona ospitata privata non FIPS corrispondente.
   b. Copia le informazioni associate VPCs. Ad esempio: `vpc-1234567890abcdef0 | us-east-2`.
   c. Trova il record del dominio wildcard. Ad esempio: `*.us-east-2.aoss.amazonaws.com`.
   d. Copia il valore e il percorso del traffico nelle informazioni. Ad esempio: `uoc1c1qsw7poexampleewjeno1pte3rw.3ym756xh7yj.aoss.searchservices.aws`.
4. Crea la zona ospitata privata FIPS:
   a. Crea una nuova zona ospitata privata con il formato FIPS. Ad esempio: `us-east-2.aoss-fips.amazonaws.com`.
   b. Per Associato VPCs, inserisci le informazioni sul VPC che hai copiato dalla zona ospitata privata non FIPS.
5. Aggiungi un nuovo record con le seguenti impostazioni:
   a. Nome del record: *
   b. Tipo di record: CNAME
   c. Valore: inserisci il valore/indirizzamento del traffico verso le informazioni che hai copiato in precedenza.

Problemi comuni

Se riscontri problemi di connettività con gli endpoint VPC conformi a FIPS, utilizza le seguenti informazioni per risolvere il problema.

- Errori di risoluzione DNS: non è possibile risolvere il nome di dominio dell'endpoint FIPS all'interno del VPC
- Timeout di connessione: scadono le tue richieste all'endpoint FIPS
- Errori di accesso negato: l'autenticazione o l'autorizzazione falliscono quando si utilizzano gli endpoint FIPS
- Record di zona ospitata privata mancanti per le raccolte solo VPC

Per risolvere i problemi di connettività degli endpoint FIPS

1. Verifica la configurazione della tua zona ospitata privata:
   a. Conferma che esiste una zona ospitata privata per il dominio dell'endpoint FIPS (. `*.region.aoss-fips.amazonaws.com`
   b. Verifica che la zona ospitata privata sia associata al VPC corretto.

      Per ulteriori informazioni, consulta le zone ospitate private nella Amazon Route 53 Developer Guide e Manage DNS names nella AWS PrivateLink Guida.
2. Verifica la risoluzione DNS:
   a. Connect a un'istanza EC2 nel tuo VPC.
   b. Esegui il comando seguente:

      ```
      nslookup collection-id.region.aoss-fips.amazonaws.com
      ```
   c. Verifica che la risposta includa l'indirizzo IP privato del tuo endpoint VPC.

      Per ulteriori informazioni, consulta le politiche degli endpoint e gli attributi DNS nella Amazon VPC User Guide.
3. Controlla le impostazioni del tuo gruppo di sicurezza:
   a. Verifica che il gruppo di sicurezza collegato all'endpoint VPC consenta il traffico HTTPS (porta 443) dalle tue risorse.

b. Verifica che i gruppi di sicurezza per le tue risorse consentano il traffico in uscita verso l'endpoint VPC.

Per ulteriori informazioni, consulta le politiche degli endpoint nella AWS PrivateLink Guida e i gruppi di sicurezza nella Amazon VPC User Guide.

4. Controlla la configurazione ACL della tua rete:

   - Verifica che la rete ACLs consenta il traffico tra le tue risorse e l'endpoint VPC.

     Per ulteriori informazioni, consulta Network ACLs in the Amazon VPC User Guide.

5. Rivedi la tua politica sugli endpoint:

   - Verifica che la policy degli endpoint VPC consenta le azioni richieste sulle tue risorse Serverless. OpenSearch

     Per ulteriori informazioni, consulta le autorizzazioni degli endpoint VPC richieste e le politiche degli endpoint nella Guida.AWS PrivateLink

> **Tip**
>
> Se utilizzi resolver DNS personalizzati nel tuo VPC, configurali per inoltrare le richieste di domini ai server. `*.amazonaws.com` AWS

Controllo dell'accesso ai dati per Amazon OpenSearch Serverless

Con il controllo dell'accesso ai dati in Amazon OpenSearch Serverless, puoi consentire agli utenti di accedere a raccolte e indici, indipendentemente dal meccanismo di accesso o dalla fonte di rete. Puoi fornire l'accesso ai ruoli IAM e alle identità SAML.

Puoi gestire le autorizzazioni di accesso tramite le policy di accesso ai dati, che si applicano alle raccolte e alle risorse dell'indice. Le policy di accesso ai dati consentono di gestire le raccolte su larga scala assegnando automaticamente autorizzazioni di accesso a raccolte e indici che corrispondono a uno schema specifico. È possibile applicare più policy di accesso ai dati a una singola risorsa. Tieni presente che devi disporre di una politica di accesso ai dati per la tua raccolta per poter accedere all'URL delle dashboard OpenSearch .

Argomenti

- Policy di accesso ai dati rispetto alle policy IAM
- Autorizzazioni IAM necessarie per configurare le politiche di accesso ai dati
- Sintassi delle policy
- Autorizzazioni delle policy supportate
- Set di dati di esempio nelle dashboard OpenSearch
- Creazione di policy di accesso ai dati (console)
- Creazione di policy di accesso ai dati (AWS CLI)
- Visualizzazione di policy di accesso ai dati
- Aggiornamento delle policy di accesso ai dati
- Eliminazione delle policy di accesso ai dati
- Accesso ai dati tra account

Policy di accesso ai dati rispetto alle policy IAM

Le policy di accesso ai dati sono logicamente separate dalle policy AWS Identity and Access Management (IAM). Le autorizzazioni IAM controllano l'accesso alle operazioni API serverless, come `CreateCollection` e `ListAccessPolicies`. Le politiche di accesso ai dati controllano l'accesso alle OpenSearch operazioni supportate da OpenSearch Serverless, come `PUT <index>` o. `GET _cat/indices`

Le autorizzazioni IAM che controllano l'accesso alle operazioni dell'API della policy di accesso ai dati, come `aoss:CreateAccessPolicy` e `aoss:GetAccessPolicy` (descritte nella sezione successiva), non influiscono sull'autorizzazione specificata in una policy di accesso ai dati.

Ad esempio, supponiamo che una policy IAM impedisca a un utente di creare policy di accesso ai dati per `collection-a`, ma gli consenta di creare policy di accesso ai dati per tutte le raccolte (`*`):

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Deny",
            "Action": [
                "aoss:CreateAccessPolicy"
```

```
            ],
            "Resource": "*",
            "Condition": {
                "StringLike": {
                    "aoss:collection": "collection-a"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "aoss:CreateAccessPolicy"
            ],
            "Resource": "*"
        }
    ]
}
```

Se l'utente crea una policy di accesso ai dati che consente determinate autorizzazioni per tutte le raccolte (`collection/*` o `index/*/*`), la policy si applicherà a tutte le raccolte, inclusa la raccolta A.

> **Important**
>
> La concessione delle autorizzazioni nell'ambito di una politica di accesso ai dati non è sufficiente per accedere ai dati della raccolta OpenSearch Serverless. A un principale associato deve inoltre essere concesso l'accesso alle autorizzazioni `aoss:APIAccessAll` IAM e. `aoss:DashboardsAccessAll` Entrambe le autorizzazioni garantiscono l'accesso completo alle risorse di raccolta, mentre l'autorizzazione Dashboards fornisce anche l'accesso alle dashboard. OpenSearch Se un principale non dispone di entrambe queste autorizzazioni IAM, riceverà 403 errori quando tenta di inviare richieste alla raccolta. Per ulteriori informazioni, consulta the section called “Utilizzo delle operazioni OpenSearch API”.

Autorizzazioni IAM necessarie per configurare le politiche di accesso ai dati

Il controllo dell'accesso ai dati per OpenSearch Serverless utilizza le seguenti autorizzazioni IAM. È possibile specificare condizioni IAM per limitare gli utenti a nomi di policy di accesso specifici.

- `aoss:CreateAccessPolicy`: crea una policy di accesso.

- `aoss:ListAccessPolicies`: elenca tutte le policy di accesso.
- `aoss:GetAccessPolicy`: visualizza i dettagli su una policy di accesso specifica.
- `aoss:UpdateAccessPolicy`: modifica una policy di accesso.
- `aoss:DeleteAccessPolicy`: elimina una policy di accesso.

La seguente policy di accesso basata sull'identità consente a un utente di visualizzare tutte le policy di accesso e aggiornare le policy che contengono il modello delle risorse `collection/logs`.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
                "aoss:ListAccessPolicies",
                "aoss:GetAccessPolicy"
            ],
            "Effect": "Allow",
            "Resource": "*"
        },
        {
            "Action": [
                "aoss:UpdateAccessPolicy"
            ],
            "Effect": "Allow",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aoss:collection": [
                        "logs"
                    ]
                }
            }
        }
    ]
}
```

> **Note**
>
> Inoltre, OpenSearch Serverless richiede le autorizzazioni `aoss:APIAccessAll` e le `aoss:DashboardsAccessAll` autorizzazioni per le risorse di raccolta. Per ulteriori informazioni, consulta the section called “Utilizzo delle operazioni OpenSearch API”.

Sintassi delle policy

Una policy di accesso ai dati include una serie di regole, ognuna con i seguenti elementi:

Elemento	Description
`ResourceType`	Il tipo di risorsa (raccolta o indice) a cui si applicano le autorizzazioni. Le autorizzazioni degli alias e dei modelli sono a livello di raccolta, mentre le autorizzazioni per la creazione, la modifica e la ricerca dei dati sono a livello di indice. Per ulteriori informazioni, consulta la sezione Supported policy permissions (Autorizzazioni delle policy supportate).
`Resource`	Un elenco di modelli di nomi and/or di risorse. I modelli sono prefissi seguiti da un carattere jolly (*), che consentono di applicare le autorizzazioni associate a più risorse. • Le raccolte hanno il formato `collection/` *`<name|pattern>`* . • Gli indici hanno il formato `index/`*`<collection-name|pattern>`* `/`*`<index-name|pattern/>`* .
`Permission`	Un elenco di autorizzazioni da concedere per le risorse specificate. Per un elenco completo delle autorizzazioni e delle operazioni dell'API che queste concedono, consulta la sezione the section called “Operazioni e autorizzazioni API supportate OpenSearch”.
`Principal`	Un elenco di uno o più principali a cui concedere l'accesso. I principali possono essere ruoli IAM ARNs o identità SAML. Questi principali devono essere all'interno dell' Account AWS corrente. Le policy di accesso ai dati non supportano direttamente l'accesso tra account diversi, ma puoi includere nella policy un ruolo che un utente di un altro account Account AWS può assumere

Elemento	Description
	nell'account proprietario della raccolta. Per ulteriori informazioni, consulta the section called “Accesso ai dati tra account”.

La seguente policy di esempio concede autorizzazioni di alias e modello alla raccolta denominata `autopartsinventory`, nonché a tutte le raccolte che iniziano con il prefisso `sales*`. Inoltre, concede autorizzazioni di lettura e scrittura a tutti gli indici della raccolta `autopartsinventory` e a tutti gli indici della raccolta `salesorders` che iniziano con il prefisso `orders*`.

```
[
   {
      "Description": "Rule 1",
      "Rules":[
         {
            "ResourceType":"collection",
            "Resource":[
               "collection/autopartsinventory",
               "collection/sales*"
            ],
            "Permission":[
               "aoss:CreateCollectionItems",
               "aoss:UpdateCollectionItems",
               "aoss:DescribeCollectionItems"
            ]
         },
         {
            "ResourceType":"index",
            "Resource":[
               "index/autopartsinventory/*",
               "index/salesorders/orders*"
            ],
            "Permission":[
               "aoss:*"
            ]
         }
      ],
      "Principal":[
         "arn:aws:iam::123456789012:user/Dale",
         "arn:aws:iam::123456789012:role/RegulatoryCompliance",
         "saml/123456789012/myprovider/user/Annie",
```

```
            "saml/123456789012/anotherprovider/group/Accounting"
        ]
    }
]
```

Non puoi negare esplicitamente l'accesso all'interno di una policy. Pertanto, tutte le autorizzazioni delle policy sono additive. Ad esempio, se a un utente una policy concede `aoss:ReadDocument` e un'altra concede `aoss:WriteDocument`, l'utente avrà entrambe le autorizzazioni. Se una terza policy concede `aoss:*` allo stesso utente, l'utente può eseguire tutte le azioni sull'indice associato; le autorizzazioni più restrittive non sostituiscono quelle meno restrittive.

Autorizzazioni delle policy supportate

Le seguenti autorizzazioni sono supportate nelle policy di accesso ai dati. Per le operazioni OpenSearch API consentite da ciascuna autorizzazione, consulta. the section called “Operazioni e autorizzazioni API supportate OpenSearch”

Autorizzazioni della raccolta

- `aoss:CreateCollectionItems`
- `aoss:DeleteCollectionItems`
- `aoss:UpdateCollectionItems`
- `aoss:DescribeCollectionItems`
- `aoss:*`

Autorizzazioni dell'indice

- `aoss:ReadDocument`
- `aoss:WriteDocument`
- `aoss:CreateIndex`
- `aoss:DeleteIndex`
- `aoss:UpdateIndex`
- `aoss:DescribeIndex`
- `aoss:*`

Set di dati di esempio nelle dashboard OpenSearch

OpenSearch Dashboards fornisce set di dati di esempio che includono visualizzazioni, dashboard e altri strumenti per aiutarti a esplorare le dashboard prima di aggiungere i tuoi dati. Per creare indici a partire da questi dati di esempio, è necessaria una politica di accesso ai dati che fornisca le autorizzazioni per il set di dati con cui desideri lavorare. La seguente politica utilizza un wildcard (*) per fornire le autorizzazioni per tutti e tre i set di dati di esempio.

```
[
  {
    "Rules": [
      {
        "Resource": [
          "index/<collection-name>/opensearch_dashboards_sample_data_*"
        ],
        "Permission": [
          "aoss:CreateIndex",
          "aoss:DescribeIndex",
          "aoss:ReadDocument"
        ],
        "ResourceType": "index"
      }
    ],
    "Principal": [
      "arn:aws:iam::<account-id>:user/<user>"
    ]
  }
]
```

Creazione di policy di accesso ai dati (console)

È possibile creare una policy di accesso ai dati utilizzando l'editor visivo o in formato JSON. A qualsiasi nuova raccolta che corrisponde a uno dei modelli definiti nella policy verranno assegnate le autorizzazioni corrispondenti al momento della creazione della raccolta.

Per creare una politica di accesso ai dati OpenSearch Serverless

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, espandi Serverless e in Sicurezza, scegli Politiche di accesso ai dati.
3. Scegli Create access policy (Crea policy di accesso).

4. Fornisci un nome e una descrizione per la policy.
5. Fornisci un nome per la prima regola nella policy. Ad esempio, "Accesso alla raccolta di log".
6. Scegli Add principals (Aggiungi principali) e seleziona uno o più ruoli IAM o utenti e gruppi SAML a cui fornire l'accesso ai dati.

   > **Note**
   >
   > Per poter selezionare i principali dai menu a discesa, è necessario disporre delle autorizzazioni `iam:ListUsers` e `iam:ListRoles` (per i principali IAM) e dell'autorizzazione `aoss:ListSecurityConfigs` (per le identità SAML).

7. Scegli Grant (Concedi) e seleziona le autorizzazioni per l'alias, il modello e l'indice da concedere ai principali associati. Per un elenco completo delle autorizzazioni e l'accesso che queste concedono, consulta la sezione the section called “Operazioni e autorizzazioni API supportate OpenSearch”.
8. (Facoltativo) Configurazione di regole aggiuntive per la policy.
9. Scegli Create (Crea). Potrebbe esserci circa un minuto di ritardo tra il momento in cui si crea la policy e il momento in cui vengono applicate le autorizzazioni. Se occorrono più di 5 minuti, contatta Supporto.

> **Important**
>
> Se la policy include solo le autorizzazioni dell'indice (e nessuna autorizzazione della raccolta), potresti comunque visualizzare un messaggio per le raccolte corrispondenti che indica `Collection cannot be accessed yet. Configure data access policies so that users can access the data within this collection`. Tale avviso si può ignorare. I principali autorizzati possono comunque eseguire le operazioni relative all'indice assegnate sulla raccolta.

Creazione di policy di accesso ai dati (AWS CLI)

Per creare una politica di accesso ai dati utilizzando l'API OpenSearch Serverless, usa il `CreateAccessPolicy` comando. Il comando accetta sia policy inline che file .json. Le policy inline devono essere codificate come una stringa con escape JSON.

La seguente richiesta crea una policy di accesso ai dati:

```
aws opensearchserverless create-access-policy \
    --name marketing \
    --type data \
    --policy "[{\"Rules\":[{\"ResourceType\":\"collection\",\"Resource\":
[\"collection/autopartsinventory\",\"collection/sales*\"],\"Permission\":
[\"aoss:UpdateCollectionItems\"]},{\"ResourceType\":\"index\",\"Resource\":
[\"index/autopartsinventory/*\",\"index/salesorders/orders*\"],\"Permission
\":[\"aoss:ReadDocument\",\"aoss:DescribeIndex\"]}],\"Principal\":
[\"arn:aws:iam::123456789012:user/Shaheen\"]}]"
```

Per fornire la policy all'interno di un file .json, utilizza il formato `--policy file://my-policy.json`.

I principali inclusi nella policy possono ora utilizzare le OpenSearch operazioni a cui hanno avuto accesso.

Visualizzazione di policy di accesso ai dati

Prima di creare una raccolta, hai la possibilità di visualizzare in anteprima le policy di accesso ai dati esistenti nel tuo account per vedere quali hanno un modello di risorse che corrisponde al nome della raccolta. La seguente ListAccessPoliciesrichiesta elenca tutte le politiche di accesso ai dati del tuo account:

```
aws opensearchserverless list-access-policies --type data
```

La richiesta restituisce informazioni su tutte le policy di accesso ai dati configurate. Per visualizzare le regole del modello definite in una politica specifica, trova le informazioni sulla politica nel contenuto dell'`accessPolicySummaries`elemento nella risposta. Prendi nota della `name` fine `type` di questa politica e utilizza queste proprietà in una GetAccessPolicyrichiesta per ricevere una risposta con i seguenti dettagli della politica:

```
{
    "accessPolicyDetails": [
        {
            "type": "data",
            "name": "my-policy",
            "policyVersion": "MTY2NDA1NDE4MDg1OF8x",
            "description": "My policy",
            "policy": "[{\"Rules\":[{\"ResourceType\":\"collection\",
\"Resource\":[\"collection/autopartsinventory\",\"collection/sales*\"],
```

```
\"Permission\":[\"aoss:UpdateCollectionItems\"]},{\"ResourceType\":\"index\",
\"Resource\":[\"index/autopartsinventory/*\",\"index/salesorders/orders*\"],
\"Permission\":[\"aoss:ReadDocument\",\"aoss:DescribeIndex\"]}],\"Principal\":
[\"arn:aws:iam::123456789012:user/Shaheen\"]}]",
            "createdDate": 1664054180858,
            "lastModifiedDate": 1664054180858
        }
    ]
}
```

Puoi includere filtri di risorse per limitare i risultati alle policy che contengono raccolte o indici specifici:

```
aws opensearchserverless list-access-policies --type data --resource
 "index/autopartsinventory/*"
```

Per visualizzare i dettagli su una politica specifica, utilizzare il GetAccessPolicycomando.

Aggiornamento delle policy di accesso ai dati

Quando aggiorni una policy di accesso ai dati, vengono interessate anche tutte le altre raccolte associate. Per aggiornare una politica di accesso ai dati nella console OpenSearch Serverless, scegli Controllo dell'accesso ai dati, seleziona la politica da modificare e scegli Modifica. Apporta le modifiche necessarie, quindi scegli Save (Salva).

Per aggiornare una politica di accesso ai dati utilizzando l'API OpenSearch Serverless, invia una `UpdateAccessPolicy` richiesta. È necessario includere una versione della policy, che è possibile recuperare utilizzando i comandi `ListAccessPolicies` o `GetAccessPolicy`. L'inclusione della versione più recente delle policy garantisce di non sovrascrivere inavvertitamente una modifica apportata da qualcun altro.

La seguente UpdateAccessPolicyrichiesta aggiorna una politica di accesso ai dati con un nuovo documento JSON di policy:

```
aws opensearchserverless update-access-policy \
    --name sales-inventory \
    --type data \
    --policy-version MTY2NDA1NDE4MDg1OF8x \
    --policy file://my-new-policy.json
```

Potrebbero verificarsi alcuni minuti di ritardo tra l'aggiornamento della policy e il momento in cui vengono applicate le nuove autorizzazioni.

Eliminazione delle policy di accesso ai dati

Quando si elimina una policy di accesso ai dati, tutte le raccolte associate perdono l'accesso definito nella policy stessa. Prima di eliminare una policy, assicurati che i tuoi utenti IAM e SAML abbiano l'accesso appropriato alla raccolta. Per eliminare una policy nella console OpenSearch Serverless, seleziona la policy e scegli Elimina.

Puoi anche usare il DeleteAccessPolicycomando:

```
aws opensearchserverless delete-access-policy --name my-policy --type data
```

Accesso ai dati tra account

Sebbene non sia possibile creare una politica di accesso ai dati con identità o raccolte tra più account, puoi comunque configurare l'accesso su più account con l'opzione Assumi ruolo. Ad esempio, se *account-a* possiede una raccolta a cui *account-b* deve accedere, l'utente di *account-b* può assumere un ruolo in. *account-a* Il ruolo deve disporre delle autorizzazioni IAM `aoss:APIAccessAll` ed `aoss:DashboardsAccessAll` essere incluso nella politica di accesso ai dati su*account-a*.

Accesso al piano dati tramite AWS PrivateLink

Amazon OpenSearch Serverless supporta due tipi di AWS PrivateLink connessioni per le operazioni sul piano di controllo e sul piano dati. Le operazioni del piano di controllo includono la creazione e l'eliminazione di raccolte e la gestione delle politiche di accesso. Le operazioni del piano dati servono per l'indicizzazione e l'interrogazione dei dati all'interno di una raccolta. Questa pagina descrive gli endpoint VPC del piano dati. Per informazioni sugli AWS PrivateLink endpoint del piano di controllo, vedere. the section called “Endpoint VPC del piano di controllo”

Puoi usarlo AWS PrivateLink per creare una connessione privata tra il tuo VPC e Amazon OpenSearch Serverless. Puoi accedere a OpenSearch Serverless come se fosse nel tuo VPC, senza l'uso di un gateway Internet, un dispositivo NAT, una connessione VPN o una connessione. Direct Connect Le istanze nel tuo VPC non necessitano di indirizzi IP pubblici per accedere a OpenSearch Serverless. Per ulteriori informazioni sull'accesso alla rete VPC, consulta Modelli di connettività di rete per Amazon OpenSearch Serverless.

Stabilisci questa connessione privata creando un endpoint di interfaccia attivato da AWS PrivateLink. In ciascuna sottorete viene creata un'interfaccia di rete endpoint da specificare per l'endpoint di

interfaccia. Si tratta di interfacce di rete gestite dai richiedenti che fungono da punto di ingresso per il traffico destinato a Serverless. OpenSearch

Per ulteriori informazioni, consulta la sezione Accesso a Servizi AWS tramite AWS PrivateLink nella Guida di AWS PrivateLink .

Argomenti

- Risoluzione DNS degli endpoint di raccolta
- VPCs e politiche di accesso alla rete
- VPCs e politiche relative agli endpoint
- Considerazioni
- Autorizzazioni richieste
- Crea un endpoint di interfaccia per Serverless OpenSearch
- Configurazione VPC condivisa per Amazon Serverless OpenSearch

Risoluzione DNS degli endpoint di raccolta

Quando crei un endpoint VPC del piano dati tramite la console OpenSearch Serverless, il servizio crea una nuova zona ospitata Amazon Route 53 privata e la collega al VPC. Questa zona ospitata privata è costituita da un record per risolvere il record DNS wildcard per le raccolte OpenSearch Serverless (`*.us-east-1.aoss.amazonaws.com`) negli indirizzi di interfaccia utilizzati per l'endpoint. È sufficiente un solo endpoint VPC OpenSearch serverless in un VPC per accedere a tutte le raccolte e i dashboard di ciascuno di essi. Regione AWS Ogni VPC con un endpoint per OpenSearch Serverless ha una propria zona ospitata privata collegata.

L'endpoint di interfaccia OpenSearch Serverless crea anche un record DNS wildcard pubblico di Route 53 per tutte le raccolte nella regione. Il nome DNS viene risolto negli indirizzi IP pubblici Serverless. OpenSearch I client VPCs che non dispongono di un endpoint VPC OpenSearch Serverless o i client in reti pubbliche possono utilizzare il resolver pubblico Route 53 e accedere alle raccolte e ai dashboard con tali indirizzi IP. Il tipo di indirizzo IP (IPv4o Dualstack) dell'endpoint VPC viene determinato in base alle sottoreti fornite quando si crea un endpoint di interfaccia per Serverless. IPv6 OpenSearch

> **Note**
>
> OpenSearch Serverless crea una zona ospitata privata (`<region>.opensearch.amazonaws.com`) Amazon Route 53 aggiuntiva per la

risoluzione di qualsiasi dominio di OpenSearch servizio. Puoi aggiornare il tuo endpoint IPv4 VPC esistente a Dualstack utilizzando il comando in. update-vpc-endpoint AWS CLI

L'indirizzo del resolver DNS per un determinato VPC è il secondo indirizzo IP del VPC CIDR. Qualsiasi client nel VPC deve utilizzare quel resolver per ottenere l'indirizzo dell'endpoint VPC per qualsiasi raccolta. Il resolver utilizza una zona ospitata privata creata da Serverless. OpenSearch È sufficiente utilizzare quel resolver per tutte le raccolte di qualsiasi account. È anche possibile utilizzare il resolver VPC per alcuni endpoint di raccolta e il resolver pubblico per altri, sebbene in genere non sia necessario.

VPCs e politiche di accesso alla rete

Per concedere le autorizzazioni di rete OpenSearch APIs e i dashboard per le tue raccolte, puoi utilizzare le policy di accesso alla rete OpenSearch Serverless. Puoi controllare questo accesso alla rete dai tuoi endpoint VPC o dalla rete Internet pubblica. Poiché la policy di rete controlla solo le autorizzazioni relative al traffico, è necessario impostare anche una policy di accesso ai dati che specifichi l'autorizzazione a operare sui dati di una raccolta e sui relativi indici. Pensate a un endpoint VPC OpenSearch serverless come punto di accesso al servizio, a una policy di accesso alla rete come punto di accesso a livello di rete a raccolte e dashboard e a una politica di accesso ai dati come punto di accesso per il controllo granulare degli accessi per qualsiasi operazione sui dati della raccolta.

Poiché puoi specificare più endpoint VPC IDs in una policy di rete, ti consigliamo di creare un endpoint VPC per ogni VPC che deve accedere a una raccolta. Questi VPCs possono appartenere a AWS account diversi rispetto all'account che possiede la raccolta e la policy di rete OpenSearch Serverless. Non è consigliabile creare una soluzione di VPC-to-VPC peering o altra soluzione di proxy tra due account in modo che il VPC di un account possa utilizzare l'endpoint VPC di un altro account. Si tratta di una soluzione meno sicura ed economica rispetto a ogni VPC con il proprio endpoint. Il primo VPC non sarà facilmente visibile all'amministratore dell'altro VPC, che ha impostato l'accesso all'endpoint di quel VPC nella policy di rete.

VPCs e politiche relative agli endpoint

Amazon OpenSearch Serverless supporta le policy degli endpoint per. VPCs Una policy per gli endpoint è una policy basata sulle risorse IAM che colleghi a un endpoint VPC per controllare quali AWS responsabili possono utilizzare l'endpoint per accedere al tuo servizio. AWS Per ulteriori informazioni, consulta Controllare l'accesso agli endpoint VPC utilizzando le policy degli endpoint.

Per utilizzare una policy per gli endpoint, devi prima creare un endpoint di interfaccia. È possibile creare un endpoint di interfaccia utilizzando la console Serverless o l' OpenSearch API Serverless. OpenSearch Dopo aver creato l'endpoint di interfaccia, dovrai aggiungere la policy dell'endpoint all'endpoint. Per ulteriori informazioni, consulta the section called “Creazione di un endpoint di interfaccia”.

> **Note**
>
> Non è possibile definire una policy per gli endpoint direttamente nella console di servizio. OpenSearch

Una policy per gli endpoint non sostituisce né sostituisce altre politiche basate sull'identità, politiche basate sulle risorse, politiche di rete o politiche di accesso ai dati che potresti aver configurato. Per ulteriori informazioni sull'aggiornamento delle policy degli endpoint, consulta Controllare l'accesso agli endpoint VPC utilizzando le policy degli endpoint.

Per impostazione predefinita, una policy per gli endpoint garantisce l'accesso completo all'endpoint VPC.

```
{
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": "*",
            "Action": "*",
            "Resource": "*"
        }
    ]
}
```

Sebbene la policy degli endpoint VPC predefinita garantisca l'accesso completo agli endpoint, puoi configurare una policy degli endpoint VPC per consentire l'accesso a ruoli e utenti specifici. A tale scopo, consulta l'esempio seguente:

JSON

```
{
    "Version":"2012-10-17",
```

```
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": [
                    "123456789012",
                    "987654321098"
                ]
            },
            "Action": "*",
            "Resource": "*"
        }
    ]
}
```

Puoi specificare una raccolta OpenSearch Serverless da includere come elemento condizionale nella tua policy degli endpoint VPC. A tale scopo, consulta l'esempio seguente:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": "*",
            "Action": "*",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aoss:collection": [
                        "coll-abc"
                    ]
                }
            }
        }
    ]
}
```

Il supporto per `aoss:CollectionId` è supportato.

```
Condition": {
         "StringEquals": {
               "aoss:CollectionId": "collection-id"
          }
}
```

Puoi utilizzare le identità SAML nella tua policy degli endpoint VPC per determinare l'accesso agli endpoint VPC. È necessario utilizzare un carattere jolly (*) nella sezione principale della policy degli endpoint VPC. A tale scopo, consulta l'esempio seguente:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": "*",
            "Action": "*",
            "Resource": "*",
            "Condition": {
                "ForAnyValue:StringEquals": {
                    "saml:cn": [
                        "saml/111122223333/idp123/group/football",
                        "saml/111122223333/idp123/group/soccer",
                        "saml/111122223333/idp123/group/cricket"
                    ]
                }
            }
        }
    ]
}
```

Inoltre, puoi configurare la tua policy sugli endpoint per includere una politica principale SAML specifica. A tale scopo, consulta quanto segue:

JSON

```
{
```

```
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": "*",
            "Action": "*",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:PrincipalTag/Department": [
                        "Engineering"]
                    }
                }
            }
        ]
    }
```

Per ulteriori informazioni sull'utilizzo dell'autenticazione SAML con Amazon OpenSearch Serverless, consulta Autenticazione SAML per Amazon Serverless. OpenSearch

Puoi anche includere utenti IAM e SAML nella stessa policy degli endpoint VPC. A tale scopo, consulta l'esempio seguente:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": "*",
            "Action": "*",
            "Resource": "*",
            "Condition": {
                "ForAnyValue:StringEquals": {
                    "saml:cn": [
                        "saml/111122223333/idp123/group/football",
                        "saml/111122223333/idp123/group/soccer",
                        "saml/111122223333/idp123/group/cricket"
                    ]
```

```
                }
            }
        },
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": [
                    "111122223333"
                ]
            },
            "Action": "*",
            "Resource": "*"
        }
    ]
}
```

Puoi anche accedere a una raccolta Amazon OpenSearch Serverless da Amazon EC2 tramite endpoint VPC di interfaccia. Per ulteriori informazioni, consulta Accedere a una raccolta OpenSearch serverless da Amazon EC2 (tramite endpoint VPC di interfaccia).

Considerazioni

Prima di configurare un endpoint di interfaccia per OpenSearch Serverless, considera quanto segue:

- OpenSearch Serverless supporta l'esecuzione di chiamate a tutte le operazioni OpenSearch API supportate (non le operazioni API di configurazione) tramite l'endpoint dell'interfaccia.
- Dopo aver creato un endpoint di interfaccia per OpenSearch Serverless, è comunque necessario includerlo nelle politiche di accesso alla rete per consentirgli di accedere alle raccolte serverless.
- Per impostazione predefinita, l'accesso completo a OpenSearch Serverless è consentito tramite l'endpoint dell'interfaccia. È possibile associare un gruppo di sicurezza alle interfacce di rete degli endpoint per controllare il traffico verso OpenSearch Serverless attraverso l'endpoint dell'interfaccia.
- Un singolo dispositivo Account AWS può avere un massimo di 50 endpoint VPC OpenSearch serverless.
- Se abiliti l'accesso pubblico a Internet all'API o alle dashboard della tua raccolta in una politica di rete, la raccolta è accessibile da qualsiasi VPC e dalla rete Internet pubblica.
- Se sei in locale e all'esterno del VPC, non puoi utilizzare direttamente un resolver DNS per OpenSearch la risoluzione degli endpoint VPC Serverless. Se hai bisogno di un accesso VPN,

il VPC necessita di un resolver proxy DNS per l'utilizzo da parte di client esterni. Route 53 offre un'opzione di endpoint in entrata che puoi utilizzare per risolvere le query DNS sul tuo VPC dalla tua rete locale o da un altro VPC.

- La zona ospitata privata che OpenSearch Serverless crea e collega al VPC è gestita dal servizio, ma compare nelle tue Amazon Route 53 risorse e viene fatturata sul tuo account.
- Per altre considerazioni, consulta la sezione Considerazioni nella Guida AWS PrivateLink .

Autorizzazioni richieste

L'accesso VPC per OpenSearch Serverless utilizza le seguenti autorizzazioni AWS Identity and Access Management (IAM). È possibile specificare le condizioni IAM per limitare gli utenti a raccolte specifiche.

- `aoss:CreateVpcEndpoint` - Creazione di un endpoint VPC.
- `aoss:ListVpcEndpoints` - Elencazione di tutti gli endpoint VPC.
- `aoss:BatchGetVpcEndpoint` - Visualizzazione dei dettagli su un sottoinsieme di endpoint VPC.
- `aoss:UpdateVpcEndpoint` - Modifica di un endpoint VPC.
- `aoss:DeleteVpcEndpoint` - Eliminazione di un endpoint VPC.

Inoltre, sono necessarie le seguenti autorizzazioni Amazon EC2 e Route 53 per creare un endpoint VPC.

- `ec2:CreateTags`
- `ec2:CreateVpcEndpoint`
- `ec2:DeleteVpcEndPoints`
- `ec2:DescribeSecurityGroups`
- `ec2:DescribeSubnets`
- `ec2:DescribeVpcEndpoints`
- `ec2:DescribeVpcs`
- `ec2:ModifyVpcEndPoint`
- `route53:AssociateVPCWithHostedZone`
- `route53:ChangeResourceRecordSets`
- `route53:CreateHostedZone`

- `route53:DeleteHostedZone`
- `route53:GetChange`
- `route53:GetHostedZone`
- `route53:ListHostedZonesByName`
- `route53:ListHostedZonesByVPC`
- `route53:ListResourceRecordSets`

Crea un endpoint di interfaccia per Serverless OpenSearch

Puoi creare un endpoint di interfaccia per OpenSearch Serverless utilizzando la console o l'API Serverless. OpenSearch

Per creare un endpoint di interfaccia per una raccolta Serverless OpenSearch

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Nel pannello di navigazione a sinistra, espandi Serverless e seleziona endpoint VPC.
3. Scegli Create VPC endpoint (Crea endpoint VPC).
4. Fornisci un nome per l'endpoint.
5. Per VPC, seleziona il VPC da cui accederai Serverless. OpenSearch
6. Per le sottoreti, seleziona una sottorete da cui accederai a Serverless. OpenSearch
   - L'indirizzo IP e il tipo DNS dell'endpoint si basano sul tipo di sottorete
     - Dualstack: se tutte le sottoreti hanno entrambi gli intervalli di indirizzi IPv4 IPv6
     - IPv6: Se tutte le sottoreti sono solo sottoreti IPv6
     - IPv4: Se tutte le sottoreti hanno intervalli di indirizzi IPv4
7. Per Security groups (Gruppi di sicurezza), seleziona i gruppi di sicurezza da associare alle interfacce di rete dell'endpoint. Questo è un passaggio fondamentale in cui limiti le porte, i protocolli e le fonti per il traffico in entrata che stai autorizzando verso il tuo endpoint. Assicurati che le regole del gruppo di sicurezza consentano alle risorse che utilizzeranno l'endpoint VPC di comunicare con OpenSearch Serverless di comunicare con l'interfaccia di rete dell'endpoint.
8. Seleziona Crea endpoint.

Per creare un endpoint VPC utilizzando l'API OpenSearch Serverless, usa il comando. `CreateVpcEndpoint`

> **Note**
>
> Dopo aver creato un endpoint, annotane l'ID (ad esempio, `vpce-abc123def4EXAMPLE`). Per fornire all'endpoint l'accesso alle raccolte, è necessario includere questo ID in una o più policy di accesso alla rete.

Dopo aver creato un endpoint di interfaccia, è necessario fornirgli l'accesso alle raccolte tramite policy di accesso alla rete. Per ulteriori informazioni, consulta the section called “Accesso alla rete”.

Configurazione VPC condivisa per Amazon Serverless OpenSearch

Puoi utilizzare Amazon Virtual Private Cloud (VPC) per condividere sottoreti VPC con altri Account AWS membri dell'organizzazione, nonché condividere infrastrutture di rete come una VPN tra più risorse. Account AWS

Attualmente, Amazon OpenSearch Serverless non supporta la creazione di una AWS PrivateLink connessione a un VPC condiviso a meno che tu non sia il proprietario di quel VPC. AWS PrivateLink inoltre non supporta la condivisione di connessioni tra. Account AWS

Tuttavia, sulla base dell'architettura flessibile e modulare di OpenSearch Serverless, puoi comunque configurare un VPC condiviso. Questo perché l'infrastruttura di rete OpenSearch Serverless è separata da quella dell'infrastruttura di raccolta individuale (OpenSearch Service). È quindi possibile creare un AWS PrivateLink VPCe endpoint per un account in cui si trova un VPC e quindi utilizzare VPCe un ID nella politica di rete di altri account per limitare il traffico proveniente solo da quel VPC condiviso.

Le seguenti procedure si riferiscono a un account proprietario e a un account consumatore.

Un account proprietario funge da account di rete comune in cui è possibile configurare un VPC e condividerlo con altri account. Gli account consumer sono quegli account che creano e gestiscono le proprie raccolte OpenSearch Serverless nel VPC condiviso con loro dall'account proprietario.

Prerequisiti

Assicurati che siano soddisfatti i seguenti requisiti prima di configurare il VPC condiviso:

- L'account proprietario desiderato deve aver già configurato un VPC, sottoreti, tabella di routing e altre risorse richieste in Amazon Virtual Private Cloud. Per ulteriori informazioni, consulta la Guida utente Amazon VPC.

- L'account proprietario e gli account consumer previsti devono appartenere alla stessa organizzazione in. AWS Organizations Per ulteriori informazioni, consulta la Guida per l'utente AWS Organizations.

Per configurare un VPC condiviso in un account di account/common rete proprietario.

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Segui la procedura riportata in the section called “Creazione di un endpoint di interfaccia”. Mentre lo fai, effettua le seguenti selezioni:
   - Seleziona un VPC e delle sottoreti condivise con gli account consumer della tua organizzazione.
3. Dopo aver creato l'endpoint, prendi nota dell' VPCe ID generato e forniscilo agli amministratori che devono eseguire l'attività di configurazione negli account consumer.

   VPCe IDs sono nel formato. `vpce-abc123def4EXAMPLE`

Per configurare un VPC condiviso in un account consumer

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Usa le informazioni the section called “Gestione delle raccolte” per creare una collezione, se non ne hai già una.
3. Usa le informazioni contenute the section called “Creazione di policy di rete (console)” per creare una politica di rete. Mentre lo fai, effettua le seguenti selezioni.

   > **Note**
   >
   > A tale scopo è inoltre possibile aggiornare una politica di rete esistente.

   a. Per Tipo di accesso, seleziona VPC (consigliato).

   b. Per gli endpoint VPC per l'accesso, scegli l' VPCe ID fornito dall'account proprietario, nel formato. `vpce-abc123def4EXAMPLE`

   c. Nell'area Tipo di risorsa, procedi come segue:

- Seleziona la casella Abilita l'accesso all' OpenSearchendpoint, quindi seleziona il nome o il modello di raccolta da utilizzare per abilitare l'accesso da quel VPC condiviso.
- Seleziona la casella Abilita l'accesso alla OpenSearch dashboard, quindi seleziona il nome o il modello di raccolta da utilizzare per abilitare l'accesso da quel VPC condiviso.

4. Per una nuova politica, scegli Crea. Per una politica esistente, scegli Aggiorna.

Controlla l'accesso al piano tramite AWS PrivateLink

Amazon OpenSearch Serverless supporta due tipi di AWS PrivateLink connessioni per le operazioni sul piano di controllo e sul piano dati. Le operazioni del piano di controllo includono la creazione e l'eliminazione di raccolte e la gestione delle politiche di accesso. Le operazioni sul piano dati servono per l'indicizzazione e l'interrogazione dei dati all'interno di una raccolta. Questa pagina copre l'endpoint del piano di controllo. AWS PrivateLink Per informazioni sugli endpoint VPC del piano dati, vedere. the section called “Endpoint VPC del piano dati”

Creazione di un endpoint del piano di controllo AWS PrivateLink

Puoi migliorare il livello di sicurezza del tuo VPC OpenSearch configurando Serverless per utilizzare un endpoint VPC di interfaccia. Gli endpoint di interfaccia sono alimentati da. AWS PrivateLink Questa tecnologia consente di accedere privatamente a OpenSearch Serverless APIs senza un gateway Internet, un dispositivo NAT, una connessione VPN o una connessione AWS Direct Connect.

Per ulteriori informazioni sugli AWS PrivateLink endpoint VPC, consulta la sezione Endpoint VPC nella Amazon VPC User Guide.

Considerazioni

- Gli endpoint VPC sono supportati solo all'interno della stessa regione.
- Gli endpoint VPC supportano solo il DNS fornito da Amazon tramite Amazon Route 53.
- Gli endpoint VPC supportano le policy degli endpoint per controllare l'accesso a OpenSearch Serverless Collections, Policies e. VpcEndpoints
- OpenSearch Serverless supporta solo gli endpoint di interfaccia. Gli endpoint gateway non sono supportati.

Creazione dell'endpoint VPC

Per creare l'endpoint VPC del piano di controllo per Amazon OpenSearch Serverless, utilizza la procedura Access an AWS service using an interface VPC endpoint nella Amazon VPC Developer Guide. Crea il seguente endpoint:

- `com.amazonaws.region.aoss`

Per creare un endpoint VPC del piano di controllo utilizzando la console

1. Apri la console Amazon VPC all'indirizzo https://console.aws.amazon.com/vpc/.
2. Nel pannello di navigazione, seleziona Endpoints (Endpoint).
3. Scegliere Create Endpoint (Crea endpoint).
4. Per Service category (Categoria servizio), scegli Servizi AWS.
5. Per Servizi, scegli. `com.amazonaws.region.aoss` Ad esempio, `com.amazonaws.us-east-1.aoss`.
6. Per VPC, scegli un VPC in cui creare l'endpoint.
7. In Subnets (Sottoreti), scegliere le sottoreti (zone di disponibilità) in cui creare le interfacce di rete degli endpoint.
8. Per i gruppi di sicurezza, scegli i gruppi di sicurezza da associare alle interfacce di rete degli endpoint. Assicurati che HTTPS (porta 443) sia consentito.
9. Per Policy, scegli Accesso completo per consentire tutte le operazioni oppure scegli Personalizzato per allegare una politica personalizzata.
10. Seleziona Crea endpoint.

Creazione di una policy per gli endpoint

Puoi allegare una policy per gli endpoint al tuo endpoint VPC che controlla l'accesso ad Amazon Serverless. OpenSearch La policy specifica le informazioni riportate di seguito:

- Il principale che può eseguire operazioni.
- Le azioni che possono essere eseguite.
- Le risorse sui cui si possono eseguire operazioni.

Per ulteriori informazioni, consulta Controllo degli accessi ai servizi con endpoint VPC in Guida per l'utente di Amazon VPC.

Example Policy degli endpoint VPC per Serverless OpenSearch

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": "*",
      "Action": [
        "aoss:ListCollections",
        "aoss:BatchGetCollection"
      ],
      "Resource": "*"
    }
  ]
}
```

Example Politica restrittiva che consente solo le operazioni di elenco

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": "*",
      "Action": "aoss:ListCollections",
      "Resource": "*"
    }
  ]
}
```

Autenticazione SAML per Amazon Serverless OpenSearch

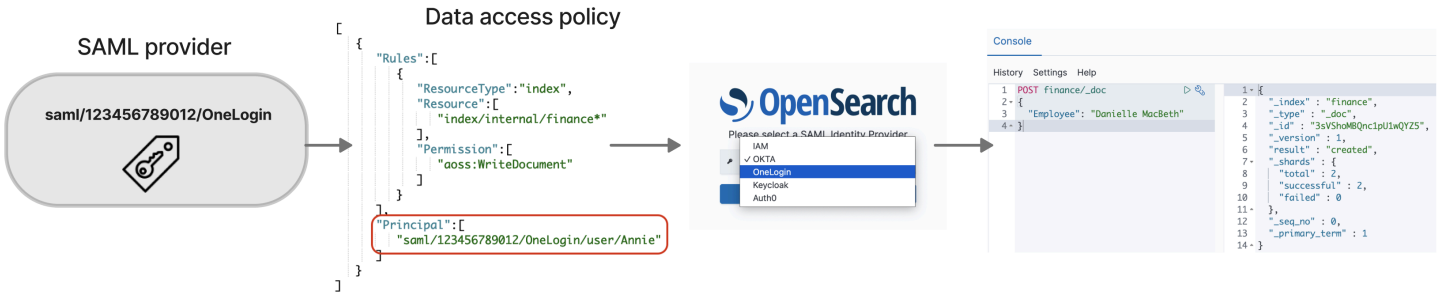
Con l'autenticazione SAML per Amazon OpenSearch Serverless, puoi utilizzare il tuo provider di identità esistente per offrire il Single Sign-On (SSO) per gli endpoint OpenSearch Dashboards delle raccolte serverless.

L'autenticazione SAML consente di utilizzare provider di identità di terze parti per accedere alle dashboard per indicizzare e cercare dati. OpenSearch OpenSearch Serverless supporta i provider

che utilizzano lo standard SAML 2.0, come IAM Identity Center, Okta, Keycloak, Active Directory Federation Services (AD FS) e Auth0. Puoi configurare IAM Identity Center per sincronizzare utenti e gruppi da altre fonti di identità come Okta e Microsoft Entra ID. OneLogin Per un elenco delle fonti di identità supportate da IAM Identity Center e i passaggi per configurarle, consulta i tutorial introduttivi nella Guida per l'utente di IAM Identity Center.

> **Note**
>
> L'autenticazione SAML serve solo per accedere alle OpenSearch dashboard tramite un browser web. Gli utenti autenticati possono effettuare richieste alle operazioni OpenSearch API solo tramite Dev Tools in Dashboards. OpenSearch Le tue credenziali SAML non ti consentono di effettuare richieste HTTP dirette alle operazioni API. OpenSearch

Per configurare l'autenticazione SAML, è necessario configurare prima un provider di identità (IdP) SAML. Quindi includi uno o più utenti di quell'IdP in una policy di accesso ai dati. Questa politica concede determinate autorizzazioni per gli indici delle raccolte. and/or Un utente può quindi accedere alle OpenSearch dashboard ed eseguire le azioni consentite nella politica di accesso ai dati.



Argomenti

- Considerazioni
- Autorizzazioni richieste
- Creazione di provider SAML (console)
- Accesso ai pannelli OpenSearch di controllo
- Concessione alle identità SAML dell'accesso ai dati della raccolta
- Creazione di provider SAML (AWS CLI)
- Visualizzazione di provider SAML
- Aggiornamento dei provider SAML

- Eliminazione di provider SAML

Considerazioni

Durante la configurazione dell'autenticazione SAML tieni presente quanto segue:

- Le richieste firmate e crittografate non sono supportate.
- Le asserzioni crittografate non sono supportate.
- L'autenticazione e la disconnessione avviate dal'IdP non sono supportate.
- Le policy di controllo dei servizi (SCP) non saranno applicabili o valutate nel caso di identità non IAM (come SAML in Amazon OpenSearch Serverless e SAML e l'autorizzazione utente interna di base per Amazon Service). OpenSearch

Autorizzazioni richieste

L'autenticazione SAML per OpenSearch Serverless utilizza le seguenti autorizzazioni (IAM): AWS Identity and Access Management

- `aoss:CreateSecurityConfig`: crea un provider SAML.
- `aoss:ListSecurityConfig`: elenca tutti i provider SAML nell'account corrente.
- `aoss:GetSecurityConfig`: visualizza le informazioni sul provider SAML.
- `aoss:UpdateSecurityConfig`: modifica una determinata configurazione del provider SAML, inclusi i metadati XML.
- `aoss:DeleteSecurityConfig`: elimina un provider SAML.

La seguente policy di accesso basata sull'identità consente a un utente di gestire tutte le configurazioni dell'IdP:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
```

```
                "aoss:CreateSecurityConfig",
                "aoss:DeleteSecurityConfig",
                "aoss:GetSecurityConfig",
                "aoss:UpdateSecurityConfig",
                "aoss:ListSecurityConfigs"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

Nota bene: l'elemento `Resource` deve essere un carattere jolly.

Creazione di provider SAML (console)

Queste fasi spiegano come creare provider SAML. Ciò consente l'autenticazione SAML con l'autenticazione avviata dal provider di servizi (SP) per le dashboard. OpenSearch L'autenticazione avviata dall'IdP non è supportata.

Per abilitare l'autenticazione SAML per le dashboard OpenSearch

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel pannello di navigazione a sinistra, espandi Serverless e scegli SAML authentication (Autenticazione SAML).
3. Scegli Add SAML provider (Aggiungi provider SAML).
4. Fornisci un nome e una descrizione per il provider.

> **Note**
>
> Il nome specificato è accessibile pubblicamente e verrà visualizzato in un menu a discesa quando gli utenti accedono alle OpenSearch dashboard. Assicurati che il nome sia facilmente riconoscibile e non riveli informazioni sensibili sul tuo provider di identità.

5. In Configure your IdP (Configura il tuo IdP), copia l'URL Assertion consumer service (ACS).
6. Utilizza l'URL ACS che hai appena copiato per configurare il provider di identità. La terminologia e le fasi variano in base al provider. Consultare la documentazione del provider.

In Okta, ad esempio, crei una "applicazione Web SAML 2.0" e specifichi l'URL ACS come l'URL di Single Sign-On, l'URL del destinatario e l'URL di destinazione. Per Auth0, lo specifichi in Allowed Callback. URLs

7. Se il tuo IdP prevede un campo apposito, fornisci la restrizione per il pubblico. La restrizione per il pubblico è un valore all'interno dell'asserzione SAML che specifica a chi è destinata l'asserzione. Con OpenSearch Serverless, puoi fare quanto segue. Assicurati di sostituire il codice riportato *content* nel seguente esempio di codice con il tuo Account AWS ID:

   1. Usa la restrizione `:opensearch:`*111122223333* predefinita per il pubblico.
   2. (Facoltativo) configura una restrizione personalizzata per il pubblico utilizzando. AWS CLI Per ulteriori informazioni, consulta Creazione di provider SAML (AWS CLI).

   Il nome del campo della restrizione per il pubblico varia in base al provider. Per Okta è l'URI del pubblico (ID entità SP). Per IAM Identity Center è il pubblico SAML dell'applicazione.
8. Se si utilizza IAM Identity Center, è necessario anche specificare la seguente mappatura degli attributi: `Subject=${user:name}`, con un formato `unspecified`.
9. Dopo aver configurato il provider di identità, viene generato un file di metadati IdP. Questo file XML contiene informazioni sul provider, ad esempio un certificato TLS, endpoint Single Sign-On e l'ID entità del provider di identità.

   Copia il testo nel file di metadati dell'IdP e incollalo nel campo Provide metadata from your IdP (Fornisci metadati dal tuo IdP). In alternativa, scegliere Importa da file XML e caricare il file. Il file dei metadati dovrebbe avere un aspetto simile al seguente:

```
<?xml version="1.0" encoding="UTF-8"?>
<md:EntityDescriptor entityID="entity-id"
 xmlns:md="urn:oasis:names:tc:SAML:2.0:metadata">
  <md:IDPSSODescriptor WantAuthnRequestsSigned="false"
 protocolSupportEnumeration="urn:oasis:names:tc:SAML:2.0:protocol">
    <md:KeyDescriptor use="signing">
      <ds:KeyInfo xmlns:ds="http://www.w3.org/2000/09/xmldsig#">
        <ds:X509Data>
          <ds:X509Certificate>tls-certificate</ds:X509Certificate>
        </ds:X509Data>
      </ds:KeyInfo>s
    </md:KeyDescriptor>
    <md:NameIDFormat>urn:oasis:names:tc:SAML:1.1:nameid-format:unspecified</
md:NameIDFormat>
```

```
    <md:NameIDFormat>urn:oasis:names:tc:SAML:1.1:nameid-format:emailAddress</
md:NameIDFormat>
    <md:SingleSignOnService Binding="urn:oasis:names:tc:SAML:2.0:bindings:HTTP-
POST" Location="idp-sso-url"/>
    <md:SingleSignOnService Binding="urn:oasis:names:tc:SAML:2.0:bindings:HTTP-
Redirect" Location="idp-sso-url"/>
  </md:IDPSSODescriptor>
</md:EntityDescriptor>
```

10. Mantieni vuoto il campo dell'attributo ID utente personalizzato per utilizzare l'`NameID`elemento dell'asserzione SAML per il nome utente. Se l'asserzione non utilizza questo elemento standard e include invece il nome utente come attributo personalizzato, specificare tale attributo qui. Gli attributi rispettano la distinzione tra maiuscole e minuscole. È supportato solo un singolo attributo dell'utente.

    L'esempio seguente mostra un attributo di sovrascrizione per `NameID` nell'asserzione SAML:

    ```
    <saml2:Attribute Name="UserId" NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-
    format:basic">
      <saml2:AttributeValue xmlns:xs="http://www.w3.org/2001/XMLSchema"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:type="xs:string">annie</saml2:AttributeValue>
    </saml2:Attribute>
    ```

11. (Facoltativo) Specifica un attributo personalizzato nel campo Group attribute (Attributo del gruppo), ad esempio `role` o `group`. È supportato solo un singolo attributo del gruppo. Non esiste un attributo del gruppo predefinito. Se non ne specifichi uno, le policy di accesso ai dati possono contenere solo utenti principali.

    L'esempio seguente mostra un attributo del gruppo nell'asserzione SAML:

    ```
    <saml2:Attribute Name="department"
     NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:basic">
        <saml2:AttributeValue xmlns:xs="http://www.w3.org/2001/XMLSchema"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:type="xs:string">finance</saml2:AttributeValue>
    </saml2:Attribute>
    ```

12. Per impostazione predefinita, OpenSearch Dashboards disconnette gli utenti dopo 24 ore. Puoi configurare questo valore su qualsiasi numero compreso tra 1 e 12 ore (15 e 720 minuti)

specificando il timeout del OpenSearch pannello di controllo. Se tenti di impostare il timeout uguale o inferiore a 15 minuti, la sessione verrà reimpostata a un'ora.

13. Scegli Create SAML provider (Crea provider SAML).

Accesso ai pannelli OpenSearch di controllo

Dopo aver configurato un provider SAML, tutti gli utenti e i gruppi associati a tale provider possono accedere all'endpoint OpenSearch Dashboards. L'URL Dashboards ha il formato *`collection-endpoint/_dashboards/`* per tutte le raccolte.

Se hai abilitato SAML, selezionando il link in basso verrai Console di gestione AWS indirizzato alla pagina di selezione IdP, dove puoi accedere utilizzando le tue credenziali SAML. Innanzitutto, utilizza il menu a discesa per selezionare un provider di identità:

Quindi accedi utilizzando le tue credenziali IdP.

Se non hai abilitato SAML, selezionando il link nelle Console di gestione AWS istruzioni potrai accedere come utente o ruolo IAM, senza alcuna opzione per SAML.

Concessione alle identità SAML dell'accesso ai dati della raccolta

Dopo aver creato un provider SAML, è comunque necessario concedere agli utenti e ai gruppi sottostanti l'accesso ai dati all'interno delle raccolte. L'accesso viene concesso tramite le policy di accesso ai dati. Finché non fornisci l'accesso agli utenti, questi non saranno in grado di leggere, scrivere o eliminare alcun dato contenuto nelle raccolte.

Per concedere l'accesso, crea una politica di accesso ai dati e specifica il tuo and/or gruppo IDs di utenti SAML nella dichiarazione: `Principal`

```
[
   {
      "Rules":[
       ...
      ],
      "Principal":[
         "saml/987654321098/myprovider/user/Shaheen",
         "saml/987654321098/myprovider/group/finance"
      ]
   }
]
```

Puoi concedere l'accesso a raccolte, a indici o a entrambi. Se desideri che utenti diversi dispongano di autorizzazioni diverse, crea più regole. Per un elenco di autorizzazioni disponibili, consulta la sezione Supported policy permissions (Autorizzazioni di policy supportate). Per ulteriori informazioni sul formato di una policy di accesso, consulta la sezione Policy syntax (Sintassi della policy).

Creazione di provider SAML (AWS CLI)

Per creare un provider SAML utilizzando l'API OpenSearch Serverless, invia una richiesta: CreateSecurityConfig

```
aws opensearchserverless create-security-config \
    --name myprovider \
    --type saml \
    --saml-options file://saml-auth0.json
```

Specifica `saml-options`, inclusi i metadati XML, come una mappa chiave-valore all'interno di un file .json. I metadati XML devono essere codificati come stringa con escape JSON.

```
{
   "sessionTimeout": 70,
   "groupAttribute": "department",
   "userAttribute": "userid",
   "openSearchServerlessEntityId": "aws:opensearch:111122223333:app1",
   "metadata": "EntityDescriptor xmlns=\"urn:oasis:names:tc:SAML:2.0:metadata
\" ... ... ... IDPSSODescriptor\r\n\/EntityDescriptor"
}
```

> **Note**
>
> (Facoltativo) configura una restrizione personalizzata per il pubblico utilizzando. AWS CLI Per ulteriori informazioni, consulta Creazione di provider SAML (AWS CLI).

Visualizzazione di provider SAML

La seguente ListSecurityConfigsrichiesta elenca tutti i provider SAML presenti nel tuo account:

```
aws opensearchserverless list-security-configs --type saml
```

La richiesta restituisce informazioni su tutti i provider SAML esistenti, inclusi i metadati completi dell'IdP generati dal provider di identità:

```
{
   "securityConfigDetails": [
      {
         "configVersion": "MTY2NDA1MjY4NDQ5M18x",
         "createdDate": 1664054180858,
         "description": "Example SAML provider",
         "id": "saml/111122223333/myprovider",
         "lastModifiedDate": 1664054180858,
         "samlOptions": {
            "groupAttribute": "department",
            "metadata": "EntityDescriptorxmlns=\"urn:oasis:names:tc:SAML:2.0:metadata
\" ..... ...IDPSSODescriptor\r\n/EntityDescriptor",
            "sessionTimeout": 120,
            "openSearchServerlessEntityId": "aws:opensearch:111122223333:app1",
            "userAttribute": "userid"
         }
      }
   ]
}
```

Per visualizzare i dettagli su un provider specifico, inclusa la `configVersion` per gli aggiornamenti futuri, invia una richiesta `GetSecurityConfig`.

Aggiornamento dei provider SAML

Per aggiornare un provider SAML utilizzando la console OpenSearch Serverless, scegli l'autenticazione SAML, seleziona il tuo provider di identità e scegli Modifica. Puoi modificare tutti i campi, inclusi i metadati e gli attributi personalizzati.

Per aggiornare un provider tramite l'API OpenSearch Serverless, invia una UpdateSecurityConfigrichiesta e includi l'identificatore della politica da aggiornare. Inoltre, è necessario includere una versione di configurazione, che è possibile recuperare utilizzando i comandi `ListSecurityConfigs` o `GetSecurityConfig`. L'inclusione della versione più recente garantisce di non sovrascrivere inavvertitamente una modifica apportata da qualcun altro.

La seguente richiesta aggiorna le opzioni SAML per un provider:

```
aws opensearchserverless update-security-config \
    --id saml/123456789012/myprovider \
    --type saml \
    --saml-options file://saml-auth0.json \
    --config-version MTY2NDA1MjY4NDQ5M18x
```

Specifica le opzioni di configurazione SAML come mappa chiave-valore all'interno di un file .json.

> **Important**
>
> Gli aggiornamenti alle opzioni SAML non sono incrementali. Se non si specifica un valore per un parametro nell'oggetto `SAMLOptions` quando si effettua un aggiornamento, i valori esistenti verranno sostituiti da valori vuoti. Ad esempio, se la configurazione corrente contiene un valore per `userAttribute`, e poi si effettua un aggiornamento e non lo si include, il valore viene rimosso dalla configurazione. Prima di effettuare un aggiornamento, assicurati di conoscere i valori esistenti richiamando l'operazione `GetSecurityConfig`.

Eliminazione di provider SAML

Quando elimini un provider SAML, qualsiasi riferimento a utenti e gruppi associati nelle policy di accesso ai dati non è più funzionale. Per evitare confusione, nelle policy di accesso ti suggeriamo di rimuovere tutti i riferimenti all'endpoint prima di eliminarlo.

Per eliminare un provider SAML utilizzando la console OpenSearch Serverless, scegli Autenticazione, seleziona il provider e scegli Elimina.

Per eliminare un provider tramite l'API OpenSearch Serverless, invia una richiesta: DeleteSecurityConfig

```
aws opensearchserverless delete-security-config --id saml/123456789012/myprovider
```

Convalida della conformità per Amazon Serverless OpenSearch

Revisori di terze parti valutano la sicurezza e la conformità di Amazon OpenSearch Serverless nell'ambito di diversi programmi di AWS conformità. Questi programmi includono SOC, PCI e HIPAA.

Per sapere se un Servizio AWS programma rientra nell'ambito di specifici programmi di conformità, consulta Servizi AWS la sezione Scope by Compliance Program Servizi AWS e scegli il programma di conformità che ti interessa. Per informazioni generali, consulta Programmi di AWS conformità Programmi di .

È possibile scaricare report di audit di terze parti utilizzando AWS Artifact. Per ulteriori informazioni, consulta Scaricamento dei report in AWS Artifact .

La vostra responsabilità di conformità durante l'utilizzo Servizi AWS è determinata dalla sensibilità dei dati, dagli obiettivi di conformità dell'azienda e dalle leggi e dai regolamenti applicabili. Per ulteriori informazioni sulla responsabilità di conformità durante l'utilizzo Servizi AWS, consulta la Documentazione AWS sulla sicurezza.

Etichettatura delle raccolte Amazon OpenSearch Serverless

I tag consentono di assegnare informazioni arbitrarie a una raccolta Amazon OpenSearch Serverless in modo da poterle classificare e filtrare in base a tali informazioni. Un tag è un'etichetta di metadati che assegni o che assegni a una risorsa. AWS AWS

Ciascun tag è formato da una chiave e da un valore, Per i tag assegnati da te, puoi definire la chiave e il valore. Ad esempio, potresti definire la chiave come `stagee` il valore di una risorsa come `test`.

Con i tag, puoi identificare e organizzare le tue risorse. AWS Molti AWS servizi supportano l'etichettatura, quindi è possibile assegnare lo stesso tag a risorse di servizi diversi per indicare che le risorse sono correlate. Ad esempio, puoi assegnare lo stesso tag a una raccolta OpenSearch Serverless che assegni a un dominio Amazon OpenSearch Service.

In OpenSearch Serverless, la risorsa principale è una raccolta. È possibile utilizzare la console di OpenSearch servizio AWS CLI, le operazioni dell'API OpenSearch Serverless o AWS SDKs aggiungere, gestire e rimuovere tag da una raccolta.

Autorizzazioni richieste

OpenSearch Serverless utilizza le seguenti autorizzazioni AWS Identity and Access Management Access Analyzer (IAM) per etichettare le raccolte:

- `aoss:TagResource`
- `aoss:ListTagsForResource`
- `aoss:UntagResource`

Etichettatura delle raccolte (console)

La console è il modo più semplice per assegnare un tag a una raccolta.

Per creare un tag (console)

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Espandi Serverless nel pannello di navigazione a sinistra e scegli Collections (Raccolte).
3. Seleziona la raccolta a cui aggiungere i tag e passa alla scheda Tag.
4. Scegli Gestisci, quindi seleziona Aggiungi nuovo tag.
5. Inserire una chiave di tag e un valore facoltativo.
6. Scegli Save (Salva).

Per eliminare un tag, esegui la stessa procedura e scegli Rimuovi nella pagina Gestisci tag.

Per ulteriori informazioni sull'utilizzo della console per il funzionamento con i tag, consulta Editor di tag nella Guida alle operazioni di base della Console di gestione AWS .

Etichettatura delle raccolte ()AWS CLI

Per etichettare una raccolta utilizzando il AWS CLI, invia una TagResourcerichiesta:

```
aws opensearchserverless tag-resource
  --resource-arn arn:aws:aoss:us-east-1:123456789012:collection/my-collection
  --tags Key=service,Value=aoss Key=source,Value=logs
```

Visualizza i tag esistenti per una raccolta con il ListTagsForResourcecomando:

```
aws opensearchserverless list-tags-for-resource
  --resource-arn arn:aws:aoss:us-east-1:123456789012:collection/my-collection
```

Rimuovi i tag da una raccolta usando il UntagResourcecomando:

```
aws opensearchserverless untag-resource
  --resource-arn arn:aws:aoss:us-east-1:123456789012:collection/my-collection
  --tag-keys service
```

Operazioni e plugin supportati in Amazon Serverless OpenSearch

Amazon OpenSearch Serverless supporta una varietà di OpenSearch plug-in, nonché un sottoinsieme delle operazioni API di indicizzazione, ricerca e metadati disponibili in. OpenSearch Per limitare l'accesso a determinate operazioni è possibile includere le autorizzazioni nella colonna sinistra della tabella all'interno delle policy di accesso ai dati.

Argomenti

- Operazioni e autorizzazioni API supportate OpenSearch
- OpenSearch Plugin supportati

Operazioni e autorizzazioni API supportate OpenSearch

La tabella seguente elenca le operazioni API supportate da OpenSearch Serverless, insieme alle corrispondenti autorizzazioni relative alle politiche di accesso ai dati:

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
`aoss:CreateIndex`	PUT <index>	Creazione di indici. Per ulteriori informazioni, consulta la sezione Creazione di indici. **Note** Questa autorizzazione si applica

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
		anche alla creazione di indici con i dati di esempio nei dashboard. OpenSearch
`aoss:DescribeIndex`	• GET <index> • GET <index>/_mapping • GET <index>/_mappings • GET <index>/_setting • GET <index>/_setting/<setting> • GET <index>/_settings • GET <index>/_settings/<setting> • GET _cat/indices • GET _mapping • GET _mappings • GET _resolve/index/<index> • TESTA <index>	Descrizione di indici. Per maggiori informazioni, consulta le seguenti risorse: • Ottenimento di un indice • Ottenimento di una mappatura • Ottenimento di impostazioni • L'indice esiste • Indici CAT (la risposta non include `status` i `health` nostri campi).

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
`aoss:WriteDocument`	• <index>ELIMINA /_doc/ <id> • POST <index>/_bulk • POST <index>/_create/<id> (for search collection types only) • POST <index>/_doc • POST <index>/_update/<id> (for search collection types only) • POST _bulk • PUT <index>/_create/<id> (for search collection types only) • PUT <index>/_doc/<id> (for search collection types only)	Scrittura e aggiornamento di documenti. Per maggiori informazioni, consulta le seguenti risorse: • Bulk • Indicizzazione di dati **Note** Alcune operazioni sono consentite solo per raccolte di tipo SEARCH. Per ulteriori informazioni, consulta the section called “Scelta di un tipo di raccolta”.

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
`aoss:ReadDocument`	• <index>OTTIENI /_analyz • GET <index>/_doc/<id> • GET <index>/_explain/<id> • GET <index>/_mget • GET <index>/_source/<id> • GET <index>/_count • GET <index>/_field_caps • GET <index>/_msearch • GET <index>/_rank_eval • GET <index>/_search • GET <index>/_validate/<query> • GET _analyze • GET _field_caps • GET _mget • GET _search • GET /_search/point_in_time/_all • HEAD <index>/_doc/<id> • HEAD <index>/_source/<id> • POST /_plugins/_sql • POST /_plugins/_ppl • POST /_plugins/_sql/_explain • POST /_plugins/_ppl/_explain • POST /_plugins/_ppl/_close • POST <index>/_analyze • POST /_search/punto_temporale • POST <index>/_explain/<id> • POST <index>/_count • POST <index>/_field_caps	Lettura di documenti. Per maggiori informazioni, consulta le seguenti risorse: • Esecuzione di analisi del testo • Ottenimento di un documento • Conteggio • Query DSL • Valutazione della classifica • Analisi di API • Spiegazione • Punto nel tempo • SQL e PPL

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
	• POST <index>/_rank_eval • POST <index>/_search • POST _analyze • POST _field_caps • POST _search • ELIMINA /_search/point_in_time/_all • ELIMINA /_search/point_in_time	
`aoss:DeleteIndex`	DELETE <target>	Eliminazione degli indici. Per ulteriori informazioni, consulta la sezione Eliminazione di indici.
`aoss:UpdateIndex`	• POST _mapping • POST <index>/_mapping/ • POST <index>/_mappings/ • POST <index>/_setting • POST <index>/_settings • POST _setting • POST _settings • PUT _mapping • PUT <index>/_mapping • PUT <index>/_mappings/ • PUT <index>/_setting • PUT <index>/_settings • PUT _setting • PUT _settings	Aggiornamento delle impostazioni dell'indice. Per maggiori informazioni, consulta le seguenti risorse: • Mappatura • Aggiornamento delle impostazioni

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
`aoss:CreateCollectionItems`	• POST _aliases • POST /_plugins/_flow_framework/workflow • <workflow_id>* POST /_plugins/_flow_framework/workflow/ /_provisioning • PUT _ingest/pipeline/ <pipeline-id> • PUT _search/pipeline/ <pipeline-id>	• Crea alias di indice, pipeline e modelli. Per ulteriori informazioni, consulta la sezione Creazione di alias. • * Fornitura o riassegnazione di modelli. I servizi ML Commons Client e OpenSearch Serverless gestiscono le politiche dipendenti.

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
`aoss:DescribeCollectionItems`	• GET <index>/_alias/<alias> • GET _alias • GET _alias/<alias> • GET _cat/aliases • GET _cat/templates • GET _cat/templates/<template_name> • GET _component_template • GET _component_template/ <component-template> • GET _index_template • GET _index_template/<index-template> • GET _ingest/pipeline/ <pipeline-id> • GET _ingest/pipeline/_simulate • OTTIENI /_plugins/_flow_framework/workflow/ <workflow-id> • OTTIENI /_plugins/_flow_framework/workflow/_search • <workflow-id>OTTIENI /_plugins/_flow_framework/workflow/ /_status • GET /_plugins/_flow_ /_search framework/workflow/state • OTTIENI /_plugins/_flow_framework/workflow/_steps • OTTIENI /_plugins/_flow_framework/workflow/_step? fase del flusso di lavoro= <step_name>	Descrive come lavorare con alias, modelli di indici e framework e pipeline. Per maggiori informazioni, consulta le seguenti risorse: • Gestione di alias • Modelli di indice

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
	• GET _search/pipeline/ <pipeline-id> • HEAD _alias/<alias> • HEAD _component_template/ <component-template> • HEAD _index_template/<name> • HEAD <index>/_alias/<alias> • POST _import/pipeline/_simulate • POST /_plugins/_flow_framework/ workflow/_ricerca • POST /_plugins/_flow_ framework/ workflow/state /_ricerca	
`aoss:UpdateCollectionItems`	• POST <index>/_alias/<alias> • POST <index>/_aliases/<alias> • POST _component_template/ <component-template> • POST _index_template/<index-template> • <workflow_id>* POST /_plugins /_flow_framework/workflow/ / _deprovision • PUT <index>/_alias/<alias> • PUT <index>/_aliases/<alias> • PUT _component_template/ <component-template> • PUT _index_template/<index-template> • INSERISCI /_plugins/_flow_framework/workflow/ <workflow_id>	Aggiorna alias, modelli di indice e modelli di framework. Per maggiori informazioni, consulta le seguenti risorse: • Alias di indice • Modelli di indice * L'API per annullare il provisioning dei modelli. I servizi ML Commons Client e OpenSearch Serverless gestiscono le politiche dipendenti.

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
`aoss:DeleteCollectionItems`	• DELETE <index>/_alias/<alias> • DELETE _component_template/<component-template> • DELETE _index_template/<index-template> • DELETE <index>/_aliases/<alias> • ELIMINA _search/pipeline/<pipeline-id> • ELIMINA _ingest/pipeline/<pipeline-id> • ELIMINA /_plugins/_flow_framework/workflow/ <workflow_id>	Eliminare alias, modelli di indice e framework e pipeline. Per maggiori informazioni, consulta le seguenti risorse: • Eliminazione di alias • Eliminazione di un modello

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
`aoss:DescribeMLResource`	• OTTIENI /_plugins/_ml/models/<model_id> • GET /_plugins_ml/models/_search • GET /_plugins/_ml/model_groups/<model_group_id> • GET /_plugins/_ml/model_groups/_search • GET /_plugins/_ml/connectors/<connector_id> • GET /_plugins/_ml/connectors/_search • OTTIENI ml/profile/tasks /_plugins/_/<task_id> • POST /_plugins/_ml/models/_search • POST /_plugins/_ml/model_groups/_search • POST /_plugins/_ml/connettori/_ricerca	Descrive GET e search per recuperare informazioni su modelli e connettori. APIs
`aoss:CreateMLResource`	• POST /_plugins/_ml/models/_register • POST /_plugins/_ml/model_groups/_register • POST /_plugins/_ml/connectors/_create	Fornisce il permesso di creare risorse ML.

Autorizzazione della policy di accesso ai dati	OpenSearch operazioni API	Descrizione e avvertenze
`aoss:UpdateMLResource`	• INSERISCI /_plugins/_ml/models/ <model_id> • <model_id>INSERISCI /_plugins/_ml/models/ /_deploy • <model_id>POST /_plugins/_ml/modelli/ /_undeploy • INSERISCI /_plugins/_ml/model_groups/ <model_group_id> • INSERISCI /_plugins/_ml/connectors/ <connector_id>	Fornisce l'autorizzazione per aggiornare le risorse ML esistenti.
`aoss:DeleteMLResource`	• ELIMINA /_plugins/_ml/models/ <model_id> • ELIMINA _plugins/_ml/model_groups/ <model_group_id> • ELIMINA /_plugins/_ml/connectors/ <connector_id> • ELIMINA /_plugins/_ml/tasks/ <task_id>	Fornisce l'autorizzazione per eliminare le risorse ML.
`aoss:ExecuteMLResource`	• <model_id>POST /_plugins/_ml/models/ /_predict	Fornisce il permesso di eseguire modelli.

OpenSearch Plugin supportati

OpenSearch Le raccolte serverless sono preconfezionate con i seguenti plugin della community. OpenSearch Serverless implementa e gestisce automaticamente i plug-in per tuo conto.

Plug-in di analisi

- ICU Analysis
- Japanese (kuromoji) Analysis

- Korean (Nori) Analysis
- Phonetic Analysis
- Smart Chinese Analysis
- Stempel Polish Analysis
- Ukrainian Analysis

Plug-in mapper

- Mapper Size
- Mapper Murmur3
- Mapper Annotated Text

Plug-in di scripting

- Painless
- Expression
- Mustache

Inoltre, OpenSearch Serverless include tutti i plugin forniti come moduli.

Monitoraggio di Amazon OpenSearch Serverless

Il monitoraggio è una parte importante per mantenere l'affidabilità, la disponibilità e le prestazioni di Amazon OpenSearch Serverless e delle altre AWS soluzioni. AWS fornisce i seguenti strumenti di monitoraggio per monitorare OpenSearch Serverless, segnalare quando qualcosa non va e intraprendere azioni automatiche se necessario:

- Amazon CloudWatch monitora AWS le tue risorse e le applicazioni su cui esegui AWS in tempo reale. È possibile raccogliere e tenere traccia dei parametri, creare pannelli di controllo personalizzati e impostare allarmi per inviare una notifica o intraprendere azioni quando un parametro specificato raggiunge una determinata soglia.

  Ad esempio, puoi tenere CloudWatch traccia dell'utilizzo della CPU o di altri parametri delle tue istanze Amazon EC2 e avviare automaticamente nuove istanze quando necessario. Per ulteriori informazioni, consulta la Amazon CloudWatch User Guide.

- AWS CloudTrail acquisisce chiamate API ed eventi correlati da parte di o per conto del tuo Account AWS. Distribuisce i file di log a un bucket Amazon S3 specificato. Puoi identificare quali utenti e account hanno effettuato la chiamata AWS, l'indirizzo IP di origine da cui sono state effettuate le chiamate e quando sono avvenute le chiamate. Per ulteriori informazioni, consulta la Guida per l'utente AWS CloudTrail.
- Amazon EventBridge offre un flusso quasi in tempo reale di eventi di sistema che descrivono le modifiche nei tuoi domini OpenSearch di servizio. Puoi creare regole che controllano determinati eventi e attivano azioni automatiche in altri Servizi AWS quando si verificano tali eventi. Per ulteriori informazioni, consulta la Amazon EventBridge User Guide.

Monitoraggio OpenSearch serverless con Amazon CloudWatch

Puoi monitorare l'utilizzo di Amazon OpenSearch Serverless CloudWatch, che raccoglie dati grezzi e li elabora in parametri leggibili quasi in tempo reale. Queste statistiche vengono conservate per un periodo di 15 mesi, per permettere l'accesso alle informazioni storiche e offrire una prospettiva migliore sulle prestazioni del servizio o dell'applicazione web.

È anche possibile impostare allarmi che controllano determinate soglie e inviare notifiche o intraprendere azioni quando queste soglie vengono raggiunte. Per ulteriori informazioni, consulta la Amazon CloudWatch User Guide.

OpenSearch Serverless riporta le seguenti metriche nel namespace. `AWS/AOSS`

Metrica	Description
`ActiveCollection`	Indica se una raccolta è attiva. Un valore pari a 1 significa che la raccolta è in uno stato `ACTIVE`. Questo valore viene emesso dopo la corretta creazione di una raccolta e rimane 1 finché non si elimina la raccolta. Il parametro non può avere un valore pari a 0. Statistiche rilevanti: Max (Massimo) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` Frequenza: 60 secondi

Metrica	Description
`DeletedDocuments`	Il numero totale di documenti eliminati. Statistiche rilevanti: Average (Media), Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frequenza: 60 secondi
`IndexingOCU`	Il numero di unità di OpenSearch calcolo (OCUs) utilizzate per importare i dati della raccolta. Questo parametro si applica a livello di account. Rappresenta l'utilizzo solo per le raccolte che non fanno parte di alcun gruppo di raccolte. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId` Frequenza: 60 secondi
`IndexingOCU`	Il numero di unità di OpenSearch calcolo (OCUs) utilizzate per importare i dati della raccolta. Questa metrica si applica a livello di gruppo di raccolta. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionGroupId` , `CollectionGroupName` Frequenza: 60 secondi

Metrica	Description
`IngestionDataRate`	La velocità di indicizzazione in GiB al secondo di una raccolta o di un indice. Questo parametro si applica solo alle richieste di indicizzazione in blocco. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frequenza: 60 secondi
`IngestionDocumentErrors`	Il numero totale di errori del documento durante l'importazione di una raccolta o di un indice. Dopo che una richiesta di indicizzazione in blocco è riuscita, le istanze di scrittura elaborano la richiesta e generano errori per tutti i documenti non riusciti all'interno della richiesta. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frequenza: 60 secondi
`IngestionDocumentRate`	La frequenza al secondo con cui i documenti vengono inseriti in una raccolta o in un indice. Questo parametro si applica solo alle richieste di indicizzazione in blocco. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frequenza: 60 secondi

Metrica	Description
`IngestionRequestErrors`	Il numero totale di errori di richiesta di indicizzazione in blocco relativi a una raccolta. OpenSearch Serverless emette questa metrica quando una richiesta di indicizza zione di massa fallisce per qualsiasi motivo, ad esempio un problema di autenticazione o disponibilità. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `Collectio nName` Frequenza: 60 secondi
`IngestionRequestLatency`	La latenza, in secondi, per le operazioni di scrittura in blocco su una raccolta. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio) Dimensioni: `ClientId`, `CollectionId` , `Collectio nName` Frequenza: 60 secondi
`IngestionRequestRate`	Il numero totale di operazioni di scrittura in blocco ricevute da una raccolta. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio) Dimensioni: `ClientId`, `CollectionId` , `Collectio nName` Frequenza: 60 secondi

Metrica	Description
`IngestionRequestSuccess`	Il numero totale di operazioni di indicizzazione riuscite su una raccolta. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` Frequenza: 60 secondi
`SearchableDocuments`	Il numero totale di documenti ricercabili in una raccolta o in un indice. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frequenza: 60 secondi
`SearchRequestErrors`	Il numero totale di errori di query al minuto per una raccolta. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` Frequenza: 60 secondi

Metrica	Description
`SearchRequestLatency`	Il tempo medio, in millisecondi, impiegato per completare un'operazione di ricerca su una raccolta. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` Frequenza: 60 secondi
`Search0CU`	Il numero di unità di OpenSearch calcolo () OCUs utilizzate per cercare i dati della raccolta. Questo parametro si applica a livello di account. Rappresenta l'utilizzo solo per le raccolte che non fanno parte di alcun gruppo di raccolte. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId` Frequenza: 60 secondi
`Search0CU`	Il numero di unità di OpenSearch calcolo (OCUs) utilizzate per cercare i dati della raccolta. Questa metrica si applica a livello di gruppo di raccolta. Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionGroupId` , `CollectionGroupName` Frequenza: 60 secondi

Metrica	Description
SearchRequestRate	Il numero totale di richieste di ricerca al minuto a una raccolta. Statistiche rilevanti: Average (Media), Maximum (Massimo) , Sum (Somma) Dimensioni: ClientId, CollectionId , CollectionName Frequenza: 60 secondi
StorageUsedInS3	La quantità, in byte, di storage Amazon S3 utilizzata. OpenSearch Serverless archivia i dati indicizzati in Amazon S3. È necessario selezionare il periodo in un minuto per ottenere un valore preciso. Statistiche rilevanti: Sum (Somma) Dimensioni: ClientId, CollectionId , CollectionName , IndexId, IndexName Frequenza: 60 secondi
VectorIndexBuildAccelerationOCU	Il numero di unità di OpenSearch calcolo (OCUs) utilizzate per accelerare l'indicizzazione vettoriale. Questa metrica si applica a livello di raccolta. Statistiche rilevanti: Sum (Somma) Dimensioni:ClientId, CollectionId Frequenza: 60 secondi

Metrica	Description
`2xx, 3xx, 4xx, 5xx`	Il numero di richieste alla raccolta che hanno prodotto il codice di risposta HTTP specificato (2xx, 3xx, 4xx, 5xx). Statistiche rilevanti: Sum (Somma) Dimensioni: `ClientId`, `CollectionId` , `CollectionName` Frequenza: 60 secondi

Registrazione delle chiamate API OpenSearch Serverless utilizzando AWS CloudTrail

Amazon OpenSearch Serverless è integrato con AWS CloudTrail, un servizio che fornisce un registro delle azioni intraprese da un utente, un ruolo o un AWS servizio in Serverless.

CloudTrail acquisisce tutte le chiamate API per OpenSearch Serverless come eventi. Le chiamate acquisite includono chiamate dalla sezione Serverless della console di OpenSearch servizio e chiamate in codice alle operazioni dell'API OpenSearch Serverless.

Se crei un trail, puoi abilitare la distribuzione continua di CloudTrail eventi a un bucket Amazon S3, inclusi gli eventi per Serverless. OpenSearch Se non configuri un percorso, puoi comunque visualizzare gli eventi più recenti nella CloudTrail console nella cronologia degli eventi.

Utilizzando le informazioni raccolte da CloudTrail, è possibile determinare la richiesta che è stata effettuata a OpenSearch Serverless, l'indirizzo IP da cui è stata effettuata la richiesta, chi ha effettuato la richiesta, quando è stata effettuata e ulteriori dettagli.

Per ulteriori informazioni CloudTrail, consulta la Guida per l'AWS CloudTrail utente.

OpenSearch Informazioni serverless in CloudTrail

CloudTrail è abilitato sul tuo Account AWS quando crei l'account. Quando l'attività si verifica in OpenSearch Serverless, tale attività viene registrata in un CloudTrail evento insieme ad altri eventi di AWS servizio nella cronologia degli eventi. Puoi visualizzare, cercare e scaricare eventi recenti in. Account AWS Per ulteriori informazioni, consulta Visualizzazione degli eventi con la cronologia degli CloudTrail eventi.

Per una registrazione continua degli eventi in tuo Account AWS, inclusi gli eventi per OpenSearch Serverless, crea un percorso. Un trail consente di CloudTrail inviare file di log a un bucket Amazon S3. Per impostazione predefinita, quando si crea un percorso nella console, questo sarà valido in tutte le Regioni AWS.

Il trail registra gli eventi di tutte le regioni della AWS partizione e consegna i file di log al bucket Amazon S3 specificato. Inoltre, puoi configurare altri AWS servizi per analizzare ulteriormente e agire in base ai dati sugli eventi raccolti nei log. CloudTrail Per ulteriori informazioni, consulta gli argomenti seguenti:

- Panoramica della creazione di un percorso
- CloudTrail servizi e integrazioni supportati
- Configurazione delle notifiche Amazon SNS per CloudTrail
- Ricezione di file di CloudTrail registro da più regioni e ricezione di file di CloudTrail registro da più account

Tutte le azioni OpenSearch Serverless vengono registrate CloudTrail e documentate nel riferimento all'API OpenSearch Serverless. Ad esempio, le chiamate alle `CreateCollection` `DeleteCollection` azioni e generano voci nei file di registro. `ListCollections` CloudTrail

Ogni evento o voce di log contiene informazioni sull'utente che ha generato la richiesta. Le informazioni di identità consentono di stabilire:

- Se la richiesta è stata effettuata con credenziali utente root o AWS Identity and Access Management (IAM).
- Se la richiesta è stata effettuata con le credenziali di sicurezza temporanee per un ruolo o un utente federato.
- Se la richiesta è stata effettuata da un altro AWS servizio.

Per ulteriori informazioni, consulta Elemento CloudTrail userIdentity.

OpenSearch Eventi di dati senza server in CloudTrail

Gli eventi relativi ai dati forniscono informazioni sulle operazioni eseguite sulle risorse su o all'interno di una risorsa (ad esempio, la ricerca o l'indicizzazione in una raccolta OpenSearch Serverless). Queste operazioni sono definite anche operazioni del piano dei dati. Gli eventi di dati sono spesso

attività che interessano volumi elevati di dati. Per impostazione predefinita, CloudTrail non registra gli eventi relativi ai dati. La cronologia CloudTrail degli eventi non registra gli eventi relativi ai dati.

Per gli eventi di dati sono previsti costi aggiuntivi. Per ulteriori informazioni sui CloudTrail prezzi, consulta la sezione AWS CloudTrail Prezzi.

Puoi registrare gli eventi relativi ai dati per i tipi di `AWS::AOSS::Collection` risorse utilizzando la CloudTrail console o AWS CLI le operazioni CloudTrail dell'API. Per ulteriori informazioni su come registrare gli eventi relativi ai dati, consulta Registrazione di eventi relativi ai dati con Console di gestione AWS e Registrazione di eventi di dati con AWS Command Line Interface nella Guida per l'utente AWS CloudTrail .

È possibile configurare selettori di eventi avanzati per filtrare i campi `eventName`, `readOnly` e `resources.`ARN per registrare solo gli eventi importanti per l'utente. Per ulteriori informazioni su questi campi, consulta AdvancedFieldSelector in Riferimento API AWS CloudTrail .

Comprendere le OpenSearch voci relative ai Serverless Data Event

Nel seguente esempio:

- Il `requestParameters` campo contiene dettagli sulla chiamata API effettuata alla raccolta. Include il percorso della richiesta di base (senza parametri di query).
- Il `responseElements` campo include un codice di stato che indica l'esito della richiesta durante la modifica delle risorse. Questo codice di stato consente di verificare se le modifiche sono state elaborate correttamente o richiedono attenzione.
- OpenSearch Serverless registra gli eventi CloudTrail relativi ai dati solo per le richieste che hanno completato con successo l'autenticazione IAM.

```
        "accessKeyId": "access-key",
        "userName": "",
        "sessionContext": {
          "sessionIssuer": {
            "type": "Role",
            "principalId": "AROA123456789EXAMPLE",
            "arn": "arn:aws:iam::111122223333:role/Admin",
            "accountId": "111122223333",
            "userName": "Admin"
          },
          "attributes": {
            "creationDate": "2025-08-15T22:57:38Z",
            "mfaAuthenticated": "false"
          },
          "sourceIdentity": "",
          "ec2RoleDelivery": "",
          "assumedRoot": ""
        },
        "identityProvider": "",
        "credentialId": ""
      },
      "eventTime": "2025-08-15T22:58:00Z",
      "eventSource": "aoss.amazonaws.com",
      "eventName": "Search",
      "awsRegion": "us-east-1",
      "sourceIPAddress": "AWS Internal",
      "userAgent": "python-requests/2.32.3",
      "requestParameters": {
        "pathPrefix": "/_search"
      },
      "responseElements": null,
      "requestID": "2cfee788-EXAM-PLE1-8617-4018cEXAMPLE",
      "eventID": "48d43617-EXAM-PLE1-9d9c-f7EXAMPLE",
      "readOnly": true,
      "resources": [
        {
          "type": "AWS::AOSS::Collection",
          "ARN": "arn:aws:aoss:us-east-1:111122223333:collection/aab9texampletu45xh77"
        }
      ],
      "eventType": "AwsApiCall",
      "managementEvent": false,
      "recipientAccountId": "111122223333",
      "eventCategory": "Data"
    }
  ]
}
```

Comprensione delle voci OpenSearch relative agli eventi di gestione serverless

Un trail è una configurazione che consente la distribuzione di eventi come file di log in un bucket Amazon S3 specificato dall'utente. CloudTrail i file di registro contengono una o più voci di registro.

Un evento rappresenta una singola richiesta da un'origine. Include informazioni sull'azione richiesta, la data e l'ora dell'azione, i parametri della richiesta e così via. CloudTrail i file di registro non sono una traccia ordinata dello stack delle chiamate API pubbliche, quindi non vengono visualizzati in un ordine specifico.

L'esempio seguente visualizza una voce di CloudTrail registro che illustra l'`CreateCollection`azione.


```
{
   "eventVersion":"1.08",
   "userIdentity":{
      "type":"AssumedRole",
      "principalId":"AIDACKCEVSQ6C2EXAMPLE",
      "arn":"arn:aws:iam::123456789012:user/test-user",
      "accountId":"123456789012",
      "accessKeyId":"access-key",
      "sessionContext":{
         "sessionIssuer":{
            "type":"Role",
            "principalId":"AIDACKCEVSQ6C2EXAMPLE",
            "arn":"arn:aws:iam::123456789012:role/Admin",
            "accountId":"123456789012",
            "userName":"Admin"
         },
         "webIdFederationData":{

         },
         "attributes":{
            "creationDate":"2022-04-08T14:11:34Z",
            "mfaAuthenticated":"false"
         }
      }
   },
   "eventTime":"2022-04-08T14:11:49Z",
   "eventSource":"aoss.amazonaws.com",
   "eventName":"CreateCollection",
   "awsRegion":"us-east-1",
   "sourceIPAddress":"AWS Internal",
```

```
   "userAgent":"aws-cli/2.1.30 Python/3.8.8 Linux/5.4.176-103.347.amzn2int.x86_64 exe/
x86_64.amzn.2 prompt/off command/aoss.create-collection",
   "errorCode":"HttpFailureException",
   "errorMessage":"An unknown error occurred",
   "requestParameters":{
      "accountId":"123456789012",
      "name":"test-collection",
      "description":"A sample collection",
      "clientToken":"d3a227d2-a2a7-49a6-8fb2-e5c8303c0718"
   },
   "responseElements": null,
   "requestID":"12345678-1234-1234-1234-987654321098",
   "eventID":"12345678-1234-1234-1234-987654321098",
   "readOnly":false,
   "eventType":"AwsApiCall",
   "managementEvent":true,
   "recipientAccountId":"123456789012",
   "eventCategory":"Management",
   "tlsDetails":{
      "clientProvidedHostHeader":"user.aoss-sample.us-east-1.amazonaws.com"
   }
}
```


Monitoraggio degli eventi OpenSearch serverless tramite Amazon EventBridge

Amazon OpenSearch Service si integra con Amazon EventBridge per notificarti determinati eventi che influiscono sui tuoi domini. Gli eventi AWS relativi ai servizi vengono forniti quasi EventBridge in tempo reale. Gli stessi eventi vengono inviati anche ad Amazon CloudWatch Events, il predecessore di Amazon EventBridge. Puoi scrivere regole per indicare quali eventi ti interessano e quali azioni automatiche intraprendere quando un evento corrisponde a una regola. Di seguito sono riportati alcuni esempi di azioni che è possibile attivare automaticamente:

- Invocare una funzione AWS Lambda
- Richiamo di un Run Command di Amazon EC2
- Inoltro dell'evento a Amazon Kinesis Data Streams
- Attivazione di una macchina a stati AWS Step Functions
- Notifica di un argomento Amazon SNS o di una coda Amazon SQS

Per ulteriori informazioni, consulta la sezione Guida introduttiva ad Amazon EventBridge nella Amazon EventBridge User Guide.

Configurazione delle notifiche

Puoi utilizzare le notifiche AWS utente per ricevere notifiche quando si verifica un evento OpenSearch Serverless. Un evento è un indicatore di un cambiamento nell'ambiente OpenSearch Serverless, ad esempio quando si raggiunge il limite massimo di utilizzo dell'OCU. Amazon EventBridge riceve l'evento e invia una notifica al Centro Console di gestione AWS notifiche e ai canali di distribuzione scelti. L'utente riceverà una notifica quando un evento corrisponde a una regola specificata.

OpenSearch eventi Compute Units (OCU)

OpenSearch Serverless invia eventi a EventBridge quando si verifica uno dei seguenti eventi relativi all'OCU.

L'utilizzo dell'OCU si avvicina al limite massimo

OpenSearch Serverless invia questo evento quando l'utilizzo dell'OCU per la ricerca o l'indicizzazione raggiunge il 75% del limite di capacità. L'utilizzo dell'OCU viene calcolato in base al limite di capacità configurato e al consumo attuale dell'OCU.

Esempio

Di seguito è riportato un esempio di evento di questo tipo (cerca OCU):

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "OCU Utilization Approaching Max Limit",
  "source": "aws.aoss",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "eventTime" : 1678943345789,
    "description": "Your search OCU usage is at 75% and is approaching the configured
 maximum limit."
  }
```

```
}
```

Di seguito è riportato un esempio di evento di questo tipo (indice OCU):

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "OCU Utilization Approaching Max Limit",
  "source": "aws.aoss",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "eventTime" : 1678943345789,
    "description": "Your indexing OCU usage is at 75% and is approaching the configured
 maximum limit."
  }
```

L'utilizzo dell'OCU ha raggiunto il limite massimo

OpenSearch Serverless invia questo evento quando l'utilizzo dell'OCU per la ricerca o l'indice raggiunge il 100% del limite di capacità. L'utilizzo dell'OCU viene calcolato in base al limite di capacità configurato e al consumo attuale dell'OCU.

Esempio

Di seguito è riportato un esempio di evento di questo tipo (cerca OCU):

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "OCU Utilization Reached Max Limit",
  "source": "aws.aoss",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "eventTime" : 1678943345789,
    "description": "Your search OCU usage has reached the configured maximum limit."
```

```
  }
}
```

Di seguito è riportato un esempio di evento di questo tipo (indice OCU):

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "OCU Utilization Reached Max Limit",
  "source": "aws.aoss",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "eventTime" : 1678943345789,
    "description": "Your indexing OCU usage has reached the configured maximum limit."
  }
}
```

Creazione e gestione di domini Amazon OpenSearch Service

Questo capitolo descrive come creare e gestire i domini Amazon OpenSearch Service. Un dominio è l'equivalente AWS fornito di un cluster open source. OpenSearch Quando si crea un dominio, si specificano le impostazioni, i tipi di istanza, il numero di istanze e l'allocazione dello storage. Per ulteriori informazioni sui cluster open source, consulta Creazione di un cluster nella OpenSearch documentazione.

A differenza delle istruzioni nel tutorial relativo alle nozioni di base, in questo capitolo sono descritte tutte le opzioni e sono fornite le informazioni di riferimento rilevanti. È possibile completare ogni procedura utilizzando le istruzioni per la console OpenSearch di servizio, il AWS Command Line Interface (AWS CLI) o il AWS SDKs.

Creazione OpenSearch di domini di servizio

Questa sezione descrive come creare domini OpenSearch di servizio utilizzando la console di OpenSearch servizio o utilizzando il comando AWS CLI with the`create-domain`.

Creazione OpenSearch di domini di servizio (console)

Utilizzare la procedura seguente per creare un dominio OpenSearch di servizio utilizzando la console.

Per creare un dominio OpenSearch di servizio (console)

1. Vai a https://aws.amazon.come scegli Accedi alla console.
2. In Analytics, scegli Amazon OpenSearch Service.
3. Scegli Crea dominio.
4. In Domain name (Nome di dominio), immettere un nome di dominio. Il nome deve soddisfare i seguenti criteri:
   - Unico per il tuo account e Regione AWS
   - Inizia con una lettera minuscola
   - Contiene da 3 a 28 caratteri
   - Contiene solo lettere minuscole a - z, i numeri da 0 a 9 e i trattini (-)
5. Per il metodo di creazione del dominio, scegli Creazione standard.

6. Per i modelli, scegli l'opzione più adatta allo scopo del tuo dominio:

   - Domini di produzione per carichi di lavoro che richiedono disponibilità e prestazioni elevate. Questi domini utilizzano Multi-AZ (con o senza standby) e nodi master dedicati per una maggiore disponibilità.
   - Sviluppo/test per lo sviluppo o il test. Questi domini possono utilizzare Multi-AZ (con o senza standby) o una singola zona di disponibilità.

   > **Important**
   >
   > I diversi tipi di distribuzione presentano diverse opzioni su pagine successive. Questi passaggi includono tutte le opzioni.

7. Per le opzioni di distribuzione, scegli Dominio con standby per configurare un dominio 3-AZ, con i nodi in una delle zone riservati come standby. Questa opzione applica una serie di best practice, come il numero di nodi di dati specificato, il conteggio dei nodi principali, il tipo di istanza, il numero di repliche e le impostazioni di aggiornamento software.
8. Per Versione, scegli la versione del OpenSearch sistema operativo Elasticsearch precedente da utilizzare. Ti consigliamo di scegliere la versione più recente di. OpenSearch Per ulteriori informazioni, consulta the section called “Versioni supportate di Elasticsearch e OpenSearch”.

   (Facoltativo) Se hai scelto una OpenSearch versione per il tuo dominio, seleziona Abilita la modalità di compatibilità per OpenSearch segnalare la versione 7.10, che consente a determinati client e plugin OSS di Elasticsearch che controllano la versione prima di connettersi di continuare a utilizzare il servizio.
9. Per Tipo di istanza, scegliere un tipo di istanza per i nodi di dati. Per ulteriori informazioni, consultare the section called “Tipi di istanze supportati”.

   > **Note**
   >
   > Non tutte le zone di disponibilità supportano tutti i tipi di istanze. Se scegli Multi-AZ con o senza Standby, ti consigliamo di scegliere i tipi di istanza della generazione attuale, come R5 o I3.

10. In Number of nodes (Numero di nodi), scegliere il numero di nodi di dati.

    Per i valori massimi, consulta Quote di dominio e istanza del servizio. OpenSearch I cluster a nodo singolo sono ideali per lo sviluppo e i test, ma non devono essere utilizzati per i carichi di

lavoro di produzione. Per ulteriori linee guida, consulta the section called “Dimensionamento dei domini” e the section called “Configurazione di un dominio Multi-AZ”.

> **Note**
>
> (Facoltativo) I nodi coordinatori dedicati supportano tutte le OpenSearch versioni e ElasticSearch le versioni dalla 6.8 alla 7.10. I nodi coordinatori dedicati sono disponibili per l'uso con domini con un gestore di cluster dedicato abilitato. Per abilitare i nodi coordinatori dedicati, dovrai selezionare il tipo e il numero di istanze. Come procedura ottimale, è consigliabile mantenere la stessa famiglia di istanze per il nodo coordinatore dedicato e i nodi dati (istanze basate su Intel o istanze basate su Graviton).

11. Per Tipo di storage, seleziona Amazon EBS. I tipi di volume disponibili nell'elenco dipendono dal tipo di istanza scelta. Per istruzioni su come creare domini di grandi dimensioni, consultare the section called “Scala in petabyte”.
12. Per lo storage EBS, configura le seguenti impostazioni aggiuntive. Alcune impostazioni potrebbero non essere visualizzate a seconda del tipo di volume scelto.

Impostazione	Description
Tipi di volume EBS	Scegli tra Scopo generico (SSD) - gp3 e Scopo generico (SSD) - gp2 o la generazione precedente Capacità di IOPS allocata (SSD) e Magnetico (standard).
Dimensioni dello storage EBS per ogni nodo	Inserisci le dimensioni del volume EBS da collegare a ogni nodo di dati. EBS volume size (Dimensione del volume EBS) è per nodo. È possibile calcolare la dimensione totale del cluster per il dominio del OpenSearch servizio moltiplicando il numero di nodi di dati per la dimensione del volume EBS. Le dimensioni minime e massime di un volume EBS dipendono sia dal tipo di volume EBS specificato sia dal tipo di istanza collegata. Per ulteriori informazioni, consulta Limiti delle dimensioni dei volumi EBS.

Impostazione	Description
IOPS con provisioning	Se hai selezionato un tipo di volume SSD Provisioned IOPS, inserisci il numero di I/O operazioni al secondo (IOPS) che il volume può supportare.

13. (Facoltativo) Se hai selezionato un tipo di `gp3` volume, espandi le Impostazioni avanzate e specifica IOPS (fino a 16.000 per ogni dimensione di volume da 3 TiB fornita per nodo di dati) e velocità effettiva (fino a 1.000 per MiB/s ogni dimensione di volume da 3 TiB fornita per nodo di dati) oltre a quanto incluso nel prezzo dello storage, a un costo aggiuntivo. Per ulteriori informazioni, consulta i prezzi OpenSearch di Amazon Service.
14. (Facoltativo) Per abilitare UltraWarm lo storage, scegli Abilita nodi di UltraWarm dati. Ogni tipo di istanza ha una quantità massima di spazio di archiviazione che può indirizzare. Moltiplicare tale importo per il numero di nodi di dati a caldo per l'archiviazione a caldo indirizzabile totale.
15. (Facoltativo) Per abilitare l'archiviazione a freddo, scegliere Abilita archiviazione a freddo. È necessario abilitare UltraWarm per abilitare la conservazione a freddo.
16. Se si utilizza Multi-AZ con Standby, tre nodi master dedicati sono già abilitati. Scegli il tipo di nodi master che desideri. Se hai scelto un dominio Multi-AZ senza Standby, seleziona Abilita nodi master dedicati e scegli il tipo e il numero di nodi master che desideri. I nodi master dedicati aumentano la stabilità del cluster e sono richiesti per domini con un conteggio istanze superiore a 10. Per i domini di produzione consigliamo tre nodi master dedicati.

    > **Note**
    >
    > È possibile scegliere tra tipi di istanze differenti per nodi master dedicati e nodi di dati. Ad esempio, è possibile selezionare istanze generiche o ottimizzate per l'archiviazione per i nodi di dati e scegliere istanze ottimizzate per l'elaborazione per i nodi principali dedicati.

17. (Facoltativo) Per i domini che eseguono OpenSearch Elasticsearch 5.3 e versioni successive, la configurazione Snapshot è irrilevante. Per ulteriori informazioni sugli snapshot automatici, consulta the section called “Creazione di snapshot di indici”.
18. Se si desidera utilizzare un endpoint personalizzato anziché quello standard di `https://search-`*`mydomain`*`-`*`1a2a3a4a5a6a7a8a9a0a9a8a7a`*`.`*`us-east-1`*`.es.amazonaws.com`, scegliere Abilita endpoint personalizzato e fornire un nome e un certificato. Per ulteriori informazioni, consulta the section called “Creazione di un endpoint personalizzato”.

19. In Rete, scegliere Accesso VPC o Accesso pubblico. Se si sceglie Public access (Accesso pubblico), andare al passaggio successivo. Se si sceglie Accesso VPC, è necessario accertarsi di soddisfare tutti i prerequisiti e configurare le seguenti impostazioni:

Impostazione	Description
VPC	Scegli l'ID per il cloud privato virtuale (VPC) da utilizzare. Il VPC e il dominio devono essere uguali Regione AWS ed è necessario seleziona re un VPC con tenancy impostata su Default. OpenSearch Il servizio non supporta ancora l'utilizzo della locazione VPCs dedicata.
Sottorete	Scegli una sottorete. Se hai abilitato Multi-AZ, devi scegliere due o tre sottoreti. OpenSearch Il servizio inserirà un endpoint VPC e interfacce di rete elastiche nelle sottoreti. È necessario riservare un numero sufficiente di indirizzi IP per le interfacc e di rete nelle sottoreti. Per ulteriori informazioni, consultare Prenotazione di indirizzi IP in una sottorete VPC.
Gruppi di sicurezza	Scegli uno o più gruppi di sicurezza VPC che consentano all'applicazione richiesta di raggiungere il dominio del OpenSearch servizio sulle porte (80 o 443) e sui protocolli (HTTP o HTTPS) esposti dal dominio. Per ulteriori informazioni, consulta the section called “Supporto per VPC”.
Ruolo IAM	Mantieni il ruolo predefinito. OpenSearch Il servizio utilizza questo ruolo predefinito (noto anche come ruolo collegato al servizio) per accedere al VPC e posizionare un endpoint VPC e interfacce di rete nella sottorete del VPC. Per ulteriori informazioni, consultare Ruolo collegato ai servizi per l'accesso VPC.
Tipo di indirizzo IP	Scegli tra dual stack o IPv4 come tipo di indirizzo IP. Il dual stack consente di condividere le risorse di dominio tra diversi tipi di IPv6 indirizzi IPv4 ed è l'opzione consigliata. Se imposti il tipo di indirizzo IP su dual stack, non puoi modificare il tipo di indirizzo in un secondo momento.

20. Abilitare o disabilitare il controllo granulare degli accessi:

    - Se si desidera utilizzare IAM per la gestione degli utenti, scegliere Imposta ARN IAM come utente principale e specificare l'ARN per un ruolo IAM.

- Se desideri utilizzare il database utenti interno, scegli Crea utente principale e specifica un nome utente e una password.

Qualunque opzione tu scelga, l'utente principale può accedere a tutti gli indici del cluster e a tutto il resto. OpenSearch APIs Per informazioni su quale opzione scegliere, vedere the section called "Concetti chiave".

Se si disattiva il controllo granulare degli accessi, è comunque possibile controllare l'accesso al dominio inserendolo in un VPC, applicando una policy di accesso restrittivo o entrambi. È necessario abilitare la node-to-node crittografia e la crittografia a riposo per utilizzare un controllo granulare degli accessi.

> **Note**
>
> Si consiglia fortemente di abilitare il controllo granulare degli accessi per proteggere i dati sul dominio. Il controllo granulare degli accessi fornisce protezione a livello di cluster, indice, documento e campo.

21. (Facoltativo) Se desideri utilizzare l'autenticazione SAML per le OpenSearch dashboard, scegli Abilita l'autenticazione SAML e configura le opzioni SAML per il dominio. Per istruzioni, consulta the section called "Autenticazione SAML per dashboard OpenSearch".
22. (Facoltativo) Se desideri utilizzare l'autenticazione Amazon Cognito per OpenSearch dashboard, scegli Abilita l'autenticazione Amazon Cognito. Quindi scegli il pool di utenti e il pool di identità di Amazon Cognito che desideri utilizzare per l'autenticazione di OpenSearch Dashboards. Per ulteriori informazioni sulla creazione di queste risorse, consultare the section called "Autenticazione Amazon Cognito per dashboard OpenSearch ".
23. (Facoltativo) Se desideri utilizzare l'autenticazione IAM Identity Center (IDC) per connettere la tua fonte di identità esistente e offrire alle tue AWS applicazioni una visione comune degli utenti, scegli Abilita l'accesso alle API autenticato con IAM Identity Center. Per ulteriori informazioni, consulta la panoramica sulla propagazione delle identità affidabili nella Guida per l'utente di IAM Identity Center.
24. (Facoltativo) Nella sezione Funzioni avanzate, lascia selezionate le funzionalità Abilita Natural Language Query Generation e Amazon Q Developer, se desideri utilizzare queste funzionalità.

a. Scegli Enable S3 Vectors come opzione del motore per opzioni di ricerca vettoriale avanzate. Per ulteriori informazioni, consulta Funzionalità di ricerca avanzate con un motore vettoriale Amazon S3.

b. Scegli Abilita l'accelerazione GPU per opzioni di ricerca vettoriale avanzate. Per ulteriori informazioni, consulta Accelerazione GPU per l'indicizzazione vettoriale.

25. Per la politica di accesso, scegli una politica di accesso o configurane una personalizzata. Se si sceglie di creare una policy personalizzata, è possibile configurarla manualmente o importarne una da un altro dominio. Per ulteriori informazioni, consultare the section called “Identity and Access Management”.

> **Note**
>
> Se è stato abilitato l'accesso VPC, non è possibile utilizzare policy basate su IP. Invece è possibile utilizzare i gruppi di sicurezza per controllare gli indirizzi IP che possono accedere al dominio. Per ulteriori informazioni, consultare the section called “Informazioni sulle policy d'accesso nei domini VPC”.

26. (Facoltativo) Per richiedere che tutte le richieste al dominio arrivino tramite HTTPS, selezionare Richiedi HTTPS per tutto il traffico verso il dominio. Per abilitare node-to-node la crittografia, seleziona ode-to-nodeCrittografia N. Per ulteriori informazioni, consulta the section called “Node-to-node crittografia”. Per abilitare la crittografia dei dati inattivi, seleziona Abilita la crittografia dei dati inattivi. Queste opzioni sono preselezionate se hai scelto l'opzione di implementazione Multi-AZ with Standby.

27. (Facoltativo) Seleziona Usa chiave AWS proprietaria per fare in modo che OpenSearch Service crei una chiave di AWS KMS crittografia per tuo conto (o utilizzi quella che ha già creato). In caso contrario, scegliere la propria chiave KMS. Per ulteriori informazioni, consulta the section called “Crittografia dei dati a riposo”.

28. Per la finestra non di punta, seleziona un orario di inizio per pianificare gli aggiornamenti del software di servizio e le ottimizzazioni Auto-Tune che richiedono un'implementazione. blue/green Gli aggiornamenti non di punta aiutano a ridurre al minimo il carico sui nodi master dedicati di un cluster durante i periodi di traffico elevato.

29. Per Auto-Tune, scegli se consentire al OpenSearch Servizio di suggerire modifiche alla configurazione relative alla memoria del tuo dominio per migliorare la velocità e la stabilità. Per ulteriori informazioni, consulta the section called “Regolazione automatica”.

(Facoltativo) Seleziona Finestra Off-peak per pianificare una finestra ricorrente durante la quale Auto-Tune aggiorna il dominio.

30. (Facoltativo) Seleziona Aggiornamento automatico del software per abilitare gli aggiornamenti software automatici.
31. (Facoltativo) Aggiungere i tag per descrivere il dominio in modo da poter categorizzare e filtrare in base a tali informazioni. Per ulteriori informazioni, consultare the section called “Assegnazione di tag ai domini”.
32. (Facoltativo) Espandi e configura Impostazioni avanzate del cluster. Per un riepilogo di queste opzioni, vedere the section called “Impostazioni avanzate del cluster”.
33. Scegli Create (Crea).

Creazione OpenSearch di domini di servizio ()AWS CLI

Invece di creare un dominio di OpenSearch servizio utilizzando la console, è possibile utilizzare il AWS CLI. Per la sintassi, consulta Amazon OpenSearch Service nel riferimento ai comandi AWS CLI a.

Comandi di esempio

Questo primo esempio dimostra la seguente configurazione del dominio di OpenSearch servizio:

- Crea un dominio OpenSearch di servizio denominato mylogs con la versione 1.2 OpenSearch
- Popola il dominio con due istanze del tipo `r6g.large.search`
- Utilizza un volume EBS a scopo generico (SSD) `gp3` da 100 GiB per lo storage di ogni nodo di dati
- Consente l'accesso anonimo, ma solo da un solo indirizzo IP: 192.0.2.0/32

```
aws opensearch create-domain \
    --domain-name mylogs \
    --engine-version OpenSearch_1.2 \
    --cluster-config  InstanceType=r6g.large.search,InstanceCount=2 \
    --ebs-options
 EBSEnabled=true,VolumeType=gp3,VolumeSize=100,Iops=3500,Throughput=125 \
    --access-policies '{"Version": "2012-10-17",        "Statement": [{"Action":
 "es:*", "Principal":"*","Effect": "Allow", "Condition": {"IpAddress":{"aws:SourceIp":
["192.0.2.0/32"]}}}]}'
```

L'esempio successivo mostra la seguente configurazione del dominio di OpenSearch servizio:

- Crea un dominio di OpenSearch servizio denominato mylogs con Elasticsearch versione 7.10
- Popola il dominio con sei istanze del tipo `r6g.large.search`
- Utilizza un volume EBS a scopo generico (SSD) gp2 da 100 GiB per lo storage di ogni nodo di dati
- Limita l'accesso al servizio a un singolo utente, identificato dall'ID dell'utente: 5555 Account AWS
- Distribuisce istanze in tre zone di disponibilità

```
aws opensearch create-domain \
    --domain-name mylogs \
    --engine-version Elasticsearch_7.10 \
    --cluster-config
 InstanceType=r6g.large.search,InstanceCount=6,ZoneAwarenessEnabled=true,ZoneAwarenessConfig={A
 \
    --ebs-options EBSEnabled=true,VolumeType=gp2,VolumeSize=100 \
    --access-policies '{"Version": "2012-10-17",        "Statement": [ { "Effect":
 "Allow", "Principal": {"AWS": "arn:aws:iam::555555555555:root" }, "Action":"es:*",
 "Resource": "arn:aws:es:us-east-1:555555555555:domain/mylogs/*" } ] }'
```

L'esempio successivo mostra la seguente configurazione del dominio di OpenSearch servizio:

- Crea un dominio OpenSearch di servizio denominato mylogs con la versione 1.0 OpenSearch
- Popola il dominio con dieci istanze del tipo `r6g.xlarge.search`
- Popola il dominio con tre istanze del tipo `r6g.large.search` per fungere come nodi master dedicati
- Utilizza un volume EBS con capacità di IOPS allocata di 100 GiB per lo storage, configurato con prestazioni di base di 1.000 IOPS per ogni nodo di dati
- Limita l'accesso a un singolo utente e a una sola subresource, l'API `_search`

```
aws opensearch create-domain \
    --domain-name mylogs \
    --engine-version OpenSearch_1.0 \
    --cluster-config
 InstanceType=r6g.xlarge.search,InstanceCount=10,DedicatedMasterEnabled=true,DedicatedMasterTyp
 \
    --ebs-options EBSEnabled=true,VolumeType=io1,VolumeSize=100,Iops=1000 \
```

```
    --access-policies '{"Version": "2012-10-17",        "Statement": [ { "Effect":
"Allow", "Principal": { "AWS": "arn:aws:iam::555555555555:root" }, "Action": "es:*",
"Resource": "arn:aws:es:us-east-1:555555555555:domain/mylogs/_search" } ] }'
```

> **Note**
>
> Se si tenta di creare un dominio di OpenSearch servizio ed esiste già un dominio con lo stesso nome, la CLI non riporta alcun errore. Restituisce invece i dettagli per il dominio esistente.

Creazione OpenSearch di domini di servizio ()AWS SDKs

AWS SDKs (eccetto Android e iOS SDKs) supportano tutte le azioni definite nell'Amazon OpenSearch Service API Reference, tra cui`CreateDomain`. Per il codice di esempio, consulta the section called “Usando il AWS SDKs”. Per ulteriori informazioni sull'installazione e l'utilizzo di AWS SDKs, consulta AWS Software Development Kits.

Creazione OpenSearch di domini di servizio ()AWS CloudFormation

OpenSearch Il servizio è integrato con AWS CloudFormation, un servizio che consente di modellare e configurare le AWS risorse in modo da dedicare meno tempo alla creazione e alla gestione delle risorse e dell'infrastruttura. Crei un modello che descrive il OpenSearch dominio che desideri creare e CloudFormation predispone e configura il dominio per te. Per ulteriori informazioni, inclusi esempi di modelli JSON e YAML per OpenSearch domini, consulta il riferimento ai tipi di risorse di Amazon OpenSearch Service nella Guida per l'utente.AWS CloudFormation

Configurazione delle policy di accesso

Amazon OpenSearch Service offre diversi modi per configurare l'accesso ai tuoi domini OpenSearch di servizio. Per ulteriori informazioni, consultare the section called “Identity and Access Management” e the section called “Controllo granulare degli accessi”.

La console offre le policy d'accesso preconfigurate che è possibile personalizzare per le esigenze specifiche del dominio. Puoi anche importare politiche di accesso da altri domini OpenSearch di servizio. Per informazioni su come queste policy d'accesso interagiscono con l'accesso VPC, consulta the section called “Informazioni sulle policy d'accesso nei domini VPC”.

Per configurare le policy d'accesso (console)

1. Vai sul sito https://aws.amazon.com e scegli Sign In to the Console (Accedi alla console).
2. In Analytics, scegli Amazon OpenSearch Service.
3. Nel pannello di navigazione, in Domini, scegli il dominio che desideri aggiornare.
4. Scegli Operazioni, quindi Modifica configurazione di sicurezza.
5. Modifica il JSON della policy di accesso o importa un'opzione preconfigurata.
6. Scegli Save changes (Salva modifiche).

Migrazione al OpenSearch servizio con Migration Assistant

Migration Assistant for Amazon OpenSearch Service è una soluzione completa che semplifica il processo di migrazione da Elasticsearch o cluster autogestiti a Service. OpenSearch OpenSearch Questo toolkit affronta la complessità operativa delle migrazioni garantendo al contempo l'integrità dei dati e convalidando le prestazioni dopo la migrazione.

Panoramica di

Che tu stia configurando un proof-of-concept ingresso AWS, trasferendo i carichi di lavoro di produzione o eseguendo l'aggiornamento alle OpenSearch versioni più recenti, Migration Assistant fornisce step-by-step indicazioni, best practice e strumenti per sfruttare tutto il potenziale delle migrazioni. OpenSearch

Migration Assistant offre i seguenti vantaggi chiave:

- Migrazione dei metadati: migra i metadati del cluster, incluse le impostazioni degli indici, le mappature dei tipi, i modelli di indice e gli alias
- Migrazione dei dati: migra i dati esistenti dai cluster legacy ai domini di servizio OpenSearch
- Gestione del traffico in tempo reale: intercetta e reindirizza il traffico in tempo reale dai cluster autogestiti ai domini di servizio con una latenza minima OpenSearch
- Replica del traffico: replica il traffico di produzione sui cluster di destinazione per convalidare precisione e prestazioni
- Test delle prestazioni: simula il traffico reale acquisendo e riproducendo i modelli di richiesta per ottimizzare le prestazioni del sistema
- Disponibilità globale: viene implementato nelle regioni più comuni per una copertura e una scalabilità globali AWS

Migration Assistant supporta la migrazione dalle versioni 6.x e 7.x e 1.x e 2.x di Elasticsearch. OpenSearch Per ulteriori informazioni, consulta Percorsi di migrazione supportati.

> **Note**
>
> Nota le seguenti informazioni aggiuntive su Migration Assistant.
>
> - Lo strumento supporta le migrazioni multi-hop (ad esempio, la migrazione da Elasticsearch 5.x a OpenSearch Service 3.x in un unico hop).
> - Puoi ripristinare una migrazione.
> - In alcuni casi d'uso, lo strumento richiede tempi di inattività minimi o nulli.
> - Lo strumento fornisce un backfill ad alte prestazioni senza influire sul cluster di origine.

Scenari di migrazione

L'Assistente alla migrazione è progettato per gestire i seguenti scenari di migrazione:

Migrazione dei metadati

Migra i metadati del cluster, ad esempio impostazioni degli indici, alias e modelli, dal cluster di origine al dominio del servizio di destinazione. OpenSearch

Migrazione Backfill

Migra i dati esistenti o storici da un cluster di origine a un dominio di OpenSearch servizio di destinazione, garantendo che tutti i dati importanti vengano preservati durante la transizione.

Migrazione del traffico in diretta

Replica il traffico in tempo reale continuo dal cluster di origine al dominio di OpenSearch servizio di destinazione, consentendoti di mantenere la disponibilità del servizio durante la migrazione.

> **Important**
>
> Le strategie di migrazione non sono universalmente applicabili. L'Assistente alla migrazione fornisce indicazioni basate sulle migliori pratiche ingegneristiche, ma è necessario valutare i requisiti specifici e testarli a fondo prima di migrare i carichi di lavoro di produzione.

Guida introduttiva a Migration Assistant

Migration Assistant per Amazon OpenSearch Service è disponibile come AWS soluzione con documentazione completa, modelli di distribuzione e codice sorgente. Per iniziare con Migration Assistant:

Per accedere alle risorse di Migration Assistant

1. Consulta la documentazione completa della soluzione nella pagina Panoramica della soluzione Migration Assistant for Amazon OpenSearch Service.
2. Comprendi i costi e i requisiti leggendo la guida alla pianificazione della distribuzione.
3. Implementa la soluzione utilizzando le istruzioni di distribuzione e segui la guida all'uso per eseguire la migrazione.

Per gli sviluppatori e gli utenti esperti, il codice sorgente di Migration Assistant e la documentazione aggiuntiva sono disponibili nell'archivio OpenSearchdelle migrazioni GitHub .

Architettura della soluzione

Quando viene distribuito in AWS, Migration Assistant utilizza diversi AWS servizi per fornire una soluzione di migrazione completa:

- AWS CloudFormation- Fornisce modelli di Infrastructure as Code (IaC) per distribuire e configurare Migration Assistant
- OpenSearch Servizio: il servizio di destinazione per i carichi di lavoro di ricerca e analisi migrati
- Amazon Managed Streaming for Apache Kafka: fornisce funzionalità di elaborazione dei flussi per lo storage durevole e il riutilizzo del traffico HTTP
- Amazon Elastic Container Service: esegue la console di gestione della migrazione e Traffic Replayer in contenitori sicuri e scalabili
- Amazon Elastic File System: fornisce storage persistente scalabile per i dati di richiesta e risposta dai cluster di origine e di destinazione
- Amazon Simple Storage Service: archivia istantanee per le attività di backfill cronologiche e il contenuto dell'infrastruttura come codice

Per informazioni dettagliate sull'architettura, consulta i dettagli dell'architettura di Migration Assistant.

Impostazioni avanzate del cluster

Utilizzare le opzioni avanzate per configurare:

Indici nei corpi delle richieste

Specifica se i riferimenti espliciti agli indici sono permessi all'interno del corpo delle richieste HTTP. L'impostazione di questa proprietà su `false` impedisce agli utenti di aggirare il controllo degli accessi per le risorse secondarie. Per impostazione predefinita, il valore è `true`. Per ulteriori informazioni, consultare the section called “Opzioni avanzate e considerazioni sulle API”.

Allocazione della cache dei dati di campo

Specifica la percentuale di spazio heap Java allocata ai dati del campo. Per impostazione predefinita, questa impostazione è il 20% dell'heap JVM.

> **Note**
>
> Molte query dei clienti effettuano la rotazione degli indici giornalieri. È consigliabile iniziare i test di benchmark con `indices.fielddata.cache.size` configurato al 40% dell'heap JVM per la maggior parte dei casi d'uso. Tuttavia, se si dispone di indici di grandi dimensioni, potrebbe essere necessaria una cache dei dati del campo di grandi dimensioni.

Numero massimo di clausole

Specifica il numero massimo di clausole permesse in una query booleana Lucene. Il valore di default è 1.024. Le query con un numero di clausole superiore a quello permesso generano un errore `TooManyClauses`. Per ulteriori informazioni, consultare la documentazione di Lucene.

Arricchimento semantico automatico per Amazon Service OpenSearch

Introduzione

Amazon OpenSearch Service utilizza la word-to-word corrispondenza (ricerca lessicale) per trovare risultati, in modo simile ad altri motori di ricerca tradizionali. Questo approccio funziona bene per

domande specifiche come codici di prodotto o numeri di modello, ma ha difficoltà con le ricerche astratte in cui la comprensione delle intenzioni dell'utente diventa fondamentale. Ad esempio, quando cerchi «scarpe per la spiaggia», la ricerca lessicale trova le singole parole «scarpe», «spiaggia», «per» e «il» negli articoli del catalogo, quindi potrebbero mancare prodotti pertinenti come «sandali impermeabili» o «calzature da surf» che non contengono i termini di ricerca esatti.

L'arricchimento semantico automatico risolve questa limitazione considerando sia le corrispondenze delle parole chiave sia il significato contestuale delle ricerche. Questa funzionalità comprende l'intento di ricerca e migliora la pertinenza della ricerca fino al 20%. Abilita questa funzionalità per i campi di testo dell'indice per migliorare i risultati di ricerca.

> **Note**
>
> L'arricchimento semantico automatico è disponibile per i domini di OpenSearch servizio che eseguono la versione 2.19 o successiva. Inoltre, anche i domini con OpenSearch versione 2.19 devono disporre dell'ultimo aggiornamento della versione del software di servizio. Attualmente, la funzionalità è disponibile per i domini pubblici e i domini VPC non sono supportati.

Dettagli del modello e benchmark delle prestazioni

Sebbene questa funzionalità gestisca le complessità tecniche dietro le quinte senza esporre il modello sottostante, forniamo trasparenza attraverso una breve descrizione del modello e risultati di benchmark per aiutarvi a prendere decisioni informate sull'adozione delle funzionalità nei vostri carichi di lavoro critici.

L'arricchimento semantico automatico utilizza un modello sparso preaddestrato e gestito dai servizi che funziona in modo efficace senza richiedere regolazioni di precisione personalizzate. Il modello analizza i campi specificati, espandendoli in vettori sparsi basati su associazioni apprese da diversi dati di addestramento. I termini estesi e i relativi pesi di significatività vengono archiviati nel formato di indice Lucene nativo per un recupero efficiente. Abbiamo ottimizzato questo processo utilizzando la modalità solo documento, in cui la codifica avviene solo durante l'inserimento dei dati. Le query di ricerca vengono semplicemente tokenizzate anziché elaborate tramite il modello sparso, il che rende la soluzione conveniente e performante.

La nostra convalida delle prestazioni durante lo sviluppo delle funzionalità ha utilizzato il set di dati MS MARCO per il recupero dei passaggi, con passaggi della lunghezza media di 334 caratteri.

Per quanto riguarda il punteggio di pertinenza, abbiamo misurato il guadagno cumulativo medio Normalized Discounted Cumulative Gain (NDCG) per i primi 10 risultati di ricerca (ndcg @10) sul benchmark BEIR per i contenuti in inglese e il ndcg @10 medio su MIRACL per i contenuti multilingue. Abbiamo valutato la latenza mediante misurazioni lato client del 90° percentile (p90) e la risposta di ricerca p90 ha rilevato dei valori. Questi benchmark forniscono indicatori prestazionali di base sia per la pertinenza della ricerca che per i tempi di risposta. Ecco i principali numeri di riferimento:

- Lingua inglese - Miglioramento della pertinenza del 20% rispetto alla ricerca lessicale. Ha inoltre ridotto la latenza di ricerca P90 del 7,7% rispetto alla ricerca lessicale (BM25 è di 26 ms e l'arricchimento semantico automatico è di 24 ms).
- Multilingue: miglioramento della pertinenza del 105% rispetto alla ricerca lessicale, mentre la latenza di ricerca P90 è aumentata del 38,4% rispetto alla ricerca lessicale (è di 26 ms e l'arricchimento semantico automatico è di 36 ms). BM25

Data la natura unica di ogni carico di lavoro, ti invitiamo a valutare questa funzionalità nel tuo ambiente di sviluppo utilizzando i tuoi criteri di benchmarking prima di prendere decisioni di implementazione.

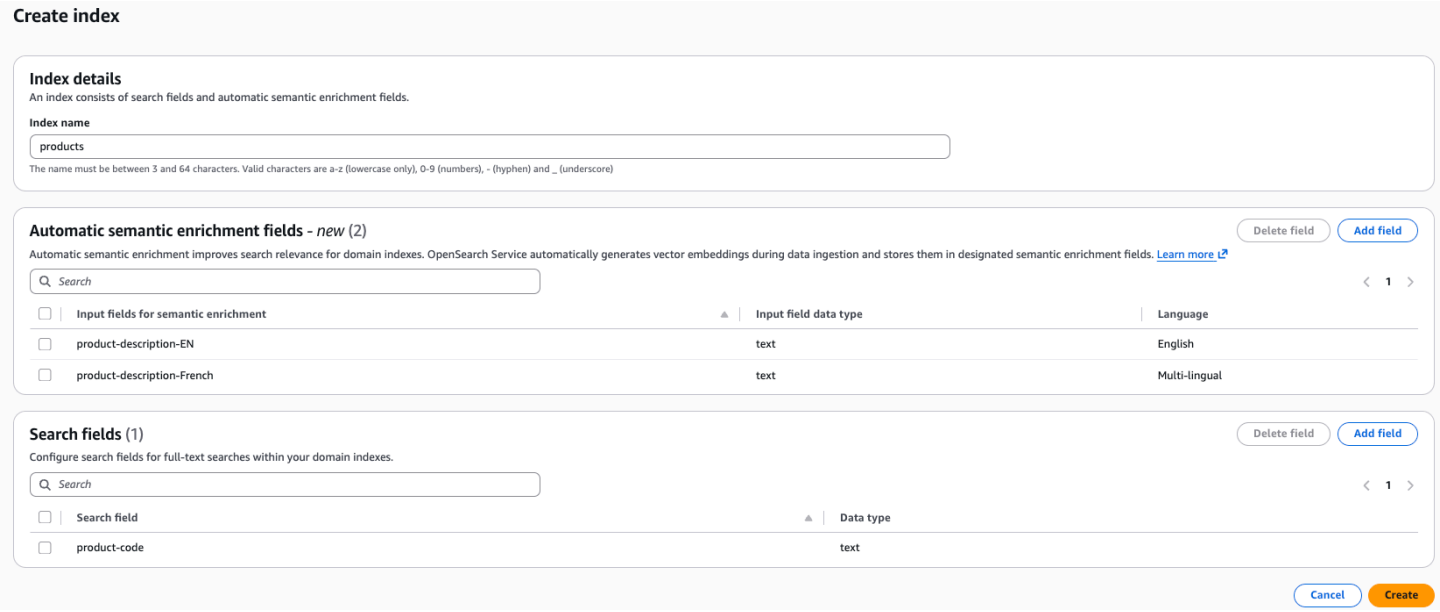
Lingue supportate

La funzionalità supporta l'inglese. Inoltre, il modello supporta anche arabo, bengalese, cinese, finlandese, francese, hindi, indonesiano, giapponese, coreano, persiano, russo, spagnolo, swahili e telugu.

Imposta un indice di arricchimento semantico automatico per i domini

La configurazione di un indice con l'arricchimento semantico automatico abilitato per i campi di testo è semplice e puoi gestirlo tramite la console e i CloudFormation modelli durante la creazione di APIs nuovi indici. Per abilitarlo per un indice esistente, è necessario ricreare l'indice con l'arricchimento semantico automatico abilitato per i campi di testo.

Esperienza da console: la AWS console consente di creare facilmente un indice con campi di arricchimento semantico automatico. Dopo aver selezionato un dominio, troverai il pulsante Crea indice nella parte superiore della console. Dopo aver fatto clic sul pulsante di creazione dell'indice, troverai le opzioni per definire campi di arricchimento semantico automatico. In un indice, puoi avere combinazioni di arricchimento semantico automatico per l'inglese e campi multilingue, oltre che lessicali.



Esperienza API - Per creare un indice di arricchimento semantico automatico utilizzando l'interfaccia a riga di comando ()AWS CLI, usa il AWS comando create-index:

```
aws opensearch create-index \
--domain-name [domain_name] \
--index-name [index_name] \
--index-schema [index_body] \
```

Nell'esempio seguente index-schema, il campo title_semantic ha un tipo di campo impostato su text e il parametro semantic_richment è impostato sullo stato ENABLED. L'impostazione del parametro semantic_richment abilita l'arricchimento semantico automatico nel campo title_semantic. Puoi usare il campo language_options per specificare l'inglese o il multilingue.

```
    aws opensearch create-index \
    --id XXXXXXXXX \
    --index-name 'product-catalog' \
    --index-schema '{
    "mappings": {
        "properties": {
            "product_id": {
                "type": "keyword"
            },
            "title_semantic": {
                "type": "text",
```

```
                "semantic_enrichment": {
                    "status": "ENABLED",
                    "language_options": "english"
                }
            },
            "title_non_semantic": {
                "type": "text"
            }
        }
    }
}'
```

Per descrivere l'indice creato, usa il seguente comando:

```
aws opensearch get-index \
--domain-name [domain_name] \
--index-name [index_name] \
```

Inserimento e ricerca di dati

Dopo aver creato un indice con l'arricchimento semantico automatico abilitato, la funzionalità funziona automaticamente durante il processo di inserimento dei dati, senza bisogno di configurazioni aggiuntive.

Inserimento di dati: quando aggiungi documenti all'indice, il sistema automaticamente:

- Analizza i campi di testo che hai designato per l'arricchimento semantico
- Genera codifiche semantiche utilizzando il modello sparse gestito dal servizio OpenSearch
- Memorizza queste rappresentazioni arricchite insieme ai dati originali

Questo processo utilizza i OpenSearch connettori ML e le pipeline di importazione integrati, che vengono creati e gestiti automaticamente dietro le quinte.

Ricerca: i dati di arricchimento semantico sono già indicizzati, quindi le query vengono eseguite in modo efficiente senza richiamare nuovamente il modello ML. Ciò significa che ottieni una maggiore pertinenza della ricerca senza alcun sovraccarico di latenza di ricerca aggiuntivo.

Configurazione delle autorizzazioni per l'arricchimento semantico automatico

Prima di creare un indice con arricchimento semantico automatico, devi configurare le autorizzazioni richieste. Questa sezione spiega le autorizzazioni necessarie per le diverse operazioni di indicizzazione e come configurarle sia per scenari di controllo degli accessi AWS Identity and Access Management (IAM) che per scenari di controllo degli accessi dettagliati.

autorizzazioni IAM

Le seguenti autorizzazioni IAM sono necessarie per le operazioni di arricchimento semantico automatico. Queste autorizzazioni variano a seconda della specifica operazione di indicizzazione che si desidera eseguire.

CreateIndex Autorizzazioni API

Per creare un indice con arricchimento semantico automatico, sono necessarie le seguenti autorizzazioni IAM:

- `es:CreateIndex`— Crea un indice con funzionalità di arricchimento semantico.
- `es:ESHttpHead`— Esegue richieste HEAD per verificare l'esistenza dell'indice.
- `es:ESHttpPut`— Esegue richieste PUT per la creazione dell'indice.
- `es:ESHttpPost`— Esegue richieste POST per le operazioni sull'indice.

UpdateIndex Autorizzazioni API

Per aggiornare un indice esistente con l'arricchimento semantico automatico, sono necessarie le seguenti autorizzazioni IAM:

- `es:UpdateIndex`— Aggiorna le impostazioni e le mappature dell'indice.
- `es:ESHttpPut`— Esegue richieste PUT per gli aggiornamenti degli indici.
- `es:ESHttpGet`— Esegue richieste GET per recuperare le informazioni sull'indice.
- `es:ESHttpPost`— Esegue richieste POST per le operazioni di indicizzazione.

GetIndex Autorizzazioni API

Per recuperare informazioni su un indice con arricchimento semantico automatico, sono necessarie le seguenti autorizzazioni IAM:

- `es:GetIndex`— Recupera le informazioni e le impostazioni dell'indice.
- `es:ESHttpGet`— Esegue richieste GET per recuperare i dati dell'indice.

DeleteIndex Autorizzazioni API

Per eliminare un indice con arricchimento semantico automatico, sono necessarie le seguenti autorizzazioni IAM:

- `es:DeleteIndex`— Eliminare un indice e i relativi componenti di arricchimento semantico.
- `es:ESHttpDelete`— Esegue richieste DELETE per la rimozione dell'indice.

Policy IAM di esempio

Il seguente esempio di politica di accesso basata sull'identità fornisce le autorizzazioni necessarie a un utente per gestire gli indici con arricchimento semantico automatico:

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowSemanticEnrichmentIndexOperations",
            "Effect": "Allow",
            "Action": [
                "es:CreateIndex",
                "es:UpdateIndex",
                "es:GetIndex",
                "es:DeleteIndex",
                "es:ESHttpHead",
                "es:ESHttpGet",
                "es:ESHttpPut",
                "es:ESHttpPost",
                "es:ESHttpDelete"
            ],
            "Resource": "arn:aws:es:aws-region:111122223333:domain/domain-name"
        }
    ]
}
```

Sostituisci *aws-region* e con i tuoi valori specifici. *111122223333* *domain-name* È possibile limitare ulteriormente l'accesso specificando particolari modelli di indice nell'ARN della risorsa.

Autorizzazioni di controllo degli accessi granulari

Se il tuo dominio Amazon OpenSearch Service ha abilitato il controllo granulare degli accessi, hai bisogno di autorizzazioni aggiuntive oltre alle autorizzazioni IAM. Le seguenti autorizzazioni sono necessarie per ogni operazione di indicizzazione.

CreateIndex Autorizzazioni API

Quando è abilitato il controllo granulare degli accessi, sono necessarie le seguenti autorizzazioni aggiuntive per creare un indice con arricchimento semantico automatico:

- `indices:admin/create`— Creare operazioni di indicizzazione.
- `indices:admin/mapping/put`— Creare e aggiornare le mappature degli indici.
- `cluster:admin/opensearch/ml/create_connector`— Crea connettori di apprendimento automatico per l'elaborazione semantica.
- `cluster:admin/opensearch/ml/register_model`— Registra modelli di apprendimento automatico per l'arricchimento semantico.
- `cluster:admin/ingest/pipeline/put`— Crea pipeline di importazione per l'elaborazione dei dati.
- `cluster:admin/search/pipeline/put`— Crea pipeline di ricerca per l'elaborazione delle query.

UpdateIndex Autorizzazioni API

Quando è abilitato il controllo granulare degli accessi, sono necessarie le seguenti autorizzazioni aggiuntive per aggiornare un indice con arricchimento semantico automatico:

- `indices:admin/get`— Recupera le informazioni sull'indice.
- `indices:admin/settings/update`— Aggiorna le impostazioni dell'indice.
- `indices:admin/mapping/put`— Aggiorna le mappature degli indici.
- `cluster:admin/opensearch/ml/create_connector`— Crea connettori per l'apprendimento automatico.
- `cluster:admin/opensearch/ml/register_model`— Registra modelli di apprendimento automatico.
- `cluster:admin/ingest/pipeline/put`— Crea pipeline di importazione.

- `cluster:admin/search/pipeline/put`— Crea pipeline di ricerca.
- `cluster:admin/ingest/pipeline/get`— Recupera le informazioni sulla pipeline di importazione.
- `cluster:admin/search/pipeline/get`— Recupera le informazioni sulla pipeline di ricerca.

GetIndex Autorizzazioni API

Quando è abilitato il controllo granulare degli accessi, sono necessarie le seguenti autorizzazioni aggiuntive per recuperare informazioni su un indice con arricchimento semantico automatico:

- `indices:admin/get`— Recupera le informazioni sull'indice.
- `cluster:admin/ingest/pipeline/get`— Recupera le informazioni sulla pipeline di importazione.
- `cluster:admin/search/pipeline/get`— Recupera le informazioni sulla pipeline di ricerca.

DeleteIndex Autorizzazioni API

Quando è abilitato il controllo granulare degli accessi, è richiesta la seguente autorizzazione aggiuntiva per eliminare un indice con arricchimento semantico automatico:

- `indices:admin/delete`— Operazioni di eliminazione dell'indice.

Riscritture delle query

L'arricchimento semantico automatico converte automaticamente le query «match» esistenti in query di ricerca semantiche senza richiedere modifiche alle query. Se una query di corrispondenza fa parte di una query composta, il sistema analizza la struttura delle query, trova le query di corrispondenza e le sostituisce con query sparse neurali. Attualmente, la funzionalità supporta solo la sostituzione delle query «match», indipendentemente dal fatto che si tratti di una query autonoma o di parte di una query composta. «multi_match» non è supportato. Inoltre, la funzionalità supporta tutte le query composte per sostituire le relative query di corrispondenza annidate. Le query composte includono: bool, boosting, constant_score, dis_max, function_score e hybrid.

Limitazioni dell'arricchimento semantico automatico

La ricerca semantica automatica è più efficace se applicata a campi di small-to-medium grandi dimensioni contenenti contenuti in linguaggio naturale, come titoli di film, descrizioni di prodotti,

recensioni e riepiloghi. Sebbene la ricerca semantica aumenti la pertinenza per la maggior parte dei casi d'uso, potrebbe non essere ottimale per determinati scenari. Prendi in considerazione le seguenti limitazioni quando decidi se implementare l'arricchimento semantico automatico per il tuo caso d'uso specifico.

- Documenti molto lunghi: l'attuale modello sparso elabora solo i primi 8.192 token di ogni documento in inglese. Per i documenti multilingue, sono 512 token. Per articoli lunghi, prendi in considerazione l'implementazione della suddivisione in blocchi dei documenti per garantire l'elaborazione completa dei contenuti.
- Carichi di lavoro di analisi dei log: l'arricchimento semantico aumenta in modo significativo la dimensione dell'indice, il che potrebbe non essere necessario per l'analisi dei log, dove in genere è sufficiente una corrispondenza esatta. Il contesto semantico aggiuntivo raramente migliora l'efficacia della ricerca nei log abbastanza da giustificare i maggiori requisiti di archiviazione.
- L'arricchimento semantico automatico non è compatibile con la funzionalità Derived Source.
- Limitazione: le richieste di inferenza di indicizzazione sono attualmente limitate a 200 TPS per i domini di servizio. OpenSearch Si tratta di un limite non vincolante; contatta il AWS servizio Support per ottenere limiti più elevati.

Prezzi

Amazon OpenSearch Service fattura l'arricchimento semantico automatico in base alle unità di OpenSearch calcolo (OCUs) consumate durante la generazione di vettori sparsi al momento dell'indicizzazione. Ti viene addebitato solo l'utilizzo effettivo durante l'indicizzazione. Puoi monitorare questo consumo utilizzando la CloudWatch metrica SemanticSearch OCU di Amazon. Per dettagli specifici sui limiti dei token del modello, sul throughput di volume per OCU ed esempi di calcolo di esempio, visita Service Pricing. OpenSearch

Apportare modifiche alla configurazione in Amazon OpenSearch Service

Amazon OpenSearch Service utilizza un processo di distribuzione blu/verde per l'aggiornamento dei domini. Una blue/green distribuzione crea un ambiente inattivo per gli aggiornamenti del dominio che copia l'ambiente di produzione e indirizza gli utenti al nuovo ambiente una volta completati gli aggiornamenti. In una blue/green distribuzione, l'ambiente blu è l'ambiente di produzione corrente. L'ambiente verde è l'ambiente inattivo.

I dati vengono migrati dall'ambiente blu all'ambiente verde. Quando il nuovo ambiente è pronto, OpenSearch Service passa da un ambiente all'altro per far sì che l'ambiente verde diventi il nuovo ambiente di produzione. Il passaggio avviene senza perdita di dati. Questa pratica riduce al minimo i tempi di inattività e mantiene l'ambiente originale nel caso in cui l'implementazione nel nuovo ambiente non abbia esito positivo.

Argomenti

- Modifiche che di solito causano implementazioni blue/green
- Modifiche che di solito non causano blue/green distribuzioni
- Opzioni di implementazione blu/verdi
- Determinare se una modifica provocherà una blue/green distribuzione
- Monitoraggio di una modifica alla configurazione
- Fasi di una modifica della configurazione
- Impatto delle implementazioni sulle prestazioni blue/green
- Costi per le modifiche di configurazione
- Risoluzione degli errori di convalida

Modifiche che di solito causano implementazioni blue/green

Le seguenti operazioni causano blue/green le distribuzioni:

- Modifica del tipo di istanza
- Abilitazione del controllo granulare degli accessi
- Esecuzione degli aggiornamenti software del servizio
- Abilitazione o disabilitazione di nodi master dedicati
- Abilitazione o disabilitazione di Multi-AZ senza Standby
- Modifica del tipo di archiviazione, del tipo di volume o della dimensione del volume
- Scelta di sottoreti VPC diverse
- Aggiunta o rimozione di gruppi di sicurezza VPC
- Aggiungere o rimuovere nodi coordinatori dedicati
- Abilitazione o disabilitazione dell'autenticazione Amazon Cognito per dashboard OpenSearch
- Scelta di un bacino d'utenza o pool di identità di Amazon Cognito differente
- Modifica delle impostazioni avanzate

- Aggiornamento a una nuova OpenSearch versione (le OpenSearch dashboard potrebbero non essere disponibili durante tutto o in parte l'aggiornamento)
- Attivazione della crittografia dei dati archiviati o della crittografia node-to-node
- Attivazione UltraWarm o disabilitazione della conservazione a freddo
- Disabilitazione della regolazione automatica e rollback delle sue modifiche
- Associazione di un plug-in opzionale a un dominio e dissociazione di un plug-in opzionale da un dominio
- Aumento del numero di nodi master dedicati per i domini Multi-AZ con due nodi master dedicati
- Riduzione della dimensione del volume EBS
- Modifica delle dimensioni, degli IOPS o del throughput del volume EBS, se l'ultima modifica apportata è in corso o è avvenuta meno di 6 ore fa
- Attivazione della pubblicazione dei registri di controllo su. CloudWatch

Per i domini Multi-AZ con standby, è possibile effettuare solo una richiesta di modifica alla volta. Se è già in corso una modifica, la nuova richiesta viene rifiutata. Puoi controllare lo stato della modifica corrente con l'`DescribeDomainChangeProgress`API.

Modifiche che di solito non causano blue/green distribuzioni

Nella maggior parte dei casi, le seguenti operazioni non causano blue/green distribuzioni:

- Modifica della politica di accesso
- Modifica dell'endpoint personalizzato
- Modifica della politica Transport Layer Security (TLS)
- Modifica dell'orario di uno snapshot automatico
- Abilitazione o disabilitazione di Require HTTPS (Richiedi HTTPS)
- Abilitazione della regolazione automatica senza il rollback delle sue modifiche
- Se il tuo dominio ha nodi master dedicati, modifica del nodo dati o del numero di UltraWarm nodi
- Se il dominio dispone di nodi master dedicati, modifica del tipo o del numero di istanze master dedicate (ad eccezione dei domini Multi-AZ con due nodi master dedicati)
- Abilitazione o disabilitazione della pubblicazione dei log degli errori o degli slow log su CloudWatch
- Disattivazione della pubblicazione dei registri di controllo su CloudWatch
- Aumento delle dimensioni del volume, degli IOPS o del throughput per il tipo gp3 EBS.

> **Note**
>
> Prima del 10 marzo 2026, gli aumenti di volume in atto erano supportati solo per volumi fino a 3 TiB per gp3. Il 10 marzo 2026, questa limitazione è stata rimossa, consentendo aumenti di volume effettivi oltre i 3 TiB. Tuttavia, se il cluster aveva una dimensione del volume superiore a 3 TiB prima che questa limitazione venisse rimossa, il primo aumento di volume richiederà una blue/green deployment. All subsequent volume increases for that cluster will be performed as in-place updates and will not require a blue/green distribuzione.

- Aggiunta e rimozione di tag

> **Note**
>
> Esistono alcune eccezioni a seconda della versione del software di servizio. Se vuoi essere sicuro che una modifica non provochi una blue/green distribuzione, esegui un'operazione a secco prima di aggiornare il dominio, se questa opzione è disponibile. Alcune modifiche non offrono un'opzione di esecuzione a secco. In genere consigliamo di apportare modifiche al cluster al di fuori delle ore di traffico di punta.

Opzioni di implementazione blu/verdi

Seleziona una strategia di implementazione per controllare il modo in cui il cluster gestisce le distribuzioni quando non è disponibile una capacità sufficiente al momento dell'aggiornamento.

1. Full Swap Blue/Green: il comportamento di distribuzione predefinito. Richiede la piena capacità iniziale dell'istanza, garantendo l'implementazione più rapida quando la capacità è disponibile. L'installazione non procederà se non è possibile allocare una capacità sufficiente.
2. Capacità ottimizzata: consigliata per cluster con più di 30 nodi di dati. Tenta innanzitutto uno blue/green swap completo e, se la capacità è insufficiente, procede con la distribuzione in batch. Garantisce il completamento delle implementazioni anche quando la capacità è limitata. Il tempo di completamento può aumentare, poiché l'implementazione verrà eseguita in batch.

Console

Per Edit Domain Flow:

1. Apri la console Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/home.
2. Nel riquadro di navigazione, in Domini, scegli il nome di dominio per aprire la configurazione del cluster.
3. Fai clic sul pulsante Modifica a destra nella scheda Configurazione del cluster.
4. Dalle opzioni della strategia di distribuzione scegli le configurazioni richieste per l'aggiornamento del dominio:

    1. Full Swap Blue/Green: il comportamento di distribuzione predefinito. Richiede la piena capacità iniziale dell'istanza, garantendo l'implementazione più rapida quando la capacità è disponibile. L'installazione non procederà se non è possibile allocare una capacità sufficiente.
    2. Capacità ottimizzata: consigliata per cluster con più di 30 nodi di dati. Tenta innanzitutto uno blue/green swap completo e, se la capacità è insufficiente, procede con la distribuzione in batch. Garantisce il completamento delle implementazioni anche quando la capacità è limitata. Il tempo di completamento può aumentare, poiché l'implementazione verrà eseguita in batch.
5. Scegli Save changes (Salva modifiche).

Per Create Domain Flow:

1. Apri la console Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/home.
2. Nel riquadro di navigazione a sinistra, scegli Domains (Domini).
3. Fai clic sul pulsante Crea dominio.
4. Seleziona tutte le configurazioni richieste per il dominio.
5. Dalle opzioni della strategia di distribuzione scegli le configurazioni richieste per l'aggiornamento del dominio:

    1. Full Swap Blue/Green: il comportamento di distribuzione predefinito. Richiede la piena capacità iniziale dell'istanza, garantendo l'implementazione più rapida quando la capacità è disponibile. L'installazione non procederà se non è possibile allocare una capacità sufficiente.
    2. Capacità ottimizzata: consigliata per cluster con più di 30 nodi di dati. Tenta innanzitutto uno blue/green swap completo e, se la capacità è insufficiente, procede con la

distribuzione in batch. Garantisce il completamento delle implementazioni anche quando la capacità è limitata. Il tempo di completamento può aumentare, poiché l'implementazione verrà eseguita in batch.

6. Fai clic sul pulsante Crea nel pannello di riepilogo del dominio sul lato destro.

API

È possibile configurare la strategia di implementazione utilizzando l'UpdateDomainConfigAPI.

Opzione ottimizzata per la capacità

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/
config
{
   "DeploymentStrategyOptions": {
    "DeploymentStrategy": "CapacityOptimized"
   }
}
```

Opzione predefinita

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/
config
{
   "DeploymentStrategyOptions": {
    "DeploymentStrategy": "Default"
   }
}
```

Determinare se una modifica provocherà una blue/green distribuzione

È possibile testare alcuni tipi di modifiche alla configurazione pianificate per determinare se causeranno una blue/green distribuzione, senza doversi impegnare ad apportare tali modifiche. Prima di avviare una modifica della configurazione, utilizza la console o un'API per eseguire un controllo di convalida per garantire che il dominio sia idoneo per un aggiornamento.

Console

Per convalidare una modifica alla configurazione

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Domains (Domini).
3. Seleziona il dominio per cui desideri avviare una modifica della configurazione. Si apre la pagina dei dettagli del dominio. Seleziona il menu a discesa Actions (Operazioni), quindi scegli Edit cluster configuration (Modifica configurazione cluster).
4. Apporta modifiche al dominio, ad esempio cambiando il tipo di istanza o il numero di nodi.
5. In Dry run analysis, scegli Esegui. Il dry run convalida la modifica della configurazione alla ricerca di errori e determina se è necessaria un' blue/green implementazione.
6. Al termine del dry run, i risultati vengono visualizzati nella parte inferiore della pagina, insieme a un ID dry run. L'analisi indica se la modifica della configurazione richiede o meno una blue/green distribuzione.

   Ogni ciclo a secco sovrascrive quello precedente. Per conservare i dettagli di ogni corsa, salvate il relativo ID di funzionamento a secco. Le esecuzioni a secco sono disponibili per 90 giorni o fino a quando non si effettua un aggiornamento della configurazione.
7. Per procedere con l'aggiornamento della configurazione, scegli Save changes (Salva modifiche). Altrimenti, scegli Cancel (Annulla). Entrambe le opzioni riportano alla scheda Cluster configuration (Configurazione del cluster). In questa scheda, puoi scegliere Dry run details (Dettagli del test) per visualizzare i dettagli del tuo ultimo test. Questa pagina include anche un side-by-side confronto tra la configurazione prima del funzionamento a secco e la configurazione del funzionamento a secco.

API

È possibile eseguire la convalida del test di analisi anche tramite l'API di configurazione. Per testare le modifiche con l'API, imposta `DryRun` su `true` e `DryRunMode` su `Verbose`. La modalità verbosa esegue un controllo di convalida oltre a determinare se la modifica avvierà una implementazione blu/verde. Ad esempio, questa UpdateDomainConfigrichiesta verifica il tipo di distribuzione risultante dall'attivazione di UltraWarm:

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/
config
```

```
{
   "ClusterConfig": {
    "WarmCount": 3,
    "WarmEnabled": true,
    "WarmType": "ultrawarm1.large.search"
   },
   "DryRun": true,
   "DryRunMode": "Verbose"
}
```

La richiesta esegue un controllo di convalida e restituisce il tipo di implementazione che sarà causata dalla modifica ma non esegue effettivamente l'aggiornamento:

```
{
   "ClusterConfig": {
     ...
    },
   "DryRunResults": {
      "DeploymentType": "Blue/Green",
      "Message": "This change will require a blue/green deployment."
    }
}
```

I possibili tipi di implementazione sono:

- `Blue/Green`: la modifica causerà una implementazione blu/verde.
- `DynamicUpdate`: la modifica non causerà una implementazione blu/verde.
- `Undetermined`: il dominio è ancora in stato di elaborazione, quindi non è possibile determinare il tipo di implementazione.
- `None`: nessuna modifica alla configurazione.

Se la convalida non riesce, viene restituito un elenco di errori di convalida.

```
{
   "ClusterConfig":{
      "..."
   },
   "DryRunProgressStatus":{
      "CreationDate":"2023-01-12T01:14:33.847Z",
      "DryRunId":"db00ca39-48b2-4774-bbd3-252cf094d205",
```

```
      "DryRunStatus":"failed",
      "UpdateDate":"2023-01-12T01:14:33.847Z",
      "ValidationFailures":[
         {
            "Code":"Cluster.Index.WriteBlock",
            "Message":"Cluster has index write blocks."
         }
      ]
   }
}
```

Se lo stato è fissopending, puoi utilizzare l'ID dry run nella UpdateDomainConfig risposta alle DescribeDryRunProgresschiamate successive per verificare lo stato della convalida.

```
GET https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/
dryRun?dryRunId=my-dry-run-id
{
    "DryRunConfig": null,
    "DryRunProgressStatus": {
        "CreationDate": "2023-01-12T01:14:42.998Z",
        "DryRunId": "db00ca39-48b2-4774-bbd3-252cf094d205",
        "DryRunStatus": "succeeded",
        "UpdateDate": "2023-01-12T01:14:49.334Z",
        "ValidationFailures": null
    },
    "DryRunResults": {
        "DeploymentType": "Blue/Green",
        "Message": "This change will require a blue/green deployment."
    }
}
```

Per eseguire un test senza un controllo di convalida, imposta `DryRunMode` su `Basic` quando usi l'API di configurazione.

Python

Il seguente codice Python utilizza l'UpdateDomainConfigAPI per eseguire un controllo di convalida dell'esecuzione a secco e, se il controllo ha esito positivo, chiama la stessa API senza esecuzione a secco per avviare l'aggiornamento. Se il controllo fallisce, lo script stampa l'errore e si interrompe.

```
import time
import boto3
```

```
client = boto3.client('opensearch')

response = client.UpdateDomainConfig(
    ClusterConfig={
        'WarmCount': 3,
        'WarmEnabled': True,
        'WarmCount': 123,
    },
    DomainName='test-domain',
    DryRun=True,
    DryRunMode='Verbose'
)

dry_run_id = response.DryRunProgressStatus.DryRunId

retry_count = 0

while True:

    if retry_count == 5:
        print('An error occured')
        break

    dry_run_progress_response = client.DescribeDryRunProgress('test-domain',
 dry_run_id)
    dry_run_status = dry_run_progress_response.DryRunProgressStatus.DryRunStatus

    if dry_run_status == 'succeeded':
        client.UpdateDomainConfig(
            ClusterConfig={
            'WarmCount': 3,
            'WarmEnabled': True,
            'WarmCount': 123,
        })
        break

    elif dry_run_status == 'failed':
        validation_failures_list =
 dry_run_progress_response.DryRunProgressStatus.ValidationFailures
        for item in validation_failures_list:
            print(f"Code: {item['Code']}, Message: {item['Message']}")
        break
```

```
    retry_count += 1
    time.sleep(30)
```

Monitoraggio di una modifica alla configurazione

È possibile richiedere una modifica alla configurazione alla volta o raggruppare più modifiche in un'unica richiesta. Utilizza i campi Stato di elaborazione del dominio e Stato di modifica della configurazione nella console per tenere traccia delle modifiche alla configurazione. Attendi che lo stato del dominio diventi `Active` tale prima di richiedere ulteriori modifiche.

Un dominio può avere i seguenti stati di elaborazione:

- `Active`— Non è in corso alcuna modifica alla configurazione. È possibile inviare una nuova richiesta di modifica della configurazione.
- `Creating`— Il dominio è in fase di creazione.
- `Modifying`— Sono in corso modifiche alla configurazione, come l'aggiunta di nuovi nodi di dati, EBS, gp3, il provisioning IOPS o la configurazione di chiavi KMS.
- `Upgrading engine version`— È in corso un aggiornamento della versione del motore.
- `Updating service software`— È in corso un aggiornamento del software di servizio.
- `Deleting`— Il dominio viene eliminato.
- `Isolated`— Il dominio è sospeso.

Un dominio può avere i seguenti stati di modifica della configurazione:

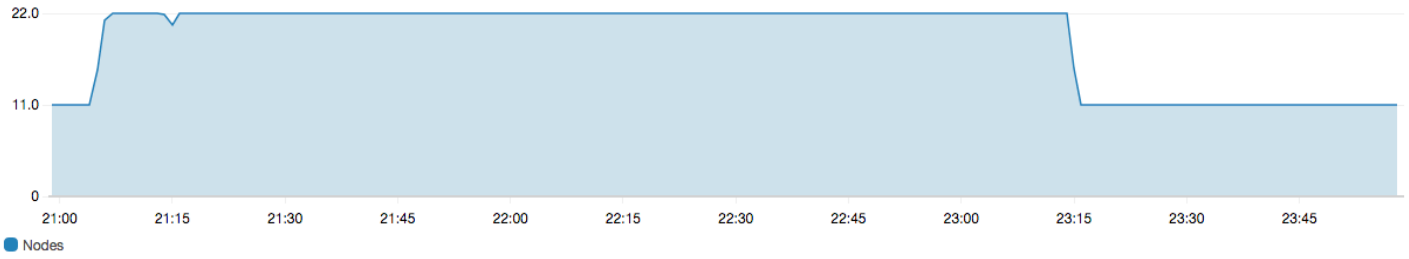
- `Pending`— È stata inviata una richiesta di modifica della configurazione.
- `Initializing`— Il servizio sta inizializzando una modifica alla configurazione.
- `Validating`— Il servizio sta convalidando le modifiche richieste e le risorse richieste.
- `Awaiting user inputs`— Il servizio prevede che le modifiche alla configurazione, ad esempio la modifica del tipo di istanza, continuino. È possibile modificare le modifiche alla configurazione.
- `Applying changes`— Il servizio sta applicando le modifiche di configurazione richieste.
- `Cancelled`— La modifica della configurazione viene annullata. Scegli Annulla e ripristina tutte le modifiche.
- `Completed`— Le modifiche alla configurazione richieste sono state completate con successo.

- `Validation failed`— Le modifiche alla configurazione richieste non sono state completate. Non è stata applicata alcuna modifica alla configurazione.

> **Note**
>
> Gli errori di convalida possono essere il risultato di indici rossi presenti nel dominio, dell'indisponibilità di un tipo di istanza scelto o di uno spazio su disco insufficiente. Per un elenco degli errori di convalida, consulta. the section called “Risoluzione degli errori di convalida” Durante un evento di errore di convalida, è possibile annullare, riprovare o modificare le modifiche alla configurazione.

Una volta completate le modifiche alla configurazione, lo stato del dominio torna a. `Active`

Puoi esaminare lo stato del cluster e i CloudWatch parametri di Amazon e vedere che il numero di nodi nel cluster aumenta temporaneamente, spesso raddoppiando, durante l'aggiornamento del dominio. L'illustrazione seguente mostra il numero di nodi che raddoppia da 11 a 22 durante una modifica di configurazione e torna a 11 al termine dell'aggiornamento.



Questo aumento temporaneo può gravare sui nodi master dedicati del cluster, che improvvisamente devono gestire molti più nodi. Può anche aumentare le latenze di ricerca e indicizzazione poiché OpenSearch Service copia i dati dal vecchio cluster a quello nuovo. È importante mantenere una capacità sufficiente sul cluster per gestire il sovraccarico associato a queste implementazioni. blue/green

> **Important**
>
> Le modifiche di configurazione e manutenzione del servizio non comportano costi aggiuntivi. Ti viene fatturato solo il numero di nodi richiesti per il cluster. Per le specifiche, consulta the section called “Costi per le modifiche di configurazione”.

Per evitare il sovraccarico dei nodi master dedicati, puoi monitorare l'utilizzo con i parametri di Amazon CloudWatch. Per scoprire i valori massimi consigliati, consulta the section called “Allarmi consigliati CloudWatch ”.

Fasi di una modifica della configurazione

Dopo aver avviato una modifica alla configurazione, OpenSearch Service esegue una serie di passaggi per aggiornare il dominio. È possibile visualizzare lo stato di avanzamento della modifica della configurazione in Stato di modifica della configurazione nella console. I passaggi esatti eseguiti da un aggiornamento dipendono dal tipo di modifica che stai apportando.
È inoltre possibile monitorare una modifica di configurazione utilizzando l’operazione API DescribeDomainChangeProgress.

Di seguito sono riportate le fasi possibili che un aggiornamento può passare durante una modifica della configurazione:

Nome fase	Description
Convalida	Convalida che il dominio è idoneo per un aggiornamento e fa emergere problemi di convalida se necessario.
Creazione di un nuovo ambiente	Completamento dei prerequisiti necessari e creazione delle risorse necessarie per avviare la blue/green distribuzione.

Nome fase	Description
Provisioning di nuovi nodi	Creazione di una nuova istanza database nell'ambiente di anteprima.
Routing del traffico su nuovi nodi	Reindiriz zamento del traffico ai nodi dati appena creati.
Routing del traffico sui vecchi nodi	Disabilitazione del traffico sui vecchi nodi dati.
Preparazione dei nodi per la rimozione	Preparazi one alla rimozione dei nodi. Questo passaggio si verifica solo quando si esegue il downscaling del dominio (ad esempio, da 8 nodi a 6 nodi).

Nome fase	Description
Copia di partizioni su nuovi nodi	Spostamento di partizioni dai vecchi nodi ai nuovi nodi.
Terminazione dei nodi	Terminare ed eliminare i vecchi nodi dopo la rimozione delle partizioni.
Eliminazione delle risorse meno recenti	Eliminazione di risorse associate al vecchio ambiente (ad esempio il load balancer).
Aggiornamento dinamico	Viene visualizzato quando l'aggiornamento non richiede una blue/green distribuzione e può essere applicato dinamicamente.

Nome fase	Description
Applicazione di modifiche dedicate relative al master	Visualizzato quando si modifica il tipo o il conteggio dell'istanza principale dedicata.
Applicazione delle modifiche relative al volume	Viene visualizzato quando le dimensioni, il tipo, gli IOPS e la velocità effettiva del volume vengono modificati.

Impatto delle implementazioni sulle prestazioni blue/green

Durante la blue/green distribuzione, il tuo cluster Amazon OpenSearch Service è disponibile per le richieste di ricerca e indicizzazione in entrata. Tuttavia, potresti riscontrare i seguenti problemi di prestazioni:

- Aumento temporaneo dell'utilizzo sui nodi leader poiché i cluster hanno più nodi da gestire.
- Maggiore latenza di ricerca e indicizzazione poiché il OpenSearch servizio copia i dati dai vecchi nodi ai nuovi nodi.
- Aumento dei rifiuti per le richieste in entrata a causa dell'aumento del carico del cluster durante le implementazioni. blue/green
- Per evitare problemi di latenza e il rifiuto delle richieste, è necessario eseguire le blue/green distribuzioni quando il cluster è integro e il traffico di rete è ridotto.

Costi per le modifiche di configurazione

Se si modifica la configurazione di un dominio, OpenSearch Service crea un nuovo cluster come descritto in. the section called “Modifiche di configurazione” Durante la migrazione dal vecchio al nuovo, saranno calcolate le spese seguenti:

- Se si modifica il tipo di istanza, saranno addebitate le spese sia per i cluster che per la prima ora. Dopo la prima ora, l'addebito è relativo solo al nuovo cluster. I volumi EBS non vengono addebitati due volte perché fanno parte del cluster, quindi la loro fatturazione segue la fatturazione delle istanze.

  Esempio: la configurazione viene modificata passando da tre istanze `m3.xlarge` a quattro istanze `m4.large`. Per la prima ora, ti viene addebitato il costo per entrambi i cluster (3 * `m3.xlarge` + 4 * `m4.large`). Dopo la prima ora, ti viene addebitato solo il nuovo cluster (4 * `m4.large`).
- Se non modifichi il tipo di istanza, ti saranno addebitate solo le spese per il cluster più grande per la prima ora. Dopo la prima ora, l'addebito è relativo solo al nuovo cluster.

  Esempio: la configurazione viene modificata passando da sei istanze `m3.xlarge` a tre istanze `m3.xlarge`. Per la prima ora, ti viene addebitato il cluster più grande (6 * `m3.xlarge`). Dopo la prima ora, ti viene addebitato solo il nuovo cluster (3 * `m3.xlarge`).

Risoluzione degli errori di convalida

Quando avvii una modifica alla configurazione o esegui un OpenSearch aggiornamento della versione di Elasticsearch, OpenSearch Service esegue innanzitutto una serie di controlli di convalida per garantire che il tuo dominio sia idoneo per un aggiornamento. Se uno di questi controlli non riesce, si riceve una notifica nella console contenente i problemi specifici che è necessario correggere prima di aggiornare il dominio.

OpenSearch Il servizio isola il dominio se rimane inutilizzabile per più di 60 giorni. Ti invieremo promemoria di notifica per risolvere questi problemi. Se non risolvi i problemi entro questo lasso di tempo, il OpenSearch Servizio elimina il tuo dominio e i relativi dati.

La tabella seguente elenca i possibili problemi di dominio che il OpenSearch Servizio potrebbe riscontrare e i passaggi per risolverli.

Problema	Codice di errore	Fasi per la risoluzione dei problemi
Gruppo di sicurezza non trovato	`SecurityGroupNotFound`	Il gruppo di sicurezza associato al dominio di OpenSearch servizio non esiste. Per risolvere questo problema, crea un gruppo di sicurezza con il nome specificato.
Sottorete non trovata	`SubnetNotFound`	La sottorete associata al dominio OpenSearch di servizio non esiste. Per risolvere questo problema, crea una sottorete nel VPC.
Ruolo collegato al servizio non configurato	`SLRNotConfigured`	Il ruolo collegato al servizio per OpenSearch Service non è configura to. Il ruolo collegato al servizio è predefinito da OpenSearch Service e include tutte le autorizzazioni richieste dal servizio per chiamare altri servizi per conto dell'utente. AWS Se il ruolo non esiste, potresti aver bisogno della Creazione manuale.
Indirizzi IP insufficienti	`InsufficientFreeIPsForSubnets`	Una o più sottoreti VPC non dispongono di indirizzi IP sufficienti per aggiornare il dominio. Per calcolare quanti indirizzi IP sono necessari, vedere the section called “Prenotazione di indirizzi IP in una sottorete VPC”.
Il pool di utenti di Cognito non esiste	`CognitoUserPoolNotFound`	OpenSearch Il servizio non riesce a trovare il pool di utenti di Amazon Cognito. Conferma di averne creato uno e di avere l'ID corretto. Per trovare l'ID, è possibile usare la console Amazon Cognito o il seguente comando della AWS CLI : `aws cognito-idp list-user-pools --max-results 60 --region us-east-1`
Il pool di identità di Cognito non esiste	`CognitoIdentityPoolNotFound`	OpenSearch Il servizio non riesce a trovare il pool di identità di Cognito. Conferma di averne creato uno e di avere l'ID corretto. Per trovare l'ID, è possibile usare la console Amazon Cognito o il seguente comando della AWS CLI : `aws cognito-identity list-identity-pools --max-results 60 --region us-east-1`

Problema	Codice di errore	Fasi per la risoluzione dei problemi
Dominio di Cognito non trovato per il pool di utenti	`CognitoDomainNotFound`	Il pool di utenti non ha un nome di dominio. Puoi configurarne uno utilizzando la console Amazon Cognito o il seguente AWS CLI comando: `aws cognito-idp create-user-pool-domain --domain my-domain --user-pool-id id`
Ruolo Cognito non configurato	`CognitoRoleNotConfigured`	Il ruolo IAM che concede l'autorizzazione al OpenSearch Servizio per configurare i pool di utenti e identità di Amazon Cognito e utilizzarli per l'autenticazione, non è configurato. Configura il ruolo con un set di autorizzazioni e una relazione di trust appropriati. Puoi utilizzare la console, che crea il CognitoAccessForAmazonOpenSearchruolo predefinito per te, oppure puoi configurare manualmente un ruolo utilizzando l'SDK AWS CLI o l' AWS SDK.
Impossibile descrivere il pool di utenti	`UserPoolNotDescribable`	Il ruolo Amazon Cognito specificato non dispone dell'autorizzazione per descrivere il pool di utenti associato al tuo dominio. Assicurarsi che la policy delle autorizzazioni dei ruoli consenta l'operazione `cognito-identity:DescribeUserPool`. Consulta the section called “Informazioni sul ruolo CognitoAccessForAmazonOpenSearch” per la policy completa delle autorizzazioni.
Impossibile descrivere il pool di identità	`IdentityPoolNotDescribable`	Il ruolo Amazon Cognito specificato non dispone dell'autorizzazione per descrivere il pool di identità associato al tuo dominio. Assicurarsi che la policy delle autorizzazioni dei ruoli consenta l'operazione `cognito-identity:DescribeIdentityPool`. Consulta the section called “Informazioni sul ruolo CognitoAccessForAmazonOpenSearch” per la policy completa delle autorizzazioni.

Problema	Codice di errore	Fasi per la risoluzione dei problemi
Impossibile descrivere il pool utente e il pool di identità	`CognitoPoolsNotDescribable`	Il ruolo Amazon Cognito specificato non dispone dell'autorizzazione per descrivere i pool di identità e utente associati al tuo dominio. Assicurarsi che la policy delle autorizzazioni dei ruoli consenta le operazioni `cognito-identity:DescribeIdentityPool` e `cognito-identity:DescribeUserPool`. Consulta the section called “Informazioni sul ruolo CognitoAccessForAmazonOpenSearch” per la policy completa delle autorizzazioni.
Chiave KMS non è abilitata	`KMSKeyNotEnabled`	La chiave AWS Key Management Service (AWS KMS) utilizzata per crittografare il dominio è disabilitata. Riattiva la chiave immediatamente.
Certificato personalizzato non in stato EMESSO	`InvalidCertificate`	Se il tuo dominio utilizza un endpoint personalizzato, lo proteggi generando un certificato SSL in AWS Certificate Manager (ACM) o importandone uno tuo. Lo stato del certificato deve essere Emesso. Se viene visualizzato questo errore, verificare lo stato del certificato nella console ACM. Se lo stato è Scaduto, Non riuscito, Inattivo o In attesa di convalida, consulta ka documentazione per la risoluzione dei problemi di ACM per risolvere il problema.
Capacità insufficiente per avviare il tipo di istanza scelto	`InsufficientInstanceCapacity`	La capacità del tipo di istanza richiesta non è disponibile. Ad esempio, potresti aver richiesto cinque `i3.16xlarge.search` nodi, ma OpenSearch Service non dispone di abbastanza `i3.16xlarge.search` host disponibili, quindi la richiesta non può essere soddisfatta. Controlla i tipi di istanza supportati in OpenSearch Service e scegli un tipo di istanza diverso.
Indici rossi nel cluster	`RedCluster`	Uno o più indici del cluster hanno uno stato rosso, che porta a uno stato generale del cluster rosso. Per risolvere il problema e risolvere questo problema, vedere the section called “Cluster in stato rosso”.

Problema	Codice di errore	Fasi per la risoluzione dei problemi
Interruttore automatico di memoria, troppe richieste	`TooManyRequests`	Ci sono troppe richieste di ricerca e scrittura sul tuo dominio, quindi OpenSearch Service non può aggiornarne la configurazione. Puoi ridurre il numero di richieste, scalare le istanze verticalmente fino a 64 GiB di RAM o scalare orizzontalmente aggiungendo le istanze.
La nuova configurazione non può contenere dati (spazio su disco insufficiente)	`InsufficientStorageCapacity`	Le dimensioni di archiviazione configurate non possono contenere tutti i dati del tuo dominio. Per risolvere questo problema, scegli un volume più grande, elimina gli indici inutilizzati o aumenta il numero di nodi nel cluster per liberare immediatamente spazio su disco.

Partizioni fissate a nodi specifici	`ShardMovementBlocked`	Uno o più indici nel tuo dominio sono collegati a nodi specifici e non possono essere riassegnati. Ciò è probabilmente accaduto perché è stato configurato il filtro di allocazione delle partizioni, che consente di specificare quali nodi sono autorizzati a ospitare le partizioni di un determinato indice.
		
		Per risolvere questo problema, rimuovere i filtri di allocazione delle partizioni da tutti gli indici interessati:
		
		`PUT my-index/_settings`
		`{`
		`  "settings": {`
		`  "index.routing.allocation.require._name": null`
		`  }`
		`}`

Problema	Codice di errore	Fasi per la risoluzione dei problemi
La nuova configurazione non può contenere tutte le partizioni (numero di partizioni)	`TooManyShards`	Il numero di shard sul tuo dominio è troppo alto, il che impedisce a OpenSearch Service di spostarli nella nuova configurazione. Per risolvere questo problema, ridimensiona il tuo dominio orizzontalmente aggiungendo nodi dello stesso tipo di configurazione dei nodi del cluster corrente. Nota che la dimensione massima dei volumi EBS dipende dal tipo di istanza del nodo. Per evitare questo problema in futuro, vedere the section called "Scelta del numero di partizioni" e definire una strategia di partizione appropriata per il tuo caso d'uso.
La sottorete associata al tuo dominio non supporta gli indirizzi IPv4	`ResultCodeIPv4BlockNotExists`	Per risolvere questo problema, crea una sottorete o aggiorna la sottorete esistente nel tuo VPC in base al tipo di indirizzo IP configurato del dominio. Se il dominio utilizza un IPv4solo tipo di indirizzo, utilizza una IPv4 sottorete -only. Se il tuo dominio utilizza la modalità Dual-stack, utilizza una sottorete dual-stack.
La sottorete associata al tuo dominio non supporta gli indirizzi IPv6	`ResultCodeIPv6BlockNotExists`	Per risolvere questo problema, crea una sottorete o aggiorna la sottorete esistente nel tuo VPC in base al tipo di indirizzo IP configurato del dominio. Se il dominio utilizza un IPv4solo tipo di indirizzo, utilizza una IPv4 sottorete -only. Se il tuo dominio utilizza la modalità Dual-stack, utilizza una sottorete dual-stack.

Aggiornamenti del software di servizio in Amazon OpenSearch Service

> **Note**
>
> Per le spiegazioni delle modifiche e delle aggiunte apportate in ogni aggiornamento principale del software del servizio (senza patch), consulta le note di rilascio.

Amazon OpenSearch Service rilascia regolarmente aggiornamenti del software di servizio che aggiungono funzionalità o migliorano in altro modo i tuoi domini. Il pannello Notifications (Notifiche) nella console è il modo più semplice per verificare se è disponibile un aggiornamento o controllare lo stato di un aggiornamento. Ogni notifica include dettagli sull'aggiornamento del software del servizio. Tutti gli aggiornamenti del software di servizio utilizzano implementazioni blu/verdi per ridurre al minimo i tempi di inattività.

Gli aggiornamenti del software di servizio differiscono dagli aggiornamenti della versione del servizio. OpenSearch Per informazioni sull'aggiornamento a una versione successiva del OpenSearch Servizio, consulta. the section called “Aggiornamento dei domini”

OpenSearch Il servizio richiede l'applicazione degli aggiornamenti software di servizio richiesti entro 30 giorni dalla loro disponibilità. Questi aggiornamenti sono fondamentali per mantenere la conformità alla sicurezza.

Se non applichi gli aggiornamenti richiesti entro il periodo di 30 giorni, riceverai notifiche di promemoria ogni 15 giorni per 30 giorni. Dopo questo periodo di inadempienza, il tuo dominio verrà isolato con i seguenti effetti:

- Tutti gli accessi di rete al tuo dominio vengono rimossi
- Lo stato del dominio diventa isolato
- Il dominio rimane inutilizzabile finché non applichi gli aggiornamenti richiesti

Durante l'isolamento, continuerai a ricevere notifiche di promemoria ogni 15 giorni per 60 giorni. Se non applichi gli aggiornamenti richiesti entro questo periodo, il dominio del OpenSearch Servizio e tutti i dati associati verranno eliminati definitivamente. Per ulteriori informazioni, consulta Risoluzione degli errori di convalida.

Aggiornamenti facoltativi e obbligatori

OpenSearch Il servizio prevede due ampie categorie di aggiornamenti software di servizio:

Aggiornamenti opzionali

Gli aggiornamenti opzionali del software di servizio generalmente includono miglioramenti e supporto per nuove caratteristiche o funzionalità. Gli aggiornamenti opzionali non vengono applicati ai tuoi domini e non è prevista una scadenza fissa per installarli. La disponibilità dell'aggiornamento viene comunicata tramite e-mail e una notifica della console. Puoi scegliere di applicare l'aggiornamento immediatamente o riprogrammarlo per una data e un'ora più appropriate. Puoi anche programmarlo durante la finestra non di punta del dominio. La maggior parte degli aggiornamenti software è facoltativa.

Indipendentemente dal fatto che pianifichi o meno un aggiornamento, se apporti una modifica al dominio che causa una distribuzione blu/verde, OpenSearch Service aggiorna automaticamente il software di servizio per te.

Puoi configurare il tuo dominio per applicare automaticamente gli aggiornamenti opzionali durante le ore non di punta. Quando questa opzione è attivata, il OpenSearch Servizio attende almeno 13 giorni dalla data in cui è disponibile un aggiornamento opzionale e quindi pianifica l'aggiornamento dopo sette giorni. Riceverai una notifica sulla console quando l'aggiornamento è pianificato e puoi scegliere di riprogrammarlo per una data successiva.

Per attivare gli aggiornamenti software automatici, seleziona Abilita l'aggiornamento automatico del software quando crei o aggiorni il tuo dominio. Per configurare la stessa impostazione utilizzando AWS CLI, imposta su `--software-update-options true` quando crei o aggiorni il dominio.

Aggiornamenti richiesti

Gli aggiornamenti software di servizio richiesti generalmente includono correzioni di sicurezza critiche o altri aggiornamenti obbligatori per garantire l'integrità e la funzionalità continue del dominio. Esempi di aggiornamenti richiesti sono le vulnerabilità e le esposizioni comuni di Log4j (CVEs) e l'applicazione della versione 2 di Instance Metadata Service (). IMDSv2 Il numero di aggiornamenti obbligatori in un anno è in genere inferiore a tre.

OpenSearch Il servizio pianifica automaticamente questi aggiornamenti e avvisa l'utente sette giorni prima dell'aggiornamento pianificato tramite e-mail e una notifica della console. Puoi scegliere di applicare l'aggiornamento immediatamente o riprogrammarlo per una data e un'ora più appropriate

entro il periodo di tempo consentito. Puoi anche programmarlo durante la prossima finestra non di punta del dominio. Se non intraprendi alcuna azione su un aggiornamento richiesto e non apporti modifiche al dominio che causino una blue/green distribuzione, OpenSearch Service può avviare l'aggiornamento in qualsiasi momento oltre la scadenza specificata (in genere 14 giorni dalla disponibilità), entro la finestra non di punta del dominio.

Indipendentemente da quando è pianificato l'aggiornamento, se apporti una modifica al dominio che causa una distribuzione blu/verde, il OpenSearch Servizio aggiorna automaticamente il dominio per te.

Aggiornamenti delle patch

Le versioni del software di servizio che terminano con «-P» e un numero, come R20211203-*P4*, sono versioni di patch. È probabile che le patch includano miglioramenti delle prestazioni, correzioni di bug lievi e correzioni di sicurezza o miglioramenti della posizione. I rilasci di patch non includono nuove funzionalità o modifiche sostanziali e in genere non hanno un impatto diretto o evidente sugli utenti. La notifica del software di servizio indica se il rilascio di una patch è facoltativo o obbligatorio.

Considerazioni

Per stabilire se aggiornare il dominio, considerare quando segue:

- L'aggiornamento manuale del dominio consente di sfruttare più rapidamente le nuove funzionalità. Quando scegli Aggiorna, OpenSearch Service mette la richiesta in coda e avvia l'aggiornamento quando è disponibile.
- Quando si avvia un aggiornamento del software di servizio, OpenSearch Service invia una notifica quando l'aggiornamento inizia e quando viene completato.
- Gli aggiornamenti software utilizzano le blue/green distribuzioni per ridurre al minimo i tempi di inattività. Gli aggiornamenti possono sovraccaricare temporaneamente i nodi principali dedicati di un cluster, quindi assicurarsi di mantenere una capacità sufficiente per gestire il sovraccarico associato.
- Gli aggiornamenti vengono generalmente completati in pochi minuti, ma possono richiedere anche diverse ore o addirittura giorni se il sistema subisce un carico pesante. Valuta la possibilità di aggiornare il dominio durante la finestra configurata non di punta per evitare lunghi periodi di aggiornamento.

Avvio di un aggiornamento del software di servizio

È possibile richiedere un aggiornamento del software di OpenSearch servizio tramite la console di servizio AWS CLI, il, o uno dei SDKs.

Console

Per richiedere un aggiornamento del software di servizio

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Seleziona il nome di dominio per aprirne la configurazione.
3. Scegli Azioni, Aggiorna e seleziona una delle seguenti opzioni:

   - Applica l'aggiornamento ora: pianifica immediatamente l'azione in modo che venga eseguita nell'ora corrente, se c'è capacità disponibile. Se la capacità non è disponibile, forniamo altre fasce orarie disponibili tra cui scegliere.
   - Pianificalo in una finestra non di punta: disponibile solo se la finestra non di punta è abilitata per il dominio. Pianifica l'aggiornamento in modo che avvenga durante la finestra non di punta configurata del dominio. Non è garantito che l'aggiornamento avvenga nella finestra immediata successiva. A seconda della capacità, potrebbe verificarsi nei giorni successivi. Per ulteriori informazioni, consulta the section called “Finestre non di punta”.
   - Pianifica per data e ora specifiche: pianifica l'aggiornamento in modo che avvenga in una data e un'ora specifiche. Se l'ora specificata non è disponibile per motivi di capacità, puoi selezionare una fascia oraria diversa.

   Se pianifichi l'aggiornamento per una data successiva (all'interno o all'esterno della finestra non di punta del dominio), puoi riprogrammarlo in qualsiasi momento. Per istruzioni, consulta the section called “Riprogrammazione delle azioni”.
4. Scegli Conferma.

AWS CLI

Invia una start-service-software-update AWS CLI richiesta per avviare un aggiornamento del software di servizio. Questo esempio aggiunge immediatamente l'aggiornamento alla coda:

```
aws opensearch start-service-software-update \
  --domain-name my-domain \
```

```
  --schedule-at "NOW"
```

Risposta:

```
{
    "ServiceSoftwareOptions": {
        "CurrentVersion": "R20220928-P1",
        "NewVersion": "R20220928-P2",
        "UpdateAvailable": true,
        "Cancellable": true,
        "UpdateStatus": "PENDING_UPDATE",
        "Description": "",
        "AutomatedUpdateDate": "1969-12-31T16:00:00-08:00",
        "OptionalDeployment": true
    }
}
```

> **Tip**
>
> Dopo aver richiesto un aggiornamento, hai a disposizione una finestra di tempo ristretta per annullarlo. La durata di questo `PENDING_UPDATE` stato può variare notevolmente e dipende dall'utente Regione AWS e dal numero di aggiornamenti simultanei eseguiti dal OpenSearch Servizio. Per annullare un aggiornamento, usa la console o il `cancel-service-software-update` AWS CLI comando.

Se la richiesta ha esito negativo con un`BaseException`, significa che l'ora specificata non è disponibile per motivi di capacità e devi specificare un'ora diversa. OpenSearch Il servizio fornisce suggerimenti alternativi sugli slot disponibili nella risposta.

AWS SDKs

Questo script Python di esempio utilizza i metodi describe_domain e start_service_software_update di per verificare se un dominio è idoneo AWS SDK per Python (Boto3) per un aggiornamento del software di servizio e, in tal caso, avvia l'aggiornamento. È necessario fornire un valore per `domain_name`:

```
import boto3
from botocore.config import Config
import time
```

```
# Build the client using the default credential configuration.
# You can use the CLI and run 'aws configure' to set access key, secret
# key, and default region.

my_config = Config(
    # Optionally lets you specify a Region other than your default.
    region_name='us-east-1'
)

domain_name = ''  # The name of the domain to check and update

client = boto3.client('opensearch', config=my_config)


def getUpdateStatus(client):
    """Determines whether the domain is eligible for an update"""
    response = client.describe_domain(
        DomainName=domain_name
    )
    sso = response['DomainStatus']['ServiceSoftwareOptions']
    if sso['UpdateStatus'] == 'ELIGIBLE':
        print('Domain [' + domain_name + '] is eligible for a service software update
 from version ' +
              sso['CurrentVersion'] + ' to version ' + sso['NewVersion'])
        updateDomain(client)
    else:
        print('Domain is not eligible for an update at this time.')


def updateDomain(client):
    """Starts a service software update for the eligible domain"""
    response = client.start_service_software_update(
        DomainName=domain_name
    )
    print('Updating domain [' + domain_name + '] to version ' +
          response['ServiceSoftwareOptions']['NewVersion'] + '...')
    waitForUpdate(client)


def waitForUpdate(client):
    """Waits for the domain to finish updating"""
    response = client.describe_domain(
        DomainName=domain_name
    )
```

```
    status = response['DomainStatus']['ServiceSoftwareOptions']['UpdateStatus']
    if status == 'PENDING_UPDATE' or status == 'IN_PROGRESS':
        time.sleep(30)
        waitForUpdate(client)
    elif status == 'COMPLETED':
        print('Domain [' + domain_name +
              '] successfully updated to the latest software version')
    else:
        print('Domain is not currently being updated.')

def main():
    getUpdateStatus(client)
```

Pianificazione degli aggiornamenti software durante le finestre non di punta

Ogni dominio OpenSearch di servizio creato dopo il 16 febbraio 2023 ha una finestra giornaliera di 10 ore tra le 22:00 e le 8:00 ora locale, periodo considerato non di punta. OpenSearch Il servizio utilizza questa finestra per pianificare gli aggiornamenti del software di servizio per il dominio. Gli aggiornamenti non di punta aiutano a ridurre al minimo il carico sui nodi master dedicati di un cluster durante i periodi di traffico più intenso. OpenSearch Il servizio non può avviare aggiornamenti al di fuori di questa finestra di 10 ore senza il tuo consenso.

- Per gli aggiornamenti opzionali, il OpenSearch Servizio notifica all'utente la disponibilità dell'aggiornamento e richiede all'utente di pianificare l'aggiornamento durante una finestra non di punta imminente.
- Per gli aggiornamenti richiesti, il OpenSearch Servizio pianifica automaticamente l'aggiornamento durante una prossima finestra non di punta e invia una notifica all'utente con tre giorni di anticipo. Puoi riprogrammare l'aggiornamento (all'interno o all'esterno della finestra non di punta), ma solo entro il periodo di tempo richiesto per il completamento dell'aggiornamento.

Per ogni dominio, puoi scegliere di sostituire l'ora di inizio predefinita delle 22:00 con un'ora personalizzata. Per istruzioni, consulta the section called “Configurazione di una finestra personalizzata non di punta”.

Console

Per pianificare un aggiornamento durante una prossima finestra non di punta

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.

2. Seleziona il nome di dominio per aprirne la configurazione.
3. Scegli Operazioni, Aggiorna.
4. Seleziona Pianificalo in una finestra non di punta.
5. Scegli Conferma.

Puoi visualizzare l'azione pianificata nella scheda della finestra Non di punta e riprogrammarla in qualsiasi momento. Per informazioni, consulta the section called “Visualizzazione delle azioni pianificate”.

CLI

Per pianificare un aggiornamento durante una prossima finestra non di punta utilizzando il AWS CLI, invia una StartServiceSoftwareUpdaterichiesta e specifica il parametro: `OFF_PEAK_WINDOW --schedule-at`

```
aws opensearch start-service-software-update \
  --domain-name my-domain \
  --schedule-at "OFF_PEAK_WINDOW"
```

Aggiornamenti del software del servizio di monitoraggio

OpenSearch Il servizio invia una notifica quando un aggiornamento del software di servizio è disponibile, richiesto, avviato, completato o non riuscito. È possibile visualizzare queste notifiche nel pannello Notifiche della console di OpenSearch servizio. La gravità della notifica è `Informational` se l'aggiornamento è facoltativo e `High` se invece è obbligatorio.

OpenSearch Il servizio invia anche eventi software di servizio ad Amazon EventBridge. Puoi utilizzarlo EventBridge per configurare regole che inviano un'e-mail o eseguono un'azione specifica quando viene ricevuto un evento. Per un esempio, consulta la guidathe section called “Tutorial: Invio di avvisi SNS per gli aggiornamenti disponibili”.

Per vedere il formato di ogni evento software di servizio inviato ad Amazon EventBridge, consultathe section called “Eventi di aggiornamento del software di servizio”.

Quando i domini non sono idonei per un aggiornamento

Il dominio potrebbe non essere idoneo all'aggiornamento del software del servizio se si trova in uno degli stati riportati di seguito:

Stato	Description
Dominio in elaborazione	Il dominio è nel mezzo di una modifica di configurazione. Controlla l'idoneità dell'aggiornamento al termine dell'operazione.
Cluster in stato rosso	Uno o più indici nel cluster sono in stato rosso. Per la risoluzione dei problemi, consulta the section called “Cluster in stato rosso”.
Elevata percentuale di errori	Il OpenSearch cluster restituisce un numero elevato di 5 xx errori quando tenta di elaborare le richieste. Questo problema è in genere il risultato di un numero eccessivo di richieste di lettura o scrittura simultanee. Valuta la possibilità di ridurre il traffico verso il cluster o di eseguire il dimensionamento del dominio.
Split brain	Il cervello diviso significa che il OpenSearch cluster ha più di un nodo master ed è suddiviso in due cluster che non si ricongiungeranno mai da soli. Puoi evitare lo split brain utilizzando il numero consigliato di nodi master dedicati. Per risolvere un problema di split brain, contatta Supporto.
Problema di integrazione di Amazon Cognito	Il tuo dominio utilizza l'autenticazione per OpenSearch le dashboard e OpenSearch Service non riesce a trovare una o più risorse Amazon Cognito. Questo problema in genere si verifica se manca il bacino d'utenza di Amazon Cognito. Per correggere il problema, ricrea la risorsa mancante e configura il dominio del OpenSearch servizio per utilizzarla.
Altro problema del servizio	Problemi relativi al OpenSearch Servizio stesso potrebbero far sì che il dominio venga visualizzato come non idoneo per un aggiornamento. Se nessuna delle condizioni precedenti si applica al dominio e il problema persiste per più di un giorno, contatta Supporto.

Definizione delle finestre non di punta per Amazon Service OpenSearch

Quando crei un dominio Amazon OpenSearch Service, definisci una finestra giornaliera di 10 ore considerata non di punta. OpenSearch Il servizio utilizza questa finestra per pianificare

gli aggiornamenti del software di servizio e le ottimizzazioni Auto-Tune che richiedono un'implementazione blu/verde durante periodi di traffico relativamente inferiori, quando possibile. Blue/green si riferisce al processo di creazione di un nuovo ambiente per gli aggiornamenti del dominio e di indirizzamento degli utenti verso il nuovo ambiente dopo il completamento di tali aggiornamenti.

Sebbene blue/green le distribuzioni non comportino interruzioni, per ridurre al minimo qualsiasi potenziale impatto sulle prestazioni dovuto al consumo di risorse per una blue/green distribuzione, si consiglia di pianificare queste distribuzioni durante la finestra non di picco configurata per il dominio. Gli aggiornamenti, come le sostituzioni dei nodi o quelli che devono essere distribuiti immediatamente nel dominio, non utilizzano la finestra di picco.

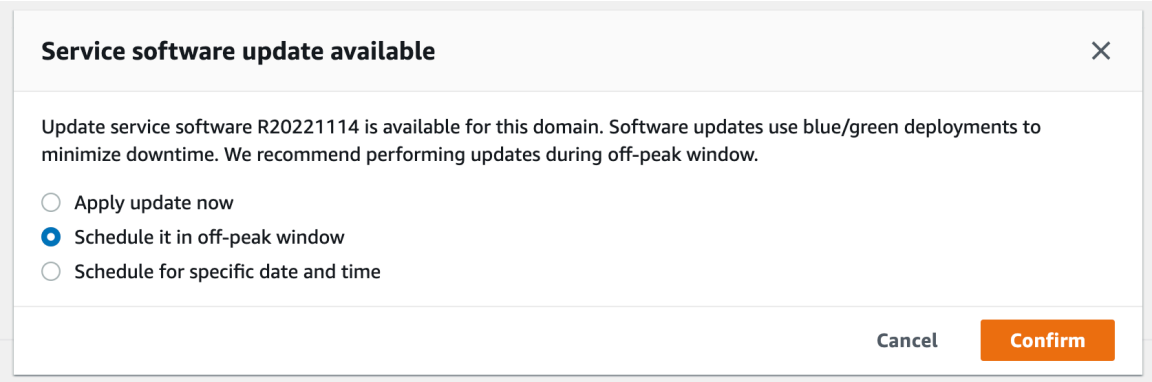
È possibile modificare l'ora di inizio della finestra non di punta, ma non la lunghezza della finestra.

> **Note**
>
> Le finestre non di punta sono state introdotte il 16 febbraio 2023. Per impostazione predefinita, per tutti i domini creati prima di questa data la finestra non di punta è disattivata. È necessario abilitare e configurare manualmente la finestra non di punta per questi domini. Per tutti i domini creati dopo questa data, la finestra non di punta sarà abilitata per impostazione predefinita. Non puoi disabilitare la finestra non di punta per un dominio dopo che è stata abilitata.

Aggiornamenti software di servizio non di punta

OpenSearch Il servizio prevede due ampie categorie di aggiornamenti del software di servizio: facoltativi e obbligatori. Entrambi i tipi richiedono blue/green implementazioni. Gli aggiornamenti opzionali non vengono applicati ai tuoi domini, mentre gli aggiornamenti richiesti vengono installati automaticamente se non intraprendi alcuna azione prima della scadenza specificata (in genere due settimane dalla disponibilità). Per ulteriori informazioni, consulta the section called “Aggiornamenti facoltativi e obbligatori”.

Quando avvii un aggiornamento facoltativo, puoi scegliere di applicarlo immediatamente, pianificarlo per una finestra successiva non di punta o specificare una data e un'ora personalizzate per applicarlo.



Per gli aggiornamenti richiesti, il OpenSearch Servizio pianifica automaticamente una data e un'ora durante le ore non di punta per eseguire l'aggiornamento. Riceverai una notifica tre giorni prima dell'aggiornamento pianificato e puoi scegliere di riprogrammarlo per una data e un'ora successive entro il periodo di distribuzione richiesto. Per istruzioni, consulta the section called “Riprogrammazione delle azioni”.

Ottimizzazioni Auto-Tune in caso di picco

In precedenza, Auto-Tune utilizzava finestre di manutenzione per pianificare le modifiche che richiedevano un'implementazione. blue/green I domini che avevano già Auto-Tune e le finestre di manutenzione abilitate prima dell'introduzione di finestre non di punta continueranno a utilizzare le finestre di manutenzione per questi aggiornamenti, a meno che non si effettui la migrazione per utilizzare la finestra non di punta.

Ti consigliamo di migrare i domini per utilizzare la finestra non di punta, poiché viene utilizzata per pianificare altre attività sul dominio, come gli aggiornamenti del software di servizio. Per istruzioni, consulta the section called “Migrazione dalle finestre di manutenzione di Auto-Tune”. Non puoi tornare a utilizzare le finestre di manutenzione dopo aver migrato il dominio alla finestra non di punta.

Tutti i domini creati dopo il 16 febbraio 2023 utilizzeranno la finestra non di punta, anziché le finestre di manutenzione precedenti, per pianificare le distribuzioni. blue/green Non puoi disabilitare la finestra non di punta per un dominio. Per un elenco delle ottimizzazioni Auto-Tune che richiedono implementazioni blu/verdi, consulta. the section called “Tipi di modifiche”

Attivazione della finestra non di punta

Per tutti i domini creati prima del 16 febbraio 2023 (quando sono state introdotte le finestre non di punta) la funzionalità è disattivata per impostazione predefinita. È necessario abilitarla manualmente per questi domini. Non puoi disabilitare la finestra non di punta dopo averla abilitata.

Console

Per abilitare la finestra non di punta per un dominio

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Seleziona il nome del dominio per aprirne la configurazione.
3. Vai alla scheda della finestra Off-peak e scegli Modifica.
4. Specificate un'ora di inizio personalizzata in UTC (Coordinated Universal Time). Ad esempio, per configurare l'ora di inizio delle 23:30 nella regione Stati Uniti occidentali (Oregon), specificare 07:30.
5. Scegli Save changes (Salva modifiche).

CLI

Per modificare la finestra non di punta utilizzando il, invia una richiesta: AWS CLIUpdateDomainConfig

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --off-peak-window-options 'Enabled=true,
 OffPeakWindow={WindowStartTime={Hours=02,Minutes=00}}'
```

Se non specifichi un'ora di inizio della finestra personalizzata, il valore predefinito è 00:00 UTC.

Configurazione di una finestra personalizzata non di punta

Specifichi una finestra non di punta personalizzata per il tuo dominio in UTC (Coordinated Universal Time). Ad esempio, se desideri che la finestra non di punta inizi alle 23:00 per un dominio nella regione Stati Uniti orientali (Virginia settentrionale), devi specificare le 04:00 UTC.

Console

Per modificare la finestra non di punta per un dominio

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Seleziona il nome del dominio per aprirne la configurazione.
3. Vai alla scheda della finestra Off-peak. Puoi visualizzare la finestra configurata non di punta e un elenco delle azioni pianificate imminenti per il dominio.

4. Scegli Modifica e specifica una nuova ora di inizio in UTC. Ad esempio, per configurare l'ora di inizio delle 21:00 nella regione Stati Uniti orientali (Virginia settentrionale), specifica 02:00 UCT.
5. Scegli Save changes (Salva modifiche).

CLI

Per configurare una finestra non di punta personalizzata utilizzando AWS CLI, invia una UpdateDomainConfigrichiesta e specifica l'ora e il minuto nel formato orario di 24 ore.

Ad esempio, la seguente richiesta modifica l'ora di inizio della finestra alle 2:00 UTC:

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --off-peak-window-options 'OffPeakWindow={WindowStartTime={Hours=02,Minutes=00}}'
```

Se non specifichi l'ora di inizio della finestra, l'impostazione predefinita è alle 22:00 ora locale in cui è stato creato il Regione AWS dominio.

Visualizzazione delle azioni pianificate

Puoi visualizzare tutte le azioni attualmente pianificate, in corso o in sospeso per ciascuno dei tuoi domini. Le azioni possono avere una gravità di `HIGH`, `MEDIUM`, e. `LOW`

Le azioni possono avere i seguenti stati:

- `Pending update`— L'azione è in coda per essere elaborata.
- `In progress`— L'azione è attualmente in corso.
- `Failed`— L'azione non è stata completata.
- `Completed`— L'azione è stata completata con successo.
- `Not eligible`— Solo per gli aggiornamenti del software di servizio. L'aggiornamento non può procedere perché lo stato del cluster non è integro.
- `Eligible`— Solo per gli aggiornamenti del software di servizio. Il dominio è idoneo per un aggiornamento.

Console

La console OpenSearch di servizio mostra tutte le azioni pianificate all'interno della configurazione del dominio, insieme alla gravità e allo stato corrente di ciascuna azione.

Per visualizzare le azioni pianificate per un dominio

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Seleziona il nome del dominio per aprirne la configurazione.
3. Vai alla scheda della finestra Off-peak.
4. In Azioni pianificate, visualizza tutte le azioni attualmente pianificate, in corso o in sospeso per il dominio.

CLI

Per visualizzare le azioni pianificate utilizzando AWS CLI, invia una ListScheduledActionsrichiesta:

```
aws opensearch list-scheduled-actions \
  --domain-name my-domain
```

Risposta:

```
{
    "ScheduledActions": [
        {
            "Cancellable": true,
            "Description": "The Deployment type is : BLUE_GREEN.",
            "ID": "R20220721-P13",
            "Mandatory": false,
            "Severity": "HIGH",
            "ScheduledBy": "CUSTOMER",
            "ScheduledTime": 1.673871601E9,
            "Status": "PENDING_UPDATE",
            "Type": "SERVICE_SOFTWARE_UPDATE",
        },
        {
            "Cancellable": true,
            "Description": "Amazon Opensearch will adjust the young generation JVM
 arguments on your domain to improve performance",
            "ID": "Auto-Tune",
            "Mandatory": true,
            "Severity": "MEDIUM",
            "ScheduledBy": "SYSTEM",
            "ScheduledTime": 1.673871601E9,
            "Status": "PENDING_UPDATE",
            "Type": "JVM_HEAP_SIZE_TUNING",
```

```
        }
    ]
}
```

Riprogrammazione delle azioni

OpenSearch Il servizio notifica all'utente gli aggiornamenti programmati del software di servizio e le ottimizzazioni di Auto-Tune. Puoi scegliere di applicare la modifica immediatamente o riprogrammarla per una data e un'ora successive.

> **Note**
>
> OpenSearch Il servizio può pianificare l'azione entro un'ora dall'orario selezionato. Ad esempio, se scegli di applicare un aggiornamento alle 17:00, puoi applicarlo tra le 17:00 e le 18:00.

Console

Per riprogrammare un'azione

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Seleziona il nome del dominio per aprirne la configurazione.
3. Vai alla scheda della finestra Off-peak.
4. In Azioni pianificate, seleziona l'azione e scegli Riprogramma.
5. Scegli una delle seguenti opzioni:

   - Applica l'aggiornamento ora: pianifica immediatamente l'azione in modo che venga eseguita nell'ora corrente, se c'è capacità disponibile. Se la capacità non è disponibile, forniamo altre fasce orarie disponibili tra cui scegliere.
   - Pianificalo in una finestra non di punta: contrassegna l'azione da riprendere durante una prossima finestra non di punta. Non c'è alcuna garanzia che la modifica venga implementata nella finestra immediatamente successiva. A seconda della capacità, potrebbe verificarsi nei giorni successivi.
   - Riprogramma questo aggiornamento: consente di specificare una data e un'ora personalizzate per applicare la modifica. Se l'ora specificata non è disponibile per motivi di capacità, è possibile selezionare una fascia oraria diversa.

- Annulla aggiornamento pianificato: annulla l'aggiornamento. Questa opzione è disponibile solo per gli aggiornamenti opzionali del software di servizio. Non è disponibile per le azioni Auto-Tune o gli aggiornamenti software obbligatori.

6. Scegli Save changes (Salva modifiche).

CLI

Per riprogrammare un'azione utilizzando il AWS CLI, invia una richiesta. UpdateScheduledAction Per recuperare l'ID dell'azione, invia una richiesta. `ListScheduledActions`

La richiesta seguente riprogramma un aggiornamento del software di servizio per una data e un'ora specifiche:

```
aws opensearch update-scheduled-action \
  --domain-name my-domain \
  --action-id R20220721-P13 \
  --action-type "SERVICE_SOFTWARE_UPDATE" \
  --desired-start-time 1677348395000 \
  --schedule-at TIMESTAMP
```

Risposta:

```
{
   "ScheduledAction": {
      "Cancellable": true,
      "Description": "Cluster status is updated.",
      "Id": "R20220721-P13",
      "Mandatory": false,
      "ScheduledBy": "CUSTOMER",
      "ScheduledTime": 1677348395000,
      "Severity": "HIGH",
      "Status": "PENDING_UPDATE",
      "Type": "SERVICE_SOFTWARE_UPDATE"
   }
}
```

Se la richiesta ha esito negativo con un`SlotNotAvailableException`, significa che l'ora specificata non è disponibile per motivi di capacità ed è necessario specificare un'ora diversa. OpenSearch Il servizio fornisce suggerimenti alternativi sugli slot disponibili nella risposta.

Migrazione dalle finestre di manutenzione di Auto-Tune

Se un dominio è stato creato prima del 16 febbraio 2023, potrebbe utilizzare le finestre di manutenzione per pianificare le ottimizzazioni di Auto-Tune che richiedono una distribuzione. blue/green Puoi migrare i domini Auto-Tune esistenti per utilizzare invece la finestra non di punta.

> **Note**
>
> Non puoi tornare a utilizzare le finestre di manutenzione dopo aver migrato il dominio per utilizzare le finestre non di punta.

Console

Per migrare un dominio per utilizzare la finestra non di punta

1. Nella console di Amazon OpenSearch Service, seleziona il nome del dominio per aprirne la configurazione.
2. Vai alla scheda Auto-Tune e scegli Modifica.
3. Seleziona Migra alla finestra non di punta.
4. Per Ora di inizio (UTC), fornisci un'ora di inizio giornaliera per la finestra non di punta dell'Universal Coordinated Time (UTC).
5. Scegli Save changes (Salva modifiche).

CLI

Per migrare da una finestra di manutenzione di Auto-Tune a una finestra non di punta utilizzando il, invia una richiesta: AWS CLIUpdateDomainConfig

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --auto-tune-options
 DesiredState=ENABLED,UseOffPeakWindow=true,MaintenanceSchedules=[]
```

La finestra di manutenzione di Auto-Tune deve essere attivata per poter migrare un dominio dalla finestra di manutenzione di Auto-Tune a quella non di punta. È possibile abilitare la finestra non di punta in una richiesta separata o nella stessa richiesta. Per istruzioni, consulta the section called “Attivazione della finestra non di punta”.

Notifiche in Amazon OpenSearch Service

Le notifiche in Amazon OpenSearch Service contengono informazioni importanti sulle prestazioni e sullo stato dei tuoi domini. OpenSearch Il servizio ti avvisa sugli aggiornamenti del software di servizio, sui miglioramenti di Auto-Tune, sugli eventi relativi allo stato di salute del cluster e sugli errori del dominio. Le notifiche sono disponibili per tutte le versioni di Elasticsearch OSS. OpenSearch

È possibile visualizzare le notifiche nel pannello Notifiche della console di OpenSearch servizio. Tutte le notifiche per OpenSearch il Servizio vengono visualizzate anche in Amazon EventBridge. Per un elenco completo delle notifiche e degli eventi di esempio, consulta la sezione the section called "Monitoraggio degli eventi".

Nozioni di base sulle notifiche

Le notifiche vengono attivate automaticamente quando si crea un dominio. Vai al pannello Notifiche della console di OpenSearch servizio per monitorare e confermare le notifiche. Ogni notifica include informazioni quali l'ora in cui è stata pubblicata, il dominio a cui si riferisce, il livello di gravità e di stato e una breve spiegazione. È possibile visualizzare le notifiche cronologiche fino a 90 giorni prima nella console.

Dopo aver effettuato l'accesso al pannello Notifiche o confermando una notifica, è possibile che venga visualizzato un messaggio di errore che indica che non si dispone di autorizzazioni per l'esecuzione di `es:ListNotificationsV2` o `es:UpdateNotificationStatus`. Per risolvere questo problema, assegna all'utente o al ruolo le seguenti autorizzazioni in IAM:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "es:DescribeDomain",
      "es:ListDomainNames"
    ],
    "Resource": "arn:aws:es:*:111122223333:domain/*"
  }]
}
```

La console IAM genera un errore ("IAM non riconosce una o più azioni") che può essere ignorato in modo sicuro. È inoltre possibile limitare l'operazione `es:UpdateNotificationStatus` per determinati domini. Per ulteriori informazioni, consulta the section called “Riferimenti agli elementi della policy”.

Gravità delle notifiche

Le notifiche nel OpenSearch Servizio possono essere informative, che si riferiscono a qualsiasi azione già intrapresa o alle operazioni del dominio, oppure utilizzabili, che richiedono l'adozione di azioni specifiche, come l'applicazione di una patch di sicurezza obbligatoria. A ogni notifica è associata una gravità, che può essere `Informational`, `Low`, `Medium`, `High` o `Critical`. Nella seguente tabella sono riepilogate le varie gravità:

Gravità	Description	Esempi
`Informational`	Informazioni relative al funzionamento del tuo dominio.	• Aggiornamento del software di servizio disponibile • Regolazione automatica avviata
`Low`	Un'azione consigliata, ma senza alcun impatto negativo sulla disponibilità o sulle prestazioni del dominio se non viene eseguita.	• Regolazione automatica annullata • Avviso di conteggio di partizioni elevate
`Medium`	Un impatto è possibile se l'azione consigliata non viene intrapresa, ma viene fornita con una finestra temporale estesa per l'azione da intraprendere.	• Aggiornamento del software di servizio non riuscito • Superamento del limite del numero di partizioni
`High`	È necessaria un'azione urgente per evitare effetti avversi.	• Aggiornamento del software di servizio richiesto • Chiave KMS inaccessibile

Gravità	Description	Esempi
`Critical`	È necessaria un'azione immediata per evitare impatti negativi o per riprendersi da questi impatti.	Nessuna attualmente disponibile

Evento di esempio EventBridge

L'esempio seguente mostra un evento OpenSearch di notifica del servizio inviato ad Amazon EventBridge. La notifica corrispondente ha una gravità pari a `Informational` perché l'aggiornamento è facoltativo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Available",
    "severity": "Informational",
    "description": "Service software update [R20200330-p1] available."
  }
}
```

Configurazione di un dominio Multi-AZ in Amazon Service OpenSearch

Per prevenire la perdita di dati e ridurre al minimo i tempi di inattività del cluster Amazon OpenSearch Service in caso di interruzione del servizio, puoi distribuire i nodi su due o tre zone di disponibilità nella stessa regione, una configurazione nota come Multi-AZ. Le zone di disponibilità sono località isolate all'interno di ciascuna regione. AWS

Per i domini che eseguono carichi di lavoro di produzione, consigliamo l'opzione di implementazione Multi-AZ with Standby, che crea la seguente configurazione:

- Il dominio è distribuito su tre zone.
- Tipi di istanze di ultima generazione per nodi master e nodi dati dedicati.
- Tre nodi master dedicati e tre (o più di tre) nodi dati.
- Almeno due repliche per ogni indice del dominio o un multiplo di tre copie di dati (inclusi sia i nodi primari che le repliche).

Il resto di questa sezione fornisce spiegazioni e contesto su queste configurazioni.

Multi-AZ con Standby

Multi-AZ with Standby è un'opzione di implementazione per i domini Amazon OpenSearch Service che offre una disponibilità del 99,99%, prestazioni costanti per i carichi di lavoro di produzione e configurazione e gestione semplificate dei domini. Quando utilizzi Multi-AZ con Standby, i domini sono resilienti ai guasti dell'infrastruttura, senza alcun impatto sulle prestazioni o sulla disponibilità. Questa opzione di implementazione raggiunge questo standard imponendo una serie di best practice, come il numero di nodi di dati specificato, il conteggio dei nodi principali, il tipo di istanze, il numero di repliche, le impostazioni di aggiornamento del software e l'attivazione di Auto-Tune.

> **Note**
>
> Multi-AZ con Standby è disponibile per la OpenSearch versione 1.3 e successive e richiede regioni con almeno tre zone di disponibilità.

Quando si utilizza Multi-AZ con Standby, OpenSearch Service crea un dominio in tre zone di disponibilità, ciascuna contenente una copia completa dei dati e con i dati distribuiti equamente in ciascuna delle zone. Il tuo dominio riserva i nodi in una di queste zone come standby, il che significa che non soddisfano le richieste di ricerca. Quando OpenSearch Service rileva un guasto nell'infrastruttura sottostante, attiva automaticamente i nodi di standby in meno di un minuto. Il dominio continua a servire le richieste di indicizzazione e ricerca e qualsiasi impatto è limitato al tempo necessario per eseguire il failover. Non vi è alcuna ridistribuzione di dati o risorse, il che si traduce in prestazioni del cluster inalterate e nessun rischio di riduzione della disponibilità. Multi-AZ con Standby è disponibile senza costi aggiuntivi.

Sono disponibili due opzioni per creare un dominio con standby su. Console di gestione AWS Innanzitutto, puoi creare un dominio con il metodo di creazione Easy create e OpenSearch Service utilizzerà automaticamente una configurazione predeterminata, che include quanto segue:

- Tre zone di disponibilità, una delle quali funge da standby
- Tre nodi master e nodi dati dedicati
- Auto-Tune abilitato sul dominio
- GP3 archiviazione per i nodi di dati

Puoi anche scegliere il metodo di creazione Standard e selezionare Dominio con standby come opzione di distribuzione. Ciò consente di personalizzare il dominio pur mantenendo obbligatorie le funzionalità chiave di standby, come tre zone e tre nodi master. Ti consigliamo di scegliere un numero di nodi dati che sia un multiplo di tre (il numero di zone di disponibilità).

Dopo aver creato il dominio, puoi accedere alle pagine dei dettagli del dominio e, nella scheda Configurazione del cluster, confermare che 3-AZ con standby sia visualizzato sotto Zone di disponibilità.

In caso di problemi durante la migrazione di un dominio esistente a Multi-AZ con Standby, consulta Errore durante la migrazione a Multi-AZ con Standby nella guida alla risoluzione dei problemi.

Limitazioni

Quando configuri un dominio con Multi-AZ with Standby, considera le seguenti limitazioni:

- Il numero totale di shard su un nodo non può superare 1000, il numero totale di shard su un cluster non può superare 75000 e la dimensione di un singolo shard non può superare i 65 GB.
- Multi-AZ con Standby funziona solo con i tipi di `c5` istanza `m5r5`,`r6g`,,`c6g`,`m6g`,`r6gd`, `r7gc7g`, `m7g` e. `r7gd` `i3` Per ulteriori informazioni sulle istanze supportate, consulta Tipi di istanze supportati.
- È possibile utilizzare solo Provisioned IOPs SSD, General Purpose SSD (GP3) o storage con backup di istanze con modalità standby.
- Se si abilita UltraWarmsu un dominio Multi-AZ con Standby, il numero di nodi caldi deve essere un multiplo del numero di zone di disponibilità utilizzate.

Multi-AZ senza Standby

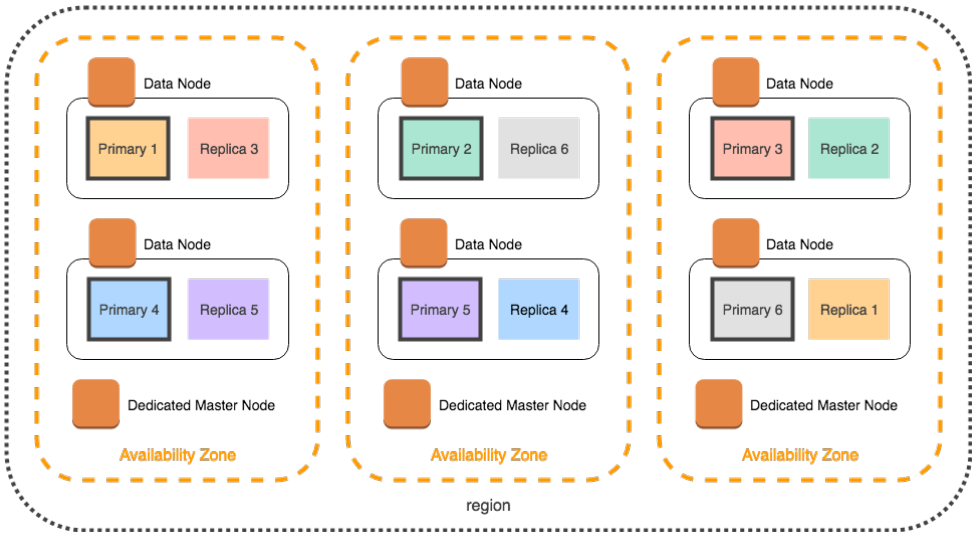
OpenSearch Il servizio supporta ancora Multi-AZ senza Standby, che offre una disponibilità del 99,9%. I nodi sono distribuiti tra le zone di disponibilità e la disponibilità dipende dal numero di zone di disponibilità e dalle copie dei dati. Mentre con lo standby è necessario configurare il dominio secondo le migliori pratiche, senza standby è possibile scegliere il numero di zone di disponibilità, nodi e repliche. Non consigliamo questa opzione a meno che non disponiate di flussi di lavoro esistenti che verrebbero interrotti dalla creazione di domini in modalità standby.

Se scegli questa opzione, ti consigliamo comunque di selezionare tre zone di disponibilità per rimanere resiliente ai guasti di nodi, dischi e Single-AZ. Quando si verifica un errore, il cluster ridistribuisce i dati tra le risorse rimanenti per mantenere la disponibilità e la ridondanza. Questo spostamento dei dati aumenta l'utilizzo delle risorse nel cluster e può avere un impatto sulle prestazioni. Se il cluster non è dimensionato correttamente, può subire una riduzione della disponibilità, il che vanifica in gran parte lo scopo di Multi-AZ.

L'unico modo per configurare un dominio senza standby su Console di gestione AWS è scegliere il metodo di creazione Standard create e selezionare Domain without standby come opzione di distribuzione.

Distribuzione di partizioni

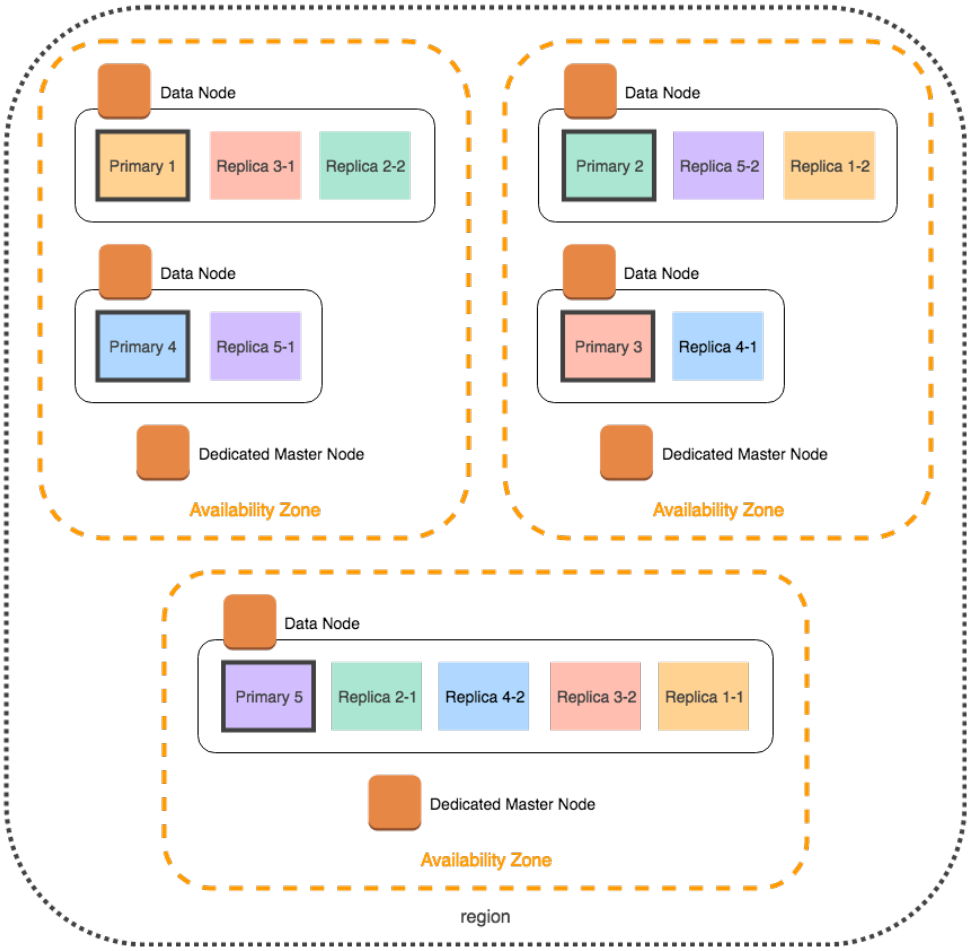
Se abiliti Multi-AZ without Standby, devi creare almeno una replica per ogni indice del cluster. Senza repliche, OpenSearch Service non può distribuire copie dei dati in altre zone di disponibilità. Fortunatamente, la configurazione predefinita per qualsiasi indice prevede un conteggio di repliche pari a 1. Come illustrato nel diagramma seguente, OpenSearch Service fa del suo meglio per distribuire gli shard primari e i corrispondenti shard di replica in zone diverse.



Oltre a distribuire gli shard per zona di disponibilità, OpenSearch Service li distribuisce per nodo. Tuttavia, alcune configurazioni di dominio possono portare a conteggi di partizioni squilibrati. Considerare il dominio seguente:

- 5 nodi di dati
- 5 partizioni primarie
- 2 repliche
- 3 zone di disponibilità

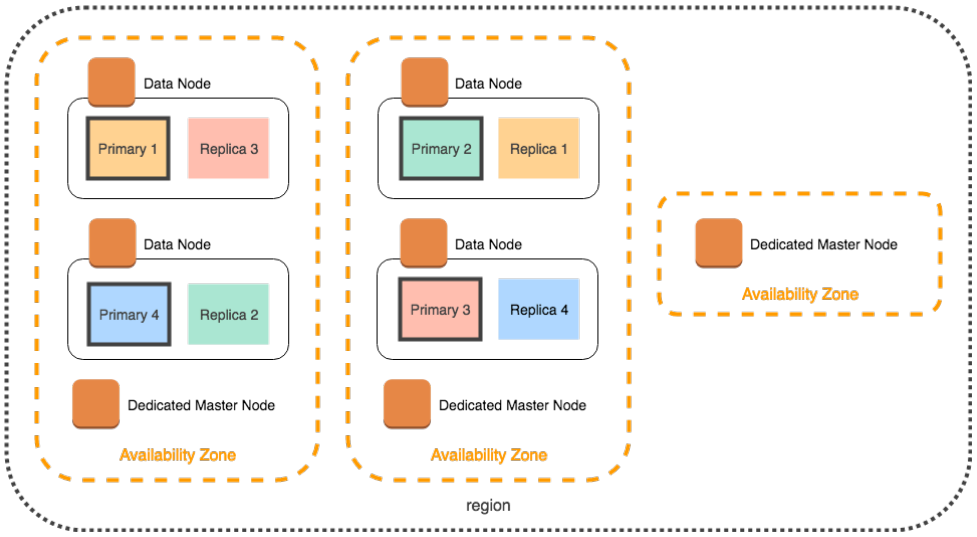
In questa situazione, OpenSearch Service deve sovraccaricare un nodo per distribuire gli shard primari e di replica tra le zone, come illustrato nel diagramma seguente.



Per evitare questo tipo di situazioni, che possono mettere a dura prova i singoli nodi e compromettere le prestazioni, si consiglia di scegliere Multi-AZ con Standby oppure di scegliere un numero di istanze multiplo di tre se si prevede di avere due o più repliche per indice.

Distribuzione dei nodi principali dedicati

Anche se si selezionano due zone di disponibilità durante la configurazione del dominio, OpenSearch Service distribuisce automaticamente i nodi master dedicati su tre zone di disponibilità. Questa distribuzione consente di evitare tempi di inattività dei cluster se in una zona si verifica un'interruzione del servizio. Se si utilizzano i tre nodi master dedicati consigliati e una zona di disponibilità è inutilizzabile, il cluster ha ancora un quorum (2) di nodi master dedicati e può eleggere un nuovo nodo master. Il seguente diagramma mostra questa configurazione.



Se si sceglie un tipo di istanza della generazione precedente che non è disponibile nelle tre zone di disponibilità, si verifica lo scenario seguente:

- Se hai scelto tre zone di disponibilità per il dominio, OpenSearch Service genera un errore. Scegliere un tipo di istanza diverso e riprovare.
- Se hai scelto due zone di disponibilità per il dominio, OpenSearch Service distribuisce i nodi master dedicati su due zone.

Interruzioni delle zone di disponibilità

I tempi di inattività per le zone di disponibilità sono rari, ma si verificano. La tabella seguente elenca diverse configurazioni con Multi-AZ e comportamenti durante i tempi di inattività. L'ultima riga della tabella si riferisce a Multi-AZ con Standby, mentre tutte le altre righe hanno configurazioni che si applicano solo a Multi-AZ senza Standby.

Numero di zone di disponibilità in una regione	Numero delle zone di disponibilità scelte	Numero di nodi principali dedicati	Comportamento nel caso in cui una zona di disponibilità subisca un'interruzione
2 o più	2	0	Tempo di inattività. Il cluster perde metà dei propri nodi di dati e deve sostituirne almeno

Numero di zone di disponibilità in una regione	Numero delle zone di disponibilità scelte	Numero di nodi principali dedicati	Comportamento nel caso in cui una zona di disponibilità subisca un'interruzione
			uno nella zona di disponibilità rimanente prima di poter eleggere un nodo master.
2	2	3	50/50 possibilità di tempi di inattività. OpenSearch Il servizio distribuisce due nodi master dedicati in una zona di disponibilità e uno nell'altra: • Se la zona di disponibilità con un nodo master dedicato subisce un'interruzione, i due nodi master dedicati nella zona di disponibilità rimanente possono scegliere un nodo master. • Se la zona di disponibilità con due nodi master dedicati subisce un'interruzione, il cluster non è disponibile fino a quando la zona di disponibilità rimanente viene ripristin ata.
3 o più	2	3	Nessun tempo di inattività. OpenSearch Il servizio distribuisce automaticamente i nodi master dedicati su tre zone di disponibilità, in modo che i restanti due nodi master dedicati possano eleggere un master.
3 o più	3	0	Nessun tempo di inattività. Circa due terzi dei nodi di dati sono ancora disponibili per eleggere un nodo master.
3 o più	3	3	Nessun tempo di inattività. I restanti due nodi master dedicati possono scegliere un nodo master.

In tutte le configurazioni, indipendentemente dalla causa, i guasti dei nodi possono causare un periodo di carico maggiore sui nodi di dati rimanenti del cluster, mentre OpenSearch Service configura automaticamente i nuovi nodi per sostituire quelli ora mancanti.

Ad esempio, nel caso in cui si verificasse l'interruzione di una zona di disponibilità in una configurazione a tre zone, due terzi dei nodi dei dati devono elaborare tutte le richieste al cluster. Mentre elaborano queste richieste, i nodi rimanenti cercano di replicare le partizioni sui nuovi nodi non appena sono disponibili online, con un ulteriore impatto sulle prestazioni. Se la disponibilità è fondamentale per il carico di lavoro, considerare l'aggiunta di risorse al cluster per alleviare questa preoccupazione.

> **Note**
>
> OpenSearch Il servizio gestisce i domini Multi-AZ in modo trasparente, quindi non è possibile simulare manualmente le interruzioni delle zone di disponibilità.

Avvio dei domini Amazon OpenSearch Service all'interno di un VPC

Puoi lanciare AWS risorse, come i domini Amazon OpenSearch Service, in un cloud privato virtuale (VPC). Un VPC è una rete virtuale dedicata a te. Account AWSÈ logicamente isolato dalle altre reti virtuali nel AWS cloud. L'inserimento di un dominio di OpenSearch servizio all'interno di un VPC consente una comunicazione sicura tra il OpenSearch Servizio e altri servizi all'interno del VPC senza la necessità di un gateway Internet, un dispositivo NAT o una connessione VPN. Tutto il traffico rimane sicuro all'interno del Cloud. AWS

> **Note**
>
> Se inserisci il tuo dominio OpenSearch di servizio all'interno di un VPC, il computer deve essere in grado di connettersi al VPC. Questa connessione spesso avviene tramite una VPN, un gateway di transito, una rete gestita o un server proxy. Non è possibile accedere direttamente ai domini dall'esterno del VPC.

VPC e domini pubblici

Di seguito sono riportati alcuni dei modi in cui i domini in VPC differiscono dai domini pubblici. Ogni differenza è descritta più avanti in modo più dettagliato.

- Grazie a loro isolamento logico, i domini che si trovano all'interno di un VPC hanno un ulteriore livello di sicurezza rispetto ai domini che usano gli endpoint pubblici.
- Mentre i domini pubblici sono accessibili da qualsiasi dispositivo connesso a Internet, i domini in VPC richiedono una qualche forma di VPN o proxy.
- Rispetto ai domini pubblici, i domini VPC visualizzano meno informazioni nella console . In particolare, la scheda Integrità del cluster non include le informazioni sulle partizioni e la scheda Indici non è presente.
- Gli endpoint del dominio assumono forme diverse (`https://search-domain-name` e `https://vpc-domain-name`).
- Poiché i gruppi di sicurezza applicano già le policy di accesso basate su IP, non è possibile applicare policy d'accesso basate su IP ai domini che si trovano all'interno di un VPC.

Limitazioni

La gestione di un dominio di OpenSearch servizio all'interno di un VPC presenta le seguenti limitazioni:

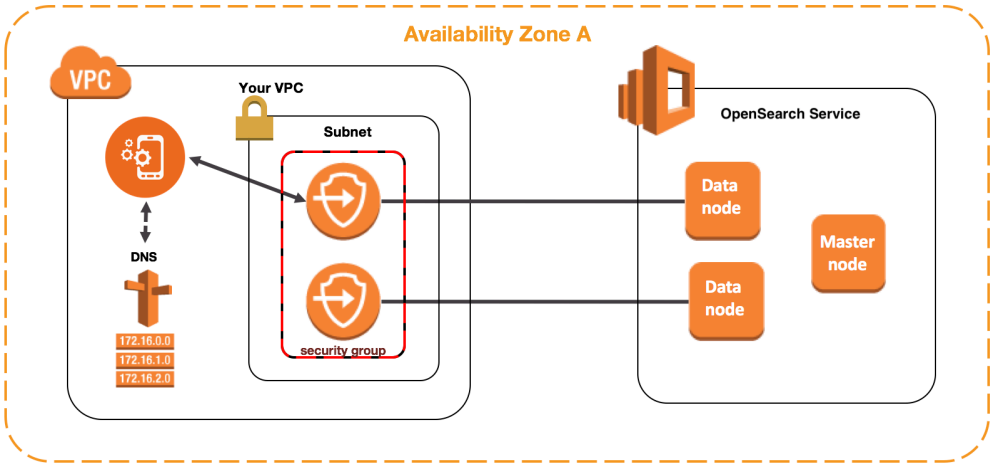
- Se avvii un nuovo dominio all'interno di un VPC, successivamente non puoi passare all'uso di un endpoint pubblico. Lo stesso vale al contrario: se crei un dominio con un endpoint pubblico, successivamente non puoi inserirlo in un VPC. Devi invece creare un nuovo dominio ed eseguire la migrazione dei dati.
- È possibile avviare il dominio all'interno di un VPC oppure usare un endpoint pubblico, ma non è possibile eseguire entrambe le operazioni. Al momento della creazione di un dominio, devi scegliere una delle due opzioni.
- Non è possibile avviare il dominio in un VPC che usa la tenancy dedicata. Devi usare un VPC con tenancy impostata su Default (Predefinita).
- Dopo aver inserito un dominio all'interno di un VPC, non è possibile spostarlo in un VPC diverso, ma è possibile modificare le sottoreti e le impostazioni del gruppo di sicurezza.
- Per accedere all'installazione predefinita di OpenSearch dashboard per un dominio che risiede all'interno di un VPC, gli utenti devono avere accesso al VPC. Questo processo varia in base alla

configurazione di rete, ma generalmente prevede la connessione a una VPN o una rete gestita o l'uso di un server proxy o di un gateway di transito. Per ulteriori informazioni, consultare the section called “Informazioni sulle policy d'accesso nei domini VPC”, la Guida per l'utente di Amazon VPC e the section called “Controllo dell'accesso ai pannelli di controllo”.

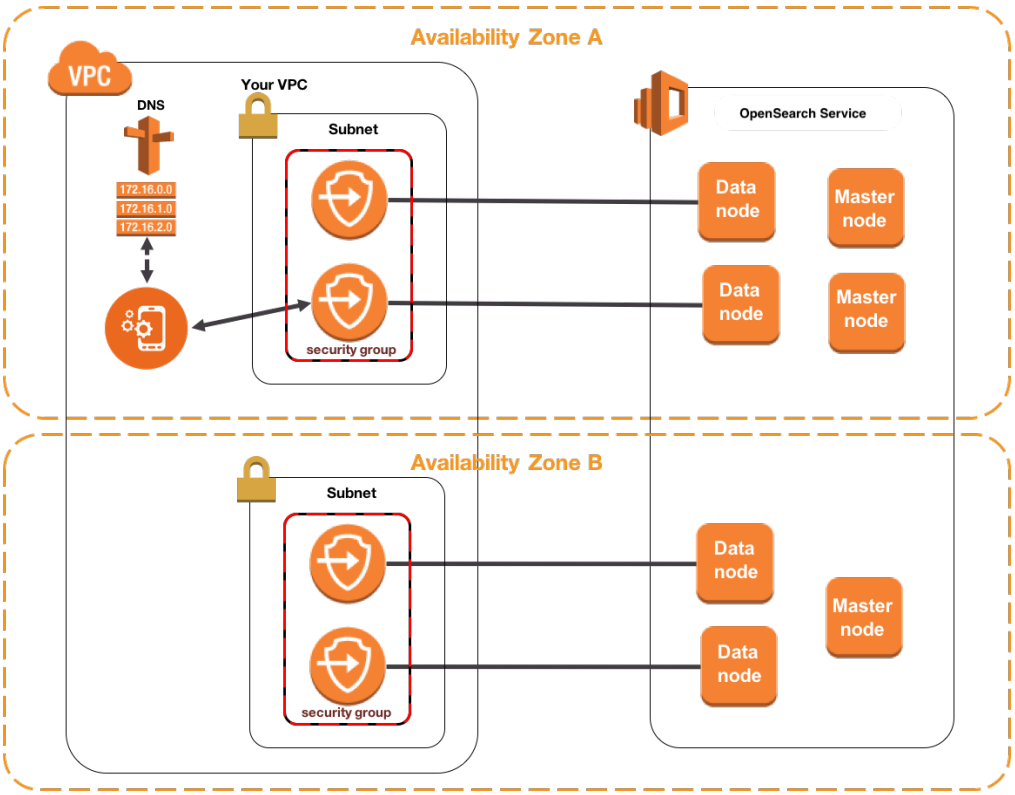
Architecture

Per supportare VPCs, OpenSearch Service colloca un endpoint in una, due o tre sottoreti del tuo VPC. Se si attivano più zone di disponibilità per il dominio, ogni sottorete deve trovarsi in una zona di disponibilità diversa nella stessa area. Se si utilizza una sola zona di disponibilità, OpenSearch Service colloca un endpoint in una sola sottorete.

La figura seguente mostra l'architettura del VPC per una zona di disponibilità.



La figura seguente mostra l'architettura del VPC per due zone di disponibilità.



OpenSearch Il servizio inserisce anche un'interfaccia di rete elastica (ENI) nel VPC per ciascuno dei nodi di dati. OpenSearch Il servizio assegna a ciascun ENI un indirizzo IP privato dall'intervallo di IPv4 indirizzi della sottorete. Il servizio, inoltre, assegna un nome host DNS pubblico (che è l'endpoint del dominio) per gli indirizzi IP. È necessario usare un servizio DNS pubblico per risolvere l'endpoint (che è un nome host DNS) negli indirizzi IP appropriati per i nodi di dati:

- Se il tuo VPC utilizza il server DNS fornito da Amazon impostando l'`enableDnsSupport`opzione su `true` (il valore predefinito), la risoluzione per l'endpoint del OpenSearch servizio avrà esito positivo.
- Se il tuo VPC utilizza un server DNS privato e il server può raggiungere i server DNS pubblici autorevoli per risolvere i nomi host DNS, anche la risoluzione per l'endpoint del servizio avrà successo. OpenSearch

Poiché gli indirizzi IP possono cambiare, è necessario risolvere l'endpoint del dominio periodicamente, in modo che sia sempre possibile accedere ai nodi di dati corretti. È consigliabile impostare l'intervallo di risoluzione DNS su un minuto. Se usi un client, assicurati anche che la cache DNS nel client venga cancellata.

Migrazione dall'accesso pubblico all'accesso VPC

Quando crei un dominio, specifichi se deve avere un endpoint pubblico o trovarsi all'interno di un VPC. Una volta completata la creazione, non è possibile passare da un'opzione all'altra. Devi invece creare un nuovo dominio e reindicizzare manualmente i dati o eseguirne la migrazione. Gli snapshot offrono una semplice modalità di migrazione dei dati. Per informazioni su come acquisire e ripristinare gli snapshot, consulta the section called “Creazione di snapshot di indici”.

Informazioni sulle policy d'accesso nei domini VPC

L'inserimento OpenSearch del dominio di servizio all'interno di un VPC fornisce un livello di sicurezza intrinseco e robusto. Quando crei un dominio con accesso pubblico, l'endpoint ha il formato seguente:

```
https://search-domain-name-identifier.region.es.amazonaws.com
```

Come suggerisce l'etichetta "pubblico", questo endpoint è accessibile da qualsiasi dispositivo connesso a Internet, anche se è possibile (e necessario) controllare l'accesso. Se accedi all'endpoint in un Web browser, potrebbe venire visualizzato un messaggio `Not Authorized`, ma la richiesta raggiunge il dominio.

Quando crei un dominio con accesso VPC, l'endpoint sembra simile a un endpoint pubblico:

```
https://vpc-domain-name-identifier.region.es.amazonaws.com
```

Se provi ad accedere all'endpoint in un Web browser, tuttavia, si potrebbe verificare il timeout della richiesta. Per eseguire richieste `GET` anche di base, il computer deve essere in grado di connettersi al VPC. Questa connessione spesso avviene tramite una VPN, un gateway di transito, una rete gestita o un server proxy. Per maggiori dettagli sulle varie forme che può avere, consultare Esempi per VPC nella Guida per l'utente di Amazon VPC. Per un esempio relativo allo sviluppo consulta the section called “Test dei domini VPC”.

Oltre a questo requisito di connettività, ti VPCs consente di gestire l'accesso al dominio tramite gruppi di sicurezza. Per molti casi d'uso, questa combinazione di caratteristiche di sicurezza è sufficiente e può essere appropriato applicare una policy d'accesso aperta al dominio.

Operare con una politica di accesso aperto non significa che chiunque su Internet possa accedere al dominio del OpenSearch Servizio. Significa piuttosto che se una richiesta raggiunge il dominio del OpenSearch Servizio e i gruppi di sicurezza associati lo consentono, il dominio accetta la richiesta.

L'unica eccezione è se si utilizza il controllo granulare degli accessi o una policy di accesso che specifichi i ruoli IAM. In questi casi, affinché il dominio accetti la richiesta, i gruppi di sicurezza devono permettere l'operazione e la richiesta deve essere firmata con credenziali valide.

> **Note**
>
> Poiché i gruppi di sicurezza applicano già criteri di accesso basati su IP, non è possibile applicare criteri di accesso basati su IP ai domini di OpenSearch servizio che risiedono all'interno di un VPC. Se usi l'accesso pubblico, le policy basate su IP sono ancora disponibili.

Prima di iniziare: Prerequisiti per l'accesso al VPC

Prima di poter abilitare una connessione tra un VPC e il nuovo dominio di OpenSearch servizio, devi fare quanto segue:

- Crea un VPC

  Per creare il tuo VPC, puoi utilizzare la console Amazon VPC, la AWS CLI o uno dei. AWS SDKs Per ulteriori informazioni, consulta Working with VPCs nella Amazon VPC User Guide. Se hai già un VPC, questa fase può essere ignorata.
- Prenota indirizzi IP

  OpenSearch Il servizio consente la connessione di un VPC a un dominio inserendo le interfacce di rete in una sottorete del VPC. Ogni interfaccia di rete è associata a un indirizzo IP. È necessario riservare un numero sufficiente di indirizzi IP nella sottorete per le interfacce di rete. Per ulteriori informazioni, consultare Prenotazione di indirizzi IP in una sottorete VPC.

Test dei domini VPC

La sicurezza avanzata di un VPC può trasformare la connessione al dominio e l'esecuzione di test di base in una vera sfida. Se disponi già di un dominio OpenSearch Service VPC e preferisci non creare un server VPN, prova la seguente procedura:

1. Per la policy di accesso del dominio, scegli Utilizza solo il controllo granulare degli accessi. È sempre possibile aggiornare questa impostazione dopo aver completato il test.
2. Crea un'istanza Amazon Linux Amazon EC2 nello stesso VPC, sottorete e gruppo di sicurezza del tuo dominio di servizio. OpenSearch

Poiché questa istanza è a scopo di test e deve fare poco lavoro, scegliere un tipo di istanza poco costoso come `t2.micro`. Assegnare all'istanza un indirizzo IP pubblico e creare una nuova coppia di chiavi o sceglierne una esistente. Se si crea una nuova chiave, scaricarla nella directory `~/.ssh`.

Per ulteriori informazioni sulla creazione di istanze, consultare Nozioni di base sulle istanze Linux di Amazon EC2.

3. Aggiungere un gateway internet al VPC.
4. Nella tabella di routing per il VPC, aggiungere un nuovo routing. Per Destination (Destinazione), specificare un blocco CIDR che contiene l'indirizzo IP pubblico del computer. In Target (Destinazione), specificare il gateway Internet appena creato.

   Ad esempio, è possibile specificare solo `123.123.123.123/32` per un computer o `123.123.123.0/24` per una serie di computer.
5. Per il gruppo di sicurezza, specificare due regole in entrata:

Tipo	Protocollo	Intervallo porte	Origine
SSH (22)	TCP (6)	22	*your-cidr-block*
HTTPS (443)	TCP (6)	443	*your-security-group-id*

   La prima regola consente l'SSH nell'istanza EC2. Il secondo consente all'istanza EC2 di comunicare con il dominio OpenSearch Service tramite HTTPS.
6. Dal terminale, esegui il comando seguente:

```
ssh -i ~/.ssh/your-key.pem ec2-user@your-ec2-instance-public-ip -N -L
 9200:vpc-domain-name-identifier.region.es.amazonaws.com:443
```

   Questo comando crea un tunnel SSH che inoltra le richieste a https://localhost:9200 al dominio del OpenSearch servizio tramite l'istanza EC2. Specificando la porta 9200 nel comando si simula un' OpenSearch installazione locale, ma usa la porta che preferisci. OpenSearch Il servizio accetta solo connessioni tramite la porta 80 (HTTP) o 443 (HTTPS).

Il comando non fornisce alcun feedback e viene eseguito a tempo indeterminato. Per arrestarlo, premere `Ctrl + C`.

7. Vai a https://localhost:9200/_dashboards/ nel tuo browser web. Potrebbe essere necessario riconoscere un'eccezione di sicurezza.

   In alternativa, è possibile inviare le richieste a https://localhost:9200 utilizzando curl, Postman o il linguaggio di programmazione preferito.

   > **Tip**
   >
   > Se si verificano errori di curl a causa di una mancata corrispondenza del certificato prova il flag `--insecure`.

Prenotazione di indirizzi IP in una sottorete VPC

OpenSearch Il servizio collega un dominio a un VPC inserendo le interfacce di rete in una sottorete del VPC (o più sottoreti del VPC se si abilitano più zone di disponibilità). Ogni interfaccia di rete è associata a un indirizzo IP. Prima di creare il dominio di OpenSearch servizio, è necessario disporre di un numero sufficiente di indirizzi IP disponibili in ciascuna sottorete per ospitare le interfacce di rete.

Ecco la formula di base: il numero di indirizzi IP che OpenSearch Service riserva in ogni sottorete è tre volte il numero di nodi di dati, diviso per il numero di zone di disponibilità.

Esempi

- Se un dominio ha nove nodi di dati su tre zone di disponibilità, il numero di indirizzi IP per sottorete è pari a 9*3/3 = 9.
- Se un dominio ha otto nodi di dati su due zone di disponibilità, il numero di indirizzi IP per sottorete è pari a 8*3/2 = 12.
- Se un dominio ha sei nodi di dati in una zona di disponibilità, il numero di indirizzi IP per sottorete è pari a 6*3/1 = 18.

Quando crei il dominio, OpenSearch Service riserva gli indirizzi IP, ne utilizza alcuni per il dominio e riserva il resto per le implementazioni blu/verdi. È possibile visualizzare le interfacce di rete e

gli indirizzi IP associati nella sezione Interfacce di rete della console Amazon EC2. La colonna Descrizione mostra a quale dominio di OpenSearch servizio è associata l'interfaccia di rete.

> **Tip**
>
> Si consiglia di creare sottoreti dedicate per gli indirizzi IP riservati del OpenSearch servizio. Usando sottoreti dedicate, puoi evitare sovrapposizioni con altri servizi e applicazioni e puoi riservare ulteriori indirizzi IP se è necessario dimensionare il cluster in futuro. Per ulteriori informazioni, consultare Creazione di una sottorete nel VPC.

Puoi anche prendere in considerazione il provisioning di nodi coordinatori dedicati per ridurre il numero di prenotazioni di indirizzi IP privati richieste per il tuo dominio VPC. OpenSearchcollega un'elastic network interface (ENI) ai nodi coordinatori dedicati anziché ai nodi dati. I nodi coordinatori dedicati rappresentano in genere circa il 10% del totale dei nodi di dati. Di conseguenza, un numero inferiore di indirizzi IP privati sarà riservato ai domini VPC.

Ruolo collegato ai servizi per l'accesso VPC

Un ruolo collegato ai servizi è un tipo specifico di ruolo IAM che delega le autorizzazioni a un servizio così che possa creare e gestire le risorse per conto dell'utente. OpenSearch Il servizio richiede un ruolo collegato al servizio per accedere al tuo VPC, creare l'endpoint del dominio e posizionare le interfacce di rete in una sottorete del tuo VPC.

OpenSearch Il servizio crea automaticamente il ruolo quando utilizzi la console di OpenSearch servizio per creare un dominio all'interno di un VPC. Affinché questa operazione di creazione automatica riesca, devi disporre delle autorizzazioni per l'operazione `iam:CreateServiceLinkedRole`. Per ulteriori informazioni, consultare Autorizzazioni del ruolo collegato ai servizi nella Guida per l'utente di IAM.

Dopo che OpenSearch Service ha creato il ruolo, puoi visualizzarlo (`AWSServiceRoleForAmazonOpenSearchService`) utilizzando la console IAM.

Per informazioni complete sulle autorizzazioni di questo ruolo e su come eliminarlo, consulta the section called “Uso di ruoli collegati ai servizi”.

Creazione di istantanee dell'indice in Amazon Service OpenSearch

Le istantanee in Amazon OpenSearch Service sono backup degli indici e dello stato di un cluster. Lo stato include le impostazioni del cluster, le informazioni sul nodo, le impostazioni degli indici e l'allocazione delle partizioni.

OpenSearch Le istantanee dei servizi sono disponibili nelle seguenti forme:

- Gli snapshot automatici sono solo per il ripristino del cluster. È possibile utilizzarli per ripristinare il dominio in caso di stato rosso del cluster o di perdita di dati. Per ulteriori informazioni, vedere Ripristino delle istantanee di seguito. OpenSearch Il servizio archivia le istantanee automatizzate in un bucket Amazon S3 preconfigurato senza costi aggiuntivi.
- Gli snapshot manuali servono per il ripristino del cluster o lo spostamento di dati da un cluster a un altro. È necessario avviare gli snapshot manuali. Questi snapshot vengono archiviati nel bucket Amazon S3 e vengono applicati i costi standard di S3. Se disponi di un'istantanea da un OpenSearch cluster autogestito, puoi utilizzarla per migrare a un dominio di servizio. OpenSearch Per ulteriori informazioni, consulta Migrazione ad Amazon OpenSearch Service.

Tutti i domini OpenSearch di servizio scattano istantanee automatiche, ma la frequenza varia nei seguenti modi:

- Per i domini che eseguono OpenSearch Elasticsearch 5.3 e versioni successive, OpenSearch Service acquisisce istantanee automatiche ogni ora e ne conserva fino a 336 per 14 giorni. Gli snapshot orari sono meno dirompenti a causa della loro natura incrementale. Forniscono inoltre un punto di ripristino più recente in caso di problemi di dominio.
- Per i domini che eseguono Elasticsearch 5.1 e versioni precedenti, OpenSearch Service acquisisce istantanee automatiche giornaliere nell'ora specificata, ne conserva fino a 14 e non conserva alcun dato sulle istantanee per più di 30 giorni.

Se il cluster entra nello stato rosso, tutti gli snapshot automatici hanno esito negativo mentre lo stato del cluster persiste. Se il problema non viene risolto entro due settimane, è possibile che i dati del cluster vengano persi definitivamente. Per la risoluzione dei problemi, consulta the section called “Cluster in stato rosso”.

Prerequisiti

Per creare manualmente gli snapshot, è necessario utilizzare IAM e Amazon S3. Verificare che siano soddisfatti i seguenti prerequisiti prima di provare ad acquisire uno snapshot.

Prerequisito	Description
Bucket S3	Crea un bucket S3 per archiviare istantanee manuali per il tuo dominio di servizio. OpenSearch Per istruzioni, consulta Creazione di un bucket generico nella Guida per l'utente di Amazon Simple Storage Service. Ricordare il nome del bucket per utilizzarlo nei seguenti punti: • L'istruzione `Resource` della policy IAM collegata al ruolo IAM • Il client Python utilizzato per registrare un repository di snapshot (se si utilizza questo metodo) **Important** Non applicare una regola del ciclo di vita di Amazon Glacier a questo bucket. Le istantanee manuali non supportano la classe di storage Amazon Glacier.

Ruolo IAM	Crea un ruolo IAM per delegare le autorizzazioni al servizio. OpenSearch Per le istruzioni, consultare Creazione di un ruolo IAM (console) nella Guida per l'utente di IAM. Nel resto di questo capitolo ci si riferisce a questo ruolo come `TheSnapshotRole` .
	
	Collegamento di una policy IAM
	
	Collegare la seguente policy a `TheSnapshotRole` per consentire l'accesso al bucket S3:
	
	`{`
	`"Version":"2012-10-17",`
	`"Statement": [{`
	`"Action": [`

Prerequisito	Description

```
        "s3:ListBucket"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3::: amzn-s3-demo-bucket  "
      ]
    },
    {
      "Action": [
        "s3:GetObject",
        "s3:PutObject",
        "s3:DeleteObject"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3::: amzn-s3-demo-bucket  /*"
      ]
    }
  ]
}
```

Per istruzioni su come allegare una policy a un ruolo, consulta Adding IAM identity permissions (console) nella IAM User Guide.

Modifica della relazione di trust

Modifica la relazione di trust `TheSnapshotRole` per specificare OpenSearch Service nell'`Principal` istruzione, come mostrato nell'esempio seguente:

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Sid": "",
    "Effect": "Allow",
    "Principal": {
      "Service": "es.amazonaws.com"
    },
    "Action": "sts:AssumeRole"
  }]
```

Prerequisito	Description
	<pre>}</pre> Per istruzioni su come modificare la relazione di trust, consulta Update a role trust policy nella IAM User Guide.

Prerequisito	Description

Permissions	Per registrare il repository di snapshot, devi essere in grado di passare `TheSnapshotRole` a OpenSearch Service. Devi inoltre disporre dell'accesso all'operazione `es:ESHttpPut` . Per concedere entrambe queste autorizzazioni, collega la policy seguente al ruolo IAM le cui credenziali vengono utilizzate per firmare la richiesta:
	
	`{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Action": "iam:PassRole", "Resource": "arn:aws:iam:: 123456789012 :role/TheSnapshotRole " }, { "Effect": "Allow", "Action": "es:ESHttpPut", "Resource": "arn:aws:es: us-east-1 :123456789012 :domain/domain-name /*" } ] }`
	
	Se l'utente o il ruolo non dispone `iam:PassRole` delle autorizzazioni necessarie per il passaggio`TheSnapshotRole` , è possibile che si verifichi il seguente errore comune quando si tenta di registrare un repository nel passaggio successivo:
	
	`$ python register-repo.py {"Message":"User: arn:aws:iam:: 123456789012 :user/MyUserAccount is not authorized to perform: iam:PassRole on resource: arn:aws:iam:: 123456789012 :role/TheSnapshotRole "}`

Registrazione di un repository di snapshot manuali

È necessario registrare un repository di istantanee con OpenSearch Service prima di poter scattare istantanee dell'indice manualmente. Questa operazione unica richiede la firma della AWS richiesta con credenziali a cui è consentito l'accesso`TheSnapshotRole`, come descritto in. the section called “Prerequisiti”

Passaggio 1: mappare il ruolo dell'istantanea nelle OpenSearch dashboard (se si utilizza un controllo di accesso granulare)

Il controllo granulare degli accessi introduce un passaggio aggiuntivo durante la registrazione di un repository. Anche se si utilizza l'autenticazione di base HTTP per tutti gli altri scopi, è necessario mappare il ruolo `manage_snapshots` al ruolo IAM che dispone delle autorizzazioni `iam:PassRole` per inviare `TheSnapshotRole`.

1. Vai al plug-in OpenSearch Dashboards per il tuo dominio di servizio. OpenSearch Puoi trovare l'endpoint Dashboards nella dashboard del tuo dominio nella OpenSearch console di servizio.
2. Dal menu principale scegliere Sicurezza, Ruoli e selezionare il ruolo manage_snapshots.
3. Scegliere Utenti mappati, Gestisci mappatura.
4. Aggiungi l'ARN del ruolo che dispone delle autorizzazioni per inviare `TheSnapshotRole`. Inserisci il ruolo nei ruoli ARNs di backend.

   ```
   arn:aws:iam::123456789123:role/role-name
   ```

5. Selezionare Mappa e confermare che l'utente o il ruolo venga visualizzato in Utenti mappati.

Fase 2: Registrazione di un repository

La seguente scheda Snapshot mostra come registrare una directory di snapshot. Per le opzioni specifiche relative alla crittografia di un'istantanea manuale e alla registrazione di un'istantanea dopo la migrazione a un nuovo dominio, consulta le schede pertinenti.

Snapshots

Per registrare un archivio di snapshot, invia una richiesta PUT all'endpoint del dominio del servizio. OpenSearch Puoi usare curl, il client Python di esempio, Postman o qualche altro metodo per inviare una richiesta firmata per registrare il repository di snapshot. Tieni presente

che non puoi utilizzare una richiesta PUT nella console OpenSearch Dashboards per registrare il repository.

La richiesta ha il seguente formato:

```
PUT domain-endpoint/_snapshot/my-snapshot-repo-name
{
  "type": "s3",
  "settings": {
    "bucket": "amzn-s3-demo-bucket",
    "base_path": "my/snapshot/directory",
    "region": "region",
    "role_arn": "arn:aws:iam::123456789012:role/TheSnapshotRole"
  }
}
```

> **Note**
>
> I nomi dei repository non possono iniziare con "cs-". Inoltre, non dovresti scrivere nello stesso repository da più domini. Solo un dominio deve avere accesso in scrittura al repository.

Se il dominio si trova all'interno di un Virtual Private Cloud (VPC), perché la richiesta registri correttamente il repository di snapshot il computer deve essere connesso al VPC. L'accesso a un VPC varia in base alla configurazione della rete, ma prevede con ogni probabilità la connessione a una VPN o a una rete aziendale. Per verificare di poter raggiungere il dominio del OpenSearch servizio, accedi a `https://your-vpc-domain.region.es.amazonaws.com` In un browser Web e verifica di ricevere la risposta JSON predefinita.

Quando il bucket Amazon S3 si trova in un OpenSearch dominio Regione AWS diverso dal tuo, aggiungi il parametro `"endpoint": "s3.amazonaws.com"` alla richiesta.

Encrypted snapshots

Al momento non puoi utilizzare le chiavi AWS Key Management Service (KMS) per crittografare le istantanee manuali, ma puoi proteggerle utilizzando la crittografia lato server (SSE).

Per attivare SSE con chiavi gestite da S3 per il bucket che usi come repository di istantanee, aggiungile al blocco della richiesta PUT. `"server_side_encryption": true` `"settings"`

Per ulteriori informazioni, consulta Uso della crittografia lato server con chiavi gestite da Amazon S3 (SSE-S3) nella Guida per l'utente di Amazon Simple Storage Service.

In alternativa, puoi utilizzare AWS KMS le chiavi per la crittografia lato server sul bucket S3 che usi come repository di istantanee. Se utilizzi questo approccio, assicurati di fornire l'`TheSnapshotRole`autorizzazione alla AWS KMS chiave utilizzata per crittografare il bucket S3. Per ulteriori informazioni, consulta Policy delle chiavi in AWS KMS.

Domain migration

La registrazione di un repository di snapshot è un'operazione che viene eseguita una tantum. Tuttavia, per eseguire la migrazione da un dominio a un altro, è necessario registrare lo stesso repository di snapshot sia sul vecchio dominio che sul nuovo. Il nome del repository è arbitrario.

Considerare le seguenti linee guida quando si esegue la migrazione a un nuovo dominio o si registra lo stesso repository con più domini:

- Quando si registra il repository nel nuovo dominio, aggiungere `"readonly": true` al blocco `"settings"` della richiesta PUT. Questa impostazione impedisce di sovrascrivere accidentalmente i dati dal vecchio dominio. Solo un dominio deve avere accesso in scrittura al repository.
- Se stai migrando i dati verso un dominio in un altro (ad esempio Regione AWS, da un vecchio dominio e bucket situato in us-east-2 a un nuovo dominio in us-west-2), `"region": "region"` sostituiscili con nell'istruzione PUT e riprova la richiesta. `"endpoint": "s3.amazonaws.com"`

Utilizzo del client Python di esempio

Il client Python è più facile da automatizzare rispetto a una semplice richiesta HTTP e ha una migliore riusabilità. Se si sceglie di utilizzare questo metodo per registrare un repository di snapshot, salvare il seguente codice Python di esempio come file Python, ad esempio `register-repo.py`. Il client richiede i pacchetti AWS SDK per Python (Boto3), richieste e requests-aws4auth. Il client contiene esempi commentati per altre operazioni con snapshot.

Aggiornare le seguenti variabili nel codice di esempio:`host`, `region`, `path` e `payload`.

```
import boto3
import requests
from requests_aws4auth import AWS4Auth
```

```
host = '' # domain endpoint
region = '' # e.g. us-west-1
service = 'es'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

# Register repository

path = '/_snapshot/my-snapshot-repo-name' # the OpenSearch API endpoint
url = host + path

payload = {
  "type": "s3",
  "settings": {
    "bucket": "amzn-s3-demo-bucket",
    "base_path": "my/snapshot/directory",
    "region": "us-west-1",
    "role_arn": "arn:aws:iam::123456789012:role/snapshot-role"
  }
}

headers = {"Content-Type": "application/json"}

r = requests.put(url, auth=awsauth, json=payload, headers=headers)

print(r.status_code)
print(r.text)

# # Take snapshot
#
# path = '/_snapshot/my-snapshot-repo-name/my-snapshot'
# url = host + path
#
# r = requests.put(url, auth=awsauth)
#
# print(r.text)
#
# # Delete index
#
# path = 'my-index'
# url = host + path
#
# r = requests.delete(url, auth=awsauth)
```

```
#
# print(r.text)
#
# # Restore snapshot (all indexes except Dashboards and fine-grained access control)
#
# path = '/_snapshot/my-snapshot-repo-name/my-snapshot/_restore'
# url = host + path
#
# payload = {
#   "indices": "-.kibana*,-.opendistro_security,-.opendistro-*",
#   "include_global_state": False
# }
#
# headers = {"Content-Type": "application/json"}
#
# r = requests.post(url, auth=awsauth, json=payload, headers=headers)
#
# print(r.text)
#
# # Restore snapshot (one index)
#
# path = '/_snapshot/my-snapshot-repo-name/my-snapshot/_restore'
# url = host + path
#
# payload = {"indices": "my-index"}
#
# headers = {"Content-Type": "application/json"}
#
# r = requests.post(url, auth=awsauth, json=payload, headers=headers)
#
# print(r.text)
```

Acquisizione di snapshot manuali

Gli snapshot non sono istantanei. Richiedono tempo per essere completati e non rappresentano point-in-time viste perfette del cluster. Mentre è in corso uno snapshot, è comunque possibile indicizzare i documenti ed effettuare altre richieste al cluster, ma i nuovi documenti (e gli aggiornamenti ai documenti esistenti) non saranno generalmente inclusi nello snapshot. L'istantanea include gli shard primari così come esistevano al momento dell' OpenSearch avvio dell'istantanea. A seconda della dimensione del pool di thread dello snapshot, potrebbero essere incluse diverse partizioni nello snapshot in momenti leggermente diversi. Per le best practice relative alle istantanee, consulta. the section called “Migliora le prestazioni delle istantanee”

> **Warning**
>
> Le istantanee manuali non includono i dati archiviati su livelli di cold storage UltraWarm o su livelli di cold storage. Se il tuo dominio utilizza una UltraWarm conservazione a freddo, esegui la migrazione di tali indici alla memorizzazione a caldo prima di scattare un'istantanea manuale se hai bisogno di conservare quei dati.

Archiviazione e prestazioni degli snapshot

OpenSearch le istantanee sono incrementali, il che significa che archiviano solo i dati modificati dall'ultima istantanea riuscita. Questa natura incrementale significa che la differenza di utilizzo del disco tra snapshot frequenti e infrequenti spesso è minima. In altre parole, effettuando snapshot orarie per una settimana (per un totale di 168 snapshot) potrebbe non essere necessario molto più spazio su disco rispetto a una snapshot singola alla fine della settimana. Inoltre, se la frequenza con cui si prendono le snapshot è alta, minore è il tempo necessario per il completamento del processo. Ad esempio, gli snapshot giornalieri possono richiedere 20-30 minuti per essere completati, mentre gli snapshot orari potrebbero essere completati in pochi minuti. Alcuni OpenSearch utenti scattano istantanee ogni mezz'ora.

Acquisisci uno snapshot

Quando si crea uno snapshot, specificare quanto segue:

- Il nome del repository di snapshot
- Un nome per lo snapshot

Per brevità e comodità, gli esempi illustrati in questo capitolo utilizzano curl, un comune client HTTP. Per passare un nome utente e una password alla tua richiesta curl, consulta il tutorial Getting started.

Se le tue politiche di accesso specificano utenti o ruoli, devi firmare le tue richieste di snapshot. Per curl, puoi usare l'`--aws-sigv4`opzione con la versione 7.75.0 o successiva. Puoi anche usare gli esempi commentati nel client Python di esempio per effettuare richieste HTTP firmate agli stessi endpoint utilizzati dai comandi curl.

Per acquisire uno snapshot manuale, procedere nel seguente modo:

1. Non puoi acquisire uno snapshot se ne è attualmente in esecuzione un altro. Per verificare, esegui il comando seguente:

```
curl -XGET 'domain-endpoint/_snapshot/_status'
```

2. Emettere il comando seguente per acquisire manualmente uno snapshot:

```
curl -XPUT 'domain-endpoint/_snapshot/repository-name/snapshot-name'
```

Per includere o escludere determinati indici e specificare altre impostazioni, aggiungere un corpo della richiesta. Per la struttura della richiesta, consulta Take snapshots nella documentazione. OpenSearch

> **Note**
>
> Il tempo necessario per scattare un'istantanea aumenta con la dimensione del dominio del OpenSearch servizio. In caso di operazioni di snapshot di lunga durata, talvolta si verifica l'errore `504 GATEWAY_TIMEOUT`. In genere, è possibile ignorare questi errori e attendere il completamento dell'operazione. Utilizzare il comando seguente per verificare lo stato di tutti gli snapshot del dominio:
>
> ```
> curl -XGET 'domain-endpoint/_snapshot/repository-name/_all?pretty'
> ```

Ripristino di snapshot

Prima di ripristinare un'istantanea, assicurati che il dominio di destinazione non utilizzi Multi-AZ con Standby. Se lo standby è abilitato, l'operazione di ripristino non riesce.

> **Warning**
>
> Se si utilizzano alias di indice, è necessario interrompere le richieste di scrittura su un alias o passare l'alias a un altro indice prima di eliminarne l'indice. Interrompere le richieste di scrittura consente di evitare il seguente scenario:
>
> 1. L'eliminazione di un indice comporta l'eliminazione anche del relativo alias.
> 2. Una richiesta di scrittura all'alias ormai eliminato crea un nuovo indice con lo stesso nome dell'alias.

3. Non puoi più utilizzare l'alias a causa di un conflitto di denominazione con il nuovo indice. Se hai passato l'alias a un altro indice, specifica `"include_aliases": false` durante il ripristino da una snapshot.

Per ripristinare una snapshot

1. Identificare lo snapshot che si desidera ripristinare. Assicurati che tutte le impostazioni di questo indice, come i pacchetti di analisi personalizzati o le impostazioni dei requisiti di allocazione, siano compatibili con il dominio. Per vedere tutti i repository di snapshot, esegui il comando seguente:

   ```
   curl -XGET 'domain-endpoint/_snapshot?pretty'
   ```

   Dopo aver identificato il repository, esegui il comando seguente per visualizzare tutte le snapshot:

   ```
   curl -XGET 'domain-endpoint/_snapshot/repository-name/_all?pretty'
   ```

   > **Note**
   >
   > La maggior parte delle snapshot automatiche viene archiviata nel repository `cs-automated`. Se il dominio prevede la crittografia dei dati a riposo, gli snapshot saranno archiviati nel repository `cs-automated-enc`. Se il repository di snapshot manuali che si sta cercando non viene trovato, verificare di averlo registrato nel dominio.

2. (Facoltativo) Eliminare o rinominare uno o più indici nel dominio del OpenSearch servizio in caso di conflitti di denominazione tra gli indici del cluster e gli indici nell'istantanea. Non è possibile ripristinare un'istantanea degli indici in un cluster che contiene già indici con gli stessi nomi. OpenSearch

   Se sono presenti conflitti di nomi degli indici, sono disponibili le opzioni seguenti:

   - Elimina gli indici nel dominio di OpenSearch servizio esistente e quindi ripristina l'istantanea.
   - Rinominare gli indici mentre vengono ripristinati dallo snapshot e in seguito reindicizzarli. Per informazioni su come rinominare gli indici, consulta questa richiesta di esempio nella documentazione. OpenSearch

- Ripristina l'istantanea in un dominio di OpenSearch servizio diverso (possibile solo con istantanee manuali).

Il comando seguente elimina tutti gli indici esistenti in un dominio:

```
curl -XDELETE 'domain-endpoint/_all'
```

Tuttavia, se non si prevede di ripristinare tutti gli indici, è possibile eliminarne uno:

```
curl -XDELETE 'domain-endpoint/index-name'
```

3. Esegui il comando seguente per ripristinare una snapshot:

```
curl -XPOST 'domain-endpoint/_snapshot/repository-name/snapshot-name/_restore'
```

A causa delle autorizzazioni speciali sui OpenSearch dashboard e degli indici di controllo degli accessi dettagliati, i tentativi di ripristinare tutti gli indici potrebbero fallire, soprattutto se si tenta di eseguire il ripristino da un'istantanea automatica. Nell'esempio seguente viene ripristinato solo un indice, `my-index`, da `2020-snapshot` nel repository di snapshot `cs-automated`:

```
curl -XPOST 'domain-endpoint/_snapshot/cs-automated/2020-snapshot/_restore' \
-d '{"indices": "my-index"}' \
-H 'Content-Type: application/json'
```

In alternativa, è possibile ripristinare tutti gli indici tranne gli indici Dashboards e quelli con controllo granulare degli accessi:

```
curl -XPOST 'domain-endpoint/_snapshot/cs-automated/2020-snapshot/_restore' \
-d '{"indices": "-.kibana*,-.opendistro*"}' \
-H 'Content-Type: application/json'
```

> **Note**
>
> A seconda della OpenSearch versione, potrebbe essere necessario escludere anche indici di sistema aggiuntivi, come e. `-.opensearch-observability*` `-.plugins-ml-config*` L'operazione di ripristino potrebbe non riuscire se questi indici esistono già nel dominio di destinazione. Per escluderli, aggiungili all'elenco di `indices` esclusione.

Ad esempio: `"indices": "-.kibana*,-.opendistro*,-.opensearch-observability*,-.plugins-ml-config*"`.

È possibile ripristinare un'istantanea senza eliminarne i dati utilizzando i `rename_pattern` parametri and. `rename_replacement` Per ulteriori informazioni su questi parametri, consulta i campi di richiesta dell'API Restore Snapshot e la richiesta di esempio nella documentazione. OpenSearch

> **Note**
>
> Se non tutte le partizioni principali sono disponibili per le istanze in questione, uno snapshot può avere `state` come `PARTIAL`. Tale valore indica che i dati provenienti da almeno una partizione non sono stati memorizzati. È comunque possibile eseguire il ripristino da una snapshot parziale, ma potrebbe essere necessario utilizzare le snapshot meno recenti per ripristinare gli indici mancanti.

Eliminazione degli snapshot manuali

Per eliminare uno snapshot manuale, emettere il seguente comando:

```
DELETE _snapshot/repository-name/snapshot-name
```

Automatizzazione delle istantanee con Snapshot Management

È possibile impostare una politica di gestione delle istantanee (SM) nelle OpenSearch dashboard per automatizzare la creazione e l'eliminazione periodiche delle istantanee. SM può creare un'istantanea di un gruppo di indici, mentre Index State Management può scattare una sola istantanea per indice. Per utilizzare SM in OpenSearch Service, devi registrare il tuo repository Amazon S3. Per istruzioni sulla registrazione del repository, consulta Registrazione di un repository di snapshot manuale.

> **Note**
>
> La gestione delle istantanee è supportata a partire dalla versione 2.5. OpenSearch I domini che eseguono versioni precedenti non supportano questa funzionalità.

Prima di SM, OpenSearch Service offriva una funzionalità di snapshot automatica e gratuita che è ancora attiva per impostazione predefinita. Questa funzionalità invia istantanee nel repository gestito dal servizio`cs-*`. Per disattivare la funzionalità, contatta. Supporto

Per ulteriori informazioni sulla funzionalità SM, consulta la sezione Gestione delle istantanee nella OpenSearch documentazione.

Attualmente SM non supporta la creazione di istantanee su più tipi di indici. Ad esempio, se provi a creare un'istantanea su più indici con * e alcuni indici si trovano nel livello «warm», la creazione dell'istantanea avrà esito negativo. Se è necessario che l'istantanea contenga più tipi di indice, utilizzate l'azione di istantanea ISM finché SM non supporta questa opzione.

Configurazione delle autorizzazioni

Se stai effettuando l'aggiornamento alla versione 2.5 da una versione precedente del dominio di OpenSearch servizio, le autorizzazioni di sicurezza per la gestione delle istantanee potrebbero non essere definite nel dominio. Gli utenti non amministratori devono essere mappati a questo ruolo per poter utilizzare la gestione delle istantanee sui domini utilizzando un controllo granulare degli accessi. Per creare manualmente il ruolo di gestione delle istantanee, procedi nel seguente modo:

1. In OpenSearch Dashboard, vai su Sicurezza e scegli Autorizzazioni.
2. Scegliere Crea gruppo di operazioni e configurare i seguenti gruppi:

Group name (Nome gruppo)	Permissions
`snapshot_management_full_access`	• `cluster:admin/opensearch/snapshot_management/*` • `cluster:admin/opensearch/notifications/feature/publish` • `cluster:admin/repository/*` • `cluster:admin/snapshot/*`
`snapshot_management_read_access`	• `cluster:admin/opensearch/snapshot_management/policy/get` • `cluster:admin/opensearch/snapshot_management/policy/search`

Group name (Nome gruppo)	Permissions
	• `cluster:admin/opensearch/snapshot_management/policy/explain` • `cluster:admin/repository/get` • `cluster:admin/snapshot/get`

3. Scegliere Ruoli, quindi selezionare Crea ruolo.
4. Assegna un nome al ruolo snapshot_management_role.
5. Per le autorizzazioni del cluster, seleziona o. `snapshot_management_full_access` `snapshot_management_read_access`
6. Scegli Create (Crea).
7. Dopo aver creato il ruolo, associalo a qualsiasi ruolo utente o di backend che gestirà le istantanee.

Considerazioni

Quando configuri la gestione delle istantanee, considera quanto segue:

- È consentita una sola policy per repository.
- Sono consentite fino a 400 istantanee per una policy.
- Questa funzionalità non verrà eseguita se il dominio ha uno stato rosso, è sottoposto a una pressione JVM elevata (85% o superiore) o ha una funzione di snapshot bloccata. Quando le prestazioni complessive di indicizzazione e ricerca del cluster ne risentono, anche SM potrebbe risentirne.
- Un'operazione di istantanea inizia solo al termine dell'operazione precedente, in modo che nessuna operazione di snapshot simultanea venga attivata da una policy.
- Più politiche con la stessa pianificazione possono causare un picco di risorse. Se gli indici istantanei delle policy si sovrappongono, le operazioni di snapshot a livello di shard possono essere eseguite solo in sequenza, il che può causare problemi di prestazioni a cascata. Se le policy condividono un repository, si verificherà un picco di operazioni di scrittura su quel repository.
- Si consiglia di pianificare l'automazione delle operazioni di snapshot su non più di una volta all'ora, a meno che non si tratti di un caso d'uso particolare.

Automazione di snapshot con Index State Management

È possibile utilizzare l'operazione di snapshot di Index State Management (ISM) per attivare automaticamente istantanee degli indici in base alle modifiche dell'età, delle dimensioni o del numero di documenti. ISM è la soluzione migliore quando è necessaria un'istantanea per indice. Se hai bisogno di un'istantanea di un gruppo di indici, vedi. Automatizzazione delle istantanee con Snapshot Management

Per utilizzare SM in OpenSearch Service, devi registrare il tuo repository Amazon S3. Per un esempio di policy ISM che utilizza l'operazione `snapshot`, consultare Policy di esempio.

Utilizzo di Curator per gli snapshot

Se ISM non funziona per la gestione di indici e snapshot, è possibile utilizzare Curator. Offre funzionalità di filtraggio avanzate utili che semplificano le attività di gestione in cluster complessi. Utilizza pip per installare Curator.

```
pip install elasticsearch-curator
```

È possibile utilizzare Curator come interfaccia a riga di comando (CLI) o API Python. Se utilizzi l'API Python, è necessario utilizzare la versione 7.13.4 o precedente del client elasticsearch-py legacy. Non supporta il client opensearch-py.

Se si utilizza l'interfaccia a riga di comando (CLI), esportare le credenziali dalla riga di comando e configurare `curator.yml` come segue:

```
client:
  hosts: search-my-domain.us-west-1.es.amazonaws.com
  port: 443
  use_ssl: True
  aws_region: us-west-1
  aws_sign_request: True
  ssl_no_validate: False
  timeout: 60

logging:
  loglevel: INFO
```

Aggiornamento dei domini Amazon Service OpenSearch

> **Note**
>
> OpenSearch e gli aggiornamenti di versione di Elasticsearch differiscono dagli aggiornamenti del software di servizio. Per informazioni sull'aggiornamento del software di servizio per il dominio di OpenSearch servizio, consulta. the section called “Aggiornamenti del software del servizio”

Amazon OpenSearch Service offre aggiornamenti sul posto per i domini con versione OpenSearch 1.0 o successiva oppure Elasticsearch 5.1 o versione successiva. Se utilizzi servizi come Amazon Data Firehose o Amazon CloudWatch Logs per lo streaming di dati su OpenSearch Service, verifica che questi servizi supportino la versione più recente di prima della migrazione. OpenSearch

Percorsi di aggiornamento supportati

Attualmente, OpenSearch Service supporta i seguenti percorsi di aggiornamento:

Dalla versione	Alla versione
OpenSearch 1.3, 2. x o 3. x	OpenSearch 3. x Se utilizzi OpenSearch 1.3 o 2.x, devi prima eseguire l'aggiornamento alla OpenSearch 2.19 prima di eseguire l'aggiornamento alla 3.x. OpenSearch Quando si esegue l'aggiornamento da OpenSearch 2.19 a OpenSearch 3.x, potrebbe essere necessario modificare le seguenti impostazioni di indice incompatibili. Queste erano obsolete in precedenza nella versione 2.x e, se non vengono rimosse, causeranno un errore di convalida del controllo di aggiornam ento. • index.knn.algo_param.ef_construction • index.knn.algo_param.m • index.knn.space_type • index.store.hybrid.mmap.extensions

Dalla versione	Alla versione
	Per ulteriori informazioni sulle impostazioni degli indici obsolete e sui passaggi per correggerle, consulta Impostazioni dell'indice. Per un elenco completo delle modifiche più importanti per OpenSearch 3.x, vedere 3.0.0. Se il tuo dominio contiene indici creati in OpenSearch 1.3, Elasticsearch 7.10 o versioni precedenti, devi reindicizzarli prima di eseguire l'aggiornamento a 3.x. OpenSearch OpenSearch 3.x non supporta questi indici precedenti, indipende ntemente dal fatto che siano archiviati a caldo o a freddo. UltraWarm Per reindiciz zare indici UltraWarm incompatibili o freddi: 1. Migra gli indici nella memorizzazione a caldo. 2. Reindicizza i dati. 3. Effettua la migrazione degli indici alla memorizzazione a caldo o a freddo. In alternativa, elimina gli indici se non ti servono più.

Dalla versione	Alla versione
OpenSearch 1.3 o 2. x	OpenSearch 2. x OpenSearch 2.17 abiliterà la ricerca simultanea di segmenti per impostazione predefinita con la modalità auto se il dominio soddisfa le seguenti condizioni: • Nessuna impostazione di ricerca simultanea precedente è stata impostata in modo esplicito. • Tutte le istanze di dati (calde e calde) sono del tipo di istanza 2.xl o superiore. • L'utilizzo medio della CPU p90 su istanze di dati (calde e calde) per più di 1 settimana è inferiore al 45%. Per maggiori dettagli sulle impostazioni di ricerca di segmenti simultanei qui, consulta Ricerca simultanea di segmenti. La versione 2.3 presenta le seguenti modifiche importanti: • Il `type` parametro è stato rimosso da tutti gli endpoint OpenSearch API nella versione 2.0. Per ulteriori informazioni, consulta la sezione modifiche importanti. • Se il tuo dominio contiene indici (caldi o freddi) originariamente creati in Elasticsearch 6.8, tali indici non sono compatibili con la versione 2.3. UltraWarm OpenSearch Prima di eseguire l'aggiornamento alla versione 2.3, è necessario reindicizzare gli indici incompatibili. Per gli indici incompatibili UltraWarm o freddi, esegui la migrazione alla memorizzazione a caldo, reindicizza i dati e quindi esegui nuovamente la migrazione alla memorizzazione a caldo o a freddo. In alternativa, puoi eliminare questi indici se non sono più necessari. Se accidentalmente effettui l'upgrade del tuo dominio alla versione 2.3 senza aver prima eseguito questi passaggi, non sarai in grado di migrare gli indici incompatibili dal livello di archiviazione attuale. La tua unica opzione è quella di eliminarli.

Dalla versione	Alla versione
OpenSearch 1. x	OpenSearch 1. x
Elasticsearch 7.x	Ricerca elastica 7. x o 1. OpenSearch x **Important** OpenSearch 1. x introduce numerose modifiche sostanziali. Per informazioni dettagliate, vedi Rinomina OpenSearch del servizio Amazon.
Elasticsearch 6.8	Elasticsearch 7. x o 1. OpenSearch x **Important** Elasticsearch 7.0 e OpenSearch 1.0 includono numerose modifiche sostanziali. Prima di iniziare un aggiornamento sul posto, consigliamo di scattare un'istantanea manuale del 6. dominio x, ripristinandolo su un test 7. x o OpenSearch 1. dominio x e utilizzo di quel dominio di test per identificare potenziali problemi di aggiornamento. Per informazioni sulle modifiche introdotte nella OpenSearch versione 1.0, vedereRinomina OpenSearch del servizio Amazon. Come Elasticsearch 6.x, gli indici possono contenere un solo tipo di mappatura, ma tale tipo ora deve essere denominato `_doc`. Di conseguenza, alcuni APIs non richiedono più un tipo di mappatura nel corpo della richiesta (come l'`_bulk`API). Per i nuovi indici, Elasticsearch 7, ospitato autonomamente. OpenSearch x e 1. x ha un numero di frammenti predefinito pari a uno. OpenSearch Domini di servizio su Elasticsearch 7. x e versioni successive mantengono il precedente valore predefinito di cinque.
Elasticsearch 6.x	Elasticsearch 6.x

Dalla versione	Alla versione
Elasticsearch 5.6	Elasticsearch 6.x **Important** Gli indici creati nella versione 6.x non supportano più tipi di mappatura diversi. Gli indici creati nella versione 5.x supportano ancora tipi di mappatura diversi quando sono ripristinati in un cluster 6.x. Controllare che il codice cliente crei un solo tipo di mappatura per indice. Per ridurre al minimo i tempi di inattività durante l'aggiornamento da Elasticsearch 5.6 a 6. x, OpenSearch Service reindicizza l'`.kibana`indice su`.kibana-6` , elimina`.kibana`, crea un alias denominato `.kibana` e associa il nuovo indice al nuovo alias.
Elasticsearch 5.x	Elasticsearch 5.x

Il processo di aggiornamento è costituito da tre fasi:

1. Controlli prima dell'aggiornamento: il OpenSearch servizio verifica la presenza di problemi che possono bloccare un aggiornamento e non procede alla fase successiva a meno che tali controlli non abbiano esito positivo.
2. Istantanea: il OpenSearch servizio scatta un'istantanea del cluster OpenSearch o Elasticsearch e non procede al passaggio successivo a meno che l'istantanea non abbia esito positivo. Se l'aggiornamento non riesce, OpenSearch Service utilizza questa istantanea per ripristinare il cluster allo stato originale. Per ulteriori informazioni, consulta the section called “Impossibile eseguire il downgrade dopo l'aggiornamento”.
3. Aggiornamento: il OpenSearch servizio avvia l'aggiornamento, che può richiedere da 15 minuti a diverse ore per essere completato. OpenSearch I dashboard potrebbero non essere disponibili durante alcuni o tutti gli upgrade.

Aggiornamento di un dominio (console)

Il processo di aggiornamento è irreversibile e non può essere sospeso né annullato. Durante un aggiornamento, non è possibile apportare modifiche di configurazione al dominio. Prima di avviare un aggiornamento, controllare attentamente che si desidera continuare. Puoi usare queste stesse fasi per eseguire il controllo di pre-aggiornamento senza di fatto avviare un aggiornamento.

Se il cluster dispone di nodi master dedicati, OpenSearch gli aggiornamenti vengono completati senza tempi di inattività. In caso contrario, il cluster potrebbe non rispondere per alcuni secondi dopo l'aggiornamento mentre elegge un nodo master.

Per aggiornare un dominio a una versione successiva di Elasticsearch OpenSearch

1. Acquisire uno snapshot manuale del dominio. Questa istantanea funge da backup che puoi ripristinare su un nuovo dominio se desideri tornare a utilizzare la versione precedente. OpenSearch
2. Vai a https://aws.amazon.come scegli Accedi alla console.
3. In Analytics, scegli Amazon OpenSearch Service.
4. Nel riquadro di navigazione, scegli il dominio da aggiornare in Domains (Domini).
5. Scegli Operazioni, quindi Aggiorna.
6. Scegli la versione a cui eseguire l'aggiornamento. Se stai effettuando l'aggiornamento a una OpenSearch versione, viene visualizzata l'opzione Abilita la modalità di compatibilità. Se abiliti questa impostazione, OpenSearch riporta la sua versione come 7.10 per consentire ai client OSS e ai plugin Elasticsearch OSS come Logstash di continuare a lavorare con Amazon Service. OpenSearch Puoi disattivare questa operazione in un secondo momento
7. Seleziona Upgrade (Aggiorna).
8. Controlla il campo Stato dominio nel pannello di controllo del dominio per monitorare lo stato dell'aggiornamento.

Aggiornamento di un dominio (CLI)

Puoi utilizzare le seguenti operazioni per identificare la versione corretta di OpenSearch Elasticsearch per il tuo dominio, avviare un aggiornamento sul posto, eseguire il controllo pre-aggiornamento e visualizzare lo stato di avanzamento:

- `get-compatible-versions` (`GetCompatibleVersions`)

- `upgrade-domain` (`UpgradeDomain`)
- `get-upgrade-status` (`GetUpgradeStatus`)
- `get-upgrade-history` (`GetUpgradeHistory`)

Per ulteriori informazioni, consulta il riferimento ai comandi AWS CLI e il riferimento all'API di Amazon OpenSearch Service.

Aggiornamento di un dominio (SDK)

Questo esempio utilizza il client Python di OpenSearchServicebasso livello di AWS SDK per Python (Boto) per verificare se un dominio è idoneo per l'aggiornamento a una versione specifica, lo aggiorna e controlla continuamente lo stato dell'aggiornamento.

```
import boto3
from botocore.config import Config
import time

# Build the client using the default credential configuration.
# You can use the CLI and run 'aws configure' to set access key, secret
# key, and default Region.

DOMAIN_NAME = ''  # The name of the domain to upgrade
TARGET_VERSION = ''  # The version you want to upgrade the domain to. For example,
 OpenSearch_1.1

my_config = Config(
    # Optionally lets you specify a Region other than your default.
    region_name='us-east-1'
)
client = boto3.client('opensearch', config=my_config)


def check_versions():
    """Determine whether domain is eligible for upgrade"""
    response = client.get_compatible_versions(
        DomainName=DOMAIN_NAME
    )
    compatible_versions = response['CompatibleVersions']
    for i in range(len(compatible_versions)):
        if TARGET_VERSION in compatible_versions[i]["TargetVersions"]:
            print('Domain is eligible for upgrade to ' + TARGET_VERSION)
```

```
            upgrade_domain()
            print(response)
        else:
            print('Domain not eligible for upgrade to ' + TARGET_VERSION)


def upgrade_domain():
    """Upgrades the domain"""
    response = client.upgrade_domain(
        DomainName=DOMAIN_NAME,
        TargetVersion=TARGET_VERSION
    )
    print('Upgrading domain to ' + TARGET_VERSION + '...' + response)
    time.sleep(5)
    wait_for_upgrade()


def wait_for_upgrade():
    """Get the status of the upgrade"""
    response = client.get_upgrade_status(
        DomainName=DOMAIN_NAME
    )
    if (response['UpgradeStep']) == 'UPGRADE' and (response['StepStatus']) ==
 'SUCCEEDED':
        print('Domain successfully upgraded to ' + TARGET_VERSION)
    elif (response['StepStatus']) == 'FAILED':
        print('Upgrade failed. Please try again.')
    elif (response['StepStatus']) == 'SUCCEEDED_WITH_ISSUES':
        print('Upgrade succeeded with issues')
    elif (response['StepStatus']) == 'IN_PROGRESS':
        time.sleep(30)
        wait_for_upgrade()


def main():
    check_versions()


if __name__ == "__main__":
    main()
```

Errori nella risoluzione dei problemi di convalida

Quando avvii un aggiornamento di una versione OpenSearch o di Elasticsearch, OpenSearch Service esegue innanzitutto una serie di controlli di convalida per garantire che il tuo dominio sia idoneo per un aggiornamento. Se uno di questi controlli non riesce, si riceve una notifica contenente i problemi specifici che è necessario correggere prima di aggiornare il dominio. Per un elenco dei potenziali problemi e dei passaggi per risolverli, consulta the section called “Risoluzione degli errori di convalida”.

Risoluzione dei problemi relativi all'aggiornamento

Aggiornamenti locali richiedono domini sani. Il dominio potrebbe essere non idoneo per un aggiornamento o non completare l'aggiornamento per diversi motivi. La tabella riportata di seguito mostra i problemi più comuni.

Problema	Description
Plugin opzionale non supportato	Quando si aggiorna un dominio con plug-in opzionali, OpenSearc h Service aggiorna automaticamente anche i plug-in. Pertanto, la versione di destinazione del dominio deve supportare anche questi plugin opzionali. Se nel dominio è installato un plug-in opzionale che non è disponibile per la versione di destinazione, la richiesta di aggiornamento ha esito negativo.
Troppe partizioni per nodo	OpenSearch, oltre a 7. x versioni di Elasticsearch hanno un'impost azione predefinita di non più di 1.000 shard per nodo. Se un nodo del cluster corrente supera questa impostazione, il OpenSearch servizio non consente l'aggiornamento. Consulta the section called “Limite massimo di partizioni superato” per le opzioni di risoluzione dei problemi.
Dominio in elaborazi one	Il dominio è nel mezzo di una modifica di configurazione. Controlla l'idoneità dell'aggiornamento al termine dell'operazione.
Cluster in stato rosso	Uno o più indici nel cluster sono in stato rosso. Per la risoluzione dei problemi, consulta the section called “Cluster in stato rosso”.
Elevata percentuale di errori	Il cluster restituisce un numero elevato di errori 5xx durante il tentativo di elaborazione delle richieste. Questo problema è in genere il risultato di un numero eccessivo di richieste di lettura o scrittura simultanee. Valuta

Problema	Description
	la possibilità di ridurre il traffico verso il cluster o di eseguire il dimension amento del dominio.
Split brain	Split brain significa che il cluster contiene più nodi principali ed è stato diviso in due cluster che non verranno mai riuniti in modo autonomo. Puoi evitare lo split brain utilizzando il numero consigliato di nodi master dedicati. Per risolvere un problema di split brain, contatta Supporto.
Impossibile trovare nodo master	OpenSearch Il servizio non riesce a trovare il nodo principale del cluster. Se il dominio utilizza il Multi-AZ, un errore nella zona di disponibilità potrebbe aver causato la perdita del quorum da parte del cluster e la conseguente incapacità di eleggere un nuovo nodo master. Se il problema non si risolve automaticamente, contatta Supporto.
Troppe attività in sospeso	Il nodo master si trova in condizioni di carico elevato e presenta numerose attività in sospeso. Valuta la possibilità di ridurre il traffico verso il cluster o di eseguire il dimensionamento del dominio.
Volume di archiviaz ione compromesso	Il volume del disco di uno o più nodi non funziona correttamente. Questo problema si verifica spesso insieme ad altri problemi, come un'elevat a percentuale di errori o troppe attività in sospeso. Se si verifica in isolamento e non si risolve automaticamente, contatta Supporto.
Problema chiave KMS	La chiave KMS utilizzata per crittografare il dominio non è accessibile o manca. Per ulteriori informazioni, consultare the section called “Monitora ggio dei domini che crittografano dati a riposo”.
Snapshot in corso	Il dominio sta attualmente acquisendo una snapshot. Controlla l'idoneit à dell'aggiornamento al termine della snapshot. Verifica anche che puoi elencare repository snapshot manuali, elencare snapshot all'interno di tali repository e acquisire snapshot manuali. Se OpenSearch il Servizio non è in grado di verificare se un'istantanea è in corso, gli aggiornamenti possono avere esito negativo.
Timeout o errore snapshot	Il pre-aggiornamento snapshot richiede troppo tempo o non è riuscito. Verifica l'integrità del cluster e riprova. Se il problema persiste, contatta Supporto.

Problema	Description
Indici incompatibili	Uno o più indici sono incompatibili con la versione di destinazione. Questo problema può verificarsi se hai migrato gli indici da una versione precedente di Elasticsearch o Elasticsearch. OpenSearch Reindicizzare gli indici e riprovare.
Elevato utilizzo del disco	L'utilizzo del disco per il cluster supera il 90%. Elimina i dati o ricalibra il dominio e riprova.
Elevato utilizzo JVM	L'utilizzo della memoria JVM è superiore al 75%. Riduci il traffico verso il cluster o ricalibra il dominio e riprova.
OpenSearch Problema con gli alias dei dashboard	`.dashboards` è già configurato come alias e corrisponde a un indice incompatibile, probabilmente proveniente da una versione precedente di Dashboards. OpenSearch Reindicizza e riprova.
Stato rosso di Dashboards	OpenSearch Lo stato del pannello di controllo è rosso. Provare a usare Dashboards al termine dell'aggiornamento. Se lo stato rosso persiste, risolvilo manualmente e riprova.
Compatibilità tra cluster	È possibile eseguire l'aggiornamento solo se viene mantenuta la compatibilità tra cluster tra i domini di origine e di destinazione dopo l'aggiornamento. Durante il processo di aggiornamento, vengono identific ate tutte le connessioni non compatibili. Per procedere, aggiornare il dominio remoto o eliminare le connessioni non compatibili. Tenere presente che se la replica è attiva sul dominio, non è possibile riprenderla dopo aver eliminato la connessione.
Altro problema relativo OpenSearc h al servizio di assistenza	Problemi relativi al OpenSearch Servizio stesso potrebbero far sì che il dominio venga visualizzato come non idoneo per un aggiornamento. Se nessuna delle condizioni precedenti si applica al dominio e il problema persiste per più di un giorno, contatta Supporto.

Utilizzo di uno snapshot per migrare i dati

Gli aggiornamenti sul posto sono il modo più semplice, veloce e affidabile per aggiornare un dominio a una versione successiva o di Elasticsearch. OpenSearch Gli snapshot sono un'ottima opzione

se occorre migrare da una versione precedente alla 5.1 di Elasticsearch o si desidera migrare a un cluster completamente nuovo.

La tabella seguente mostra come utilizzare le istantanee per migrare i dati verso un dominio che utilizza una versione diversa o di Elasticsearch. OpenSearch Per informazioni su come acquisire e ripristinare gli snapshot, consulta the section called “Creazione di snapshot di indici”.

Dalla versione	Alla versione	Processo di migrazione

OpenSearch 1.3 o 2. x	OpenSearch 2. x	1. Consulta le ultime modifiche apportate alla versione OpenSearch 2.3 per vedere se è necessario apportare modifiche agli indici o alle applicazioni.
		2. Crea un'istantanea manuale della versione 1.3 o 2. dominio x.
		3. Crea un 2. dominio x che è una versione superiore alla versione 1.3 o 2 originale. dominio x.
		4. Ripristina l'istantanea dal dominio originale al 2. dominio x. Durante l'operazione, potrebbe essere necessario ripristinare l'`.opensearch` indice con un nuovo nome:
		`POST _snapshot/ <repository-name> /<snapshot-name>/_restore`
		`{`
		`  "indices": "*",`
		`  "ignore_unavailable": true,`
		`  "rename_pattern": ".opensearch",`
		`  "rename_replacement": ".backup-opensearch"`
		`}`
		Quindi è possibile reindicizzare `.backup-opensearch` sul nuovo dominio e configurarlo come alias su `.opensearch` . Tieni presente che la chiamata `_restore` REST non include `include_global_state` perché l'impostazione predefinita in `_restore` è false. Di conseguenza, il dominio di test

Dalla versione	Alla versione	Processo di migrazione
		non includerà alcun modello di indice e non avrà lo stato completo del backup. 5. Se non si ha più bisogno del dominio originale, eliminarlo. In caso contrario, verranno addebitati i costi per il dominio.

OpenSearch 1. x	OpenSearch 1. x	1. Crea un'istantanea manuale del file 1. dominio x.
		2. Crea un 1. dominio x che è una versione superiore a quella originale 1. dominio x.
		3. Ripristina l'istantanea dal dominio originale al nuovo 1. dominio x. Durante l'operazione, potrebbe essere necessario ripristinare l'`.opensearch` indice con un nuovo nome:
		`POST _snapshot/ <repository-name> /<snapshot-name>/_restore`
		`{`
		`  "indices": "*",`
		`  "ignore_unavailable": true,`
		`  "rename_pattern": ".opensearch",`
		`  "rename_replacement": ".backup-opensearch"`
		`}`
		Quindi è possibile reindicizzare `.backup-opensearch` sul nuovo dominio e configurarlo come alias su `.opensearch` . Tieni presente che la chiamata `_restore` REST non include `include_global_state` perché l'impostazione predefinita in `_restore` è false. Di conseguenza, il dominio di test non includerà alcun modello di indice e non avrà lo stato completo del backup.
		4. Se non si ha più bisogno del dominio originale, eliminarlo. In caso contrario, verranno addebitati i costi per il dominio.

Dalla versione	Alla versione	Processo di migrazione

Elasticsearch 6.x o 7.x	OpenSearch 1. x	1. Consulta le ultime modifiche apportate alla OpenSearch versione 1.0 per vedere se è necessario apportare modifiche agli indici o alle applicazioni.
		2. Creare uno snapshot manuale del dominio Elasticsearch 7.x o 6.x.
		3. Crea un 1. OpenSearch dominio x.
		4. Ripristina l'istantanea dal dominio Elasticsearch al dominio. OpenSearch Durante l'operazione, potrebbe essere necessario ripristinare l'`.elasticsearch` indice con un nuovo nome:
		`POST _snapshot/ <repository-name> /<snapshot-name>/_restore`
		`{`
		`  "indices": "*",`
		`  "ignore_unavailable": true,`
		`  "rename_pattern": ".elasticsearch",`
		`  "rename_replacement": ".backup-opensearch"`
		`}`
		Quindi è possibile reindicizzare `.backup-opensearch` sul nuovo dominio e configurarlo come alias su `.elasticsearch` . Tieni presente che la chiamata `_restore` REST non include `include_global_state` perché l'impostazione predefinita in `_restore` è false. Di conseguenza, il dominio di test non includerà alcun modello di indice e non avrà lo stato completo del backup.
		5. Se non si ha più bisogno del dominio originale, eliminarlo. In caso contrario, verranno addebitati i costi per il dominio.

Dalla versione	Alla versione	Processo di migrazione

Elasticsearch 6.x	Elasticsearch 7.x	1. Esaminare le modifiche importanti in 7.0 per vedere se è necessario adeguare gli indici o le applicazioni.
		2. Creare una snapshot manuale del dominio 6.x.
		3. Creare un dominio 7.x.
		4. Ripristinare la snapshot dal dominio originale al dominio 7.x. Durante l'operazione, è probabile che sia necessario ripristinare l'indice `.opensearch` con un nuovo nome:
		`POST _snapshot/ <repository-name> /<snapshot-name>/_restore`
		`{`
		`  "indices": "*",`
		`  "ignore_unavailable": true,`
		`  "rename_pattern": ".elasticsearch",`
		`  "rename_replacement": ".backup-elasticsearch"`
		`}`
		Quindi è possibile reindicizzare `.backup-elasticsearch` sul nuovo dominio e configurarlo come alias su `.elasticsearch` . Tieni presente che la chiamata `_restore` REST non include `include_global_state` perché l'impostazione predefinita `_restore` è false. Di conseguenza, il dominio di test non includerà alcun modello di indice e non avrà lo stato completo del backup.
		5. Se non si ha più bisogno del dominio originale, eliminarlo. In caso contrario, verranno addebitati i costi per il dominio.

Dalla versione	Alla versione	Processo di migrazione
Elasticsearch 6.x	Elasticsearch 6.8	1. Creare una snapshot manuale del dominio 6.x. 2. Creare un dominio 6.8. 3. Ripristinare lo snapshot dal dominio originale nel dominio 6.8. 4. Se non si ha più bisogno del dominio originale, eliminarlo. In caso contrario, verranno addebitati i costi per il dominio.
Elasticsearch 5.x	Elasticsearch 6.x	1. Esaminare le modifiche importanti in 6.0 per vedere se è necessario adeguare gli indici o le applicazioni. 2. Creare una snapshot manuale del dominio 5.x. 3. Creare un dominio 6.x. 4. Ripristinare la snapshot dal dominio originale al dominio 6.x. 5. Se non si ha più bisogno del dominio 5.x, eliminarl o. In caso contrario, verranno addebitati i costi per il dominio.
Elasticsearch 5.x	Elasticsearch 5.6	1. Creare una snapshot manuale del dominio 5.x. 2. Creare un dominio 5.6. 3. Ripristinare la snapshot dal dominio originale al dominio 5.6. 4. Se non si ha più bisogno del dominio originale, eliminarlo. In caso contrario, verranno addebitati i costi per il dominio.

Dalla versione	Alla versione	Processo di migrazione
Elasticsearch 2.3	Elasticsearch 6.x	Gli snapshot di Elasticsearch 2.3 non sono compatibili con la versione 6.x. Per migrare i dati direttamente da 2.3 a 6.x, è necessario ricreare manualmente gli indici nel nuovo dominio. In alternativa, è possibile seguire i passaggi da 2.3 a 5.x in questa tabella, eseguire le operazioni `_reindex` nel nuovo dominio 5.x per convertire gli indici 2.3 negli indici 5.x e seguire i passaggi da 5.x a 6.x.
Elasticsearch 2.3	Elasticsearch 5.x	1. Esaminare le modifiche importanti in 5.0 per vedere se è necessario adeguare gli indici o le applicazioni. 2. Creare una snapshot manuale del dominio 2.3. 3. Creare un dominio 5.x. 4. Ripristinare la snapshot dal dominio 2.3 al dominio 5.x. 5. Se non si ha più bisogno del dominio 2.3, eliminarl o. In caso contrario, verranno addebitati i costi per il dominio.
Elasticsearch 1.5	Elasticsearch 5.x	Gli snapshot di Elasticsearch 1.5 non sono compatibi li con la versione 5.x. Per migrare i dati da 1.5 a 5.x, è necessario ricreare manualmente gli indici nel nuovo dominio. **Important** Le istantanee 1.5 sono compatibili con 2.3, ma i domini OpenSearch Service 2.3 non supportan o l'`_reindex`operazione. Poiché non è possibile reindicizzarli, gli indici originati in un dominio 1.5 non riescono a effettuare il ripristino dalle snapshot 2.3 a domini 5.x.

Dalla versione	Alla versione	Processo di migrazione
Elasticsearch 1.5	Elasticsearch 2.3	1. Utilizzare il plugin di migrazione per scoprire se è possibile effettuare direttamente l'upgrade alla versione 2.3. Potrebbe essere necessario modificare i dati prima della migrazione. a. In un browser Web, aprire `http://domain-endpoint /_plugin/migration/` . b. Scegliere Run checks now (Esegui controlli). c. Esaminare i risultati e, se necessario, seguire le istruzioni per modificare i dati. 2. Creare una snapshot manuale del dominio 1.5. 3. Creare un dominio 2.3. 4. Ripristinare la snapshot dal dominio 1.5 al dominio 2.3. 5. Se non si ha più bisogno del dominio 1.5, eliminarlo. In caso contrario, verranno addebitati i costi per il dominio.

Creazione di un endpoint personalizzato per Amazon Service OpenSearch

La creazione di un endpoint personalizzato per il tuo dominio Amazon OpenSearch Service ti consente di fare riferimento più facilmente alle tue OpenSearch dashboard OpenSearch URLs e a quelle di controllo. Puoi includere il marchio della tua azienda o semplicemente utilizzare un easier-to-remember endpoint più corto di quello standard.

Se è necessario passare a un nuovo dominio, aggiornare il DNS in modo che punti al nuovo URL e continuare a utilizzare lo stesso endpoint di prima.

Puoi proteggere gli endpoint personalizzati generando un certificato in AWS Certificate Manager (ACM) o importandone uno tuo.

Endpoint personalizzati per nuovi domini

È possibile abilitare un endpoint personalizzato per un nuovo dominio di OpenSearch servizio utilizzando la console di OpenSearch servizio o l'API di configurazione AWS CLI.

Come personalizzare l'endpoint (console)

1. Dalla console OpenSearch di servizio, scegli Crea dominio e fornisci un nome per il dominio.
2. In Endpoint personalizzato, selezionare Abilita endpoint personalizzato.
3. Per Nome host personalizzato, immettere il nome host dell'endpoint personalizzato preferito. Il nome host deve essere un nome di dominio completo (FQDN) come www.tuodominio.com o esempio.tuodominio.com.

   > **Note**
   >
   > Se non si dispone di un certificato jolly, sarà necessario ottenere un nuovo certificato per i domini secondari dell'endpoint personalizzato.

4. Per Certificato AWS , scegliere il certificato SSL che si desidera utilizzare per il dominio. Se non sono disponibili certificati, è possibile importarne uno in ACM o utilizzare ACM per eseguirne il provisioning. Per informazioni su come creare un certificato, consultare Emissione e gestione di certificati nella Guida per l'utente di AWS Certificate Manager.

   > **Note**
   >
   > Il certificato deve avere il nome dell'endpoint personalizzato e appartenere allo stesso account del dominio OpenSearch di servizio. Lo stato del certificato deve essere EMESSO.

   - Completa la procedura per creare il dominio e scegli Crea.
   - Al termine dell'elaborazione selezionare il dominio per visualizzare l'endpoint personalizzato.

   Per utilizzare la CLI o l'API di configurazione, utilizzare le operazioni `CreateDomain` e `UpdateDomainConfig`. Per ulteriori informazioni, consulta AWS CLI Command Reference e Amazon OpenSearch Service API Reference.

Endpoint personalizzati per domini esistenti

Per aggiungere un endpoint personalizzato a un dominio di OpenSearch servizio esistente, scegli Modifica ed esegui i passaggi da 2 a 4 precedenti.

Mappatura CNAME

Dopo aver abilitato un endpoint personalizzato per il tuo dominio di OpenSearch servizio, puoi creare una mappatura CNAME in Amazon Route 53 (o nel tuo provider di servizi DNS preferito). La creazione di una mappatura CNAME ti consentirà di indirizzare il traffico verso il tuo endpoint personalizzato e i relativi sottodomini. Senza questa mappatura, non sarai in grado di indirizzare il traffico verso il tuo endpoint personalizzato. Per i passaggi per creare questa mappatura in Route 53, vedi Configurazione del routing DNS per un nuovo dominio e Creazione di una nuova zona ospitata per un sottodominio. Per altri provider, consultare la relativa documentazione.

Crea un record CNAME che punti l'endpoint personalizzato all'endpoint del dominio generato automaticamente. Se il tuo dominio è dual stack, puoi indirizzare il record CNAME verso uno dei due endpoint generati dal servizio. La funzionalità dual stack dell'endpoint personalizzato dipende dall'endpoint generato dal servizio a cui punti il record CNAME. Il nome host dell'endpoint personalizzato è il nome del record CNAME e il nome host dell'endpoint del dominio è il valore del record CNAME.

Se utilizzi l'autenticazione SAML per le OpenSearch dashboard, devi aggiornare il tuo IdP con il nuovo URL SSO.

Puoi usare Amazon Route 53 per creare un tipo di record alias per indirizzare l'endpoint personalizzato del tuo dominio verso un endpoint di ricerca dual stack. Per creare un tipo di record alias, devi configurare il tuo dominio per utilizzare il tipo di indirizzo IP dual stack. Puoi farlo utilizzando l'API Route 53.

Per creare un tipo di record di alias utilizzando l'API Route 53, specifica la destinazione dell'alias del tuo dominio. Puoi trovare l'alias target del tuo dominio nel campo Hosted Zone (dual stack) nella sezione degli endpoint personalizzati della console di OpenSearch servizio oppure utilizzando l'`DescribeDomain`API e copiando il valore di. `DomainEndpointV2HostedZoneId`

Auto-Tune per Amazon Service OpenSearch

Auto-Tune in Amazon OpenSearch Service utilizza i parametri di prestazioni e utilizzo del OpenSearch cluster per suggerire modifiche alla configurazione relative alla memoria, tra cui le

dimensioni della coda e della cache e le impostazioni della macchina virtuale Java (JVM) sui nodi. Queste modifiche facoltative migliorano la velocità e la stabilità del cluster.

Alcune modifiche vengono implementate immediatamente, mentre altre sono pianificate durante la finestra non di punta del dominio. Puoi ripristinare le impostazioni predefinite del OpenSearch servizio in qualsiasi momento. Man mano che Auto-Tune raccoglie e analizza le metriche delle prestazioni per il tuo dominio, puoi visualizzarne i consigli nella console di OpenSearch servizio nella pagina Notifiche.

Auto-Tune è disponibile in commercio Regioni AWS su domini che eseguono qualsiasi OpenSearch versione o Elasticsearch 6.7 o versione successiva, con un tipo di istanza supportato.

Tipi di modifiche

La regolazione automatica ha due grandi categorie di modifiche:

- Modifiche senza interruzioni che applica durante l'esecuzione del cluster.
- Modifiche che richiedono una distribuzione blu/verde, da applicare durante la finestra non di punta del dominio.

In base ai parametri delle prestazioni del dominio, la regolazione automatica può suggerire regolazioni alle seguenti impostazioni:

Tipo di modifica	Categoria	Description
Dimensioni heap JVM	Blu/verde	Per impostazione predefinita, OpenSearch Service utilizza il 50% della RAM di un'istanza per l'heap JVM, fino a una dimensione dell'heap di 32 GiB. L'aumento di questa percentuale offre OpenSearch più memoria, ma ne lascia meno per il sistema operativo e altri processi. Valori maggiori possono ridurre il numero di interruzioni di garbage collection, ma aumentare la lunghezza di tali interruzioni.
Impostazioni JVM	Blu/verde	Le impostazioni di "nuova generazione" di JVM influenzano la frequenza delle garbage collection minori. Le raccolte minori

Tipo di modifica	Categoria	Description
di nuova generazione		più frequenti possono diminuire il numero di raccolte principali e interruzioni.
Dimensioni della coda	Senza interruzioni	Per impostazione predefinita, la dimensione della coda di ricerca è `1000` e la dimensione della coda di scrittura è`10000`. La regolazione automatica dimensiona automaticamente le code di ricerca e scrittura se è disponibile un heap aggiuntivo per gestire le richieste.
Dimensioni della cache	Senza interruzioni	La cache dei campi monitora le strutture dati on-heap, quindi è importante monitorare l'utilizzo della cache. La regolazione automatica dimensiona la dimensione della cache dei dati sul campo per evitare problemi di memoria esaurita e di interruttore. La cache di richieste delle partizioni viene gestita a livello di nodo e ha una dimensione massima di default pari all'1% dell'heap. La regolazione automatica dimensiona la dimensione della cache delle richieste delle partizioni per accettare più richieste di ricerca e indice rispetto a quelle gestite dal cluster configurato.
Dimensione richiesta	Senza interruzioni	Per impostazione predefinita, quando la dimensione aggregata delle richieste in corso supera il 10% del totale della JVM (2% per i tipi di `t2` istanze e 1% per`t3.small`), limita tutte OpenSearch le richieste nuove `_search` e `_bulk` fino al completamento delle richieste esistenti. Auto-Tune regola automaticamente questa soglia, tipicamente tra il 5-15%, in base alla quantità di JVM attualmente occupata nel sistema. Ad esempio, se la pressione della memoria JVM è elevata, Auto-Tune potrebbe ridurre la soglia al 5%, a quel punto si potrebbero vedere più rifiuti fino a quando il cluster non si stabilizza e la soglia aumenta.

Abilitazione o disabilitazione della regolazione automatica

OpenSearch Il servizio abilita Auto-Tune per impostazione predefinita su nuovi domini. Per abilitare o disabilitare Auto-Tune sui domini esistenti, consigliamo di utilizzare la console, che semplifica il processo. L'attivazione di Auto-Tune non causa una distribuzione. blue/green

Al momento non è possibile abilitare o disabilitare l'ottimizzazione automatica tramite AWS CloudFormation.

Console

Per abilitare Auto-Tune su un dominio esistente

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione, in Domini, scegli il nome di dominio per aprire la configurazione del cluster.
3. Scegli Attiva se Auto-Tune non è già abilitato.
4. Facoltativamente, seleziona Finestra non di punta per pianificare le ottimizzazioni che richiedono una blue/green distribuzione durante la finestra non di picco configurata del dominio. Per ulteriori informazioni, consulta the section called “Pianificazione dei miglioramenti di Auto-Tune”.
5. Scegli Salva modifiche.

CLI

Per abilitare Auto-Tune utilizzando il AWS CLI, invia una richiesta: UpdateDomainConfig

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --auto-tune-options DesiredState=ENABLED
```

Pianificazione dei miglioramenti di Auto-Tune

Prima del 16 febbraio 2023, Auto-Tune utilizzava le finestre di manutenzione per pianificare le modifiche che richiedevano un'implementazione. blue/green Le finestre di manutenzione sono ora obsolete a favore delle finestre non di punta, che sono periodi giornalieri di 10 ore durante i quali il traffico sul dominio è generalmente scarso. Puoi modificare l'ora di inizio predefinita per la finestra non di punta, ma non puoi modificare la lunghezza.

Tutti i domini che avevano le finestre di manutenzione Auto-Tune abilitate prima dell'introduzione delle finestre non di punta il 16 febbraio 2023 possono continuare a utilizzare le finestre di manutenzione precedenti senza interruzioni. Tuttavia, ti consigliamo di migrare i domini esistenti per utilizzare invece la finestra non di punta per la manutenzione del dominio. Per istruzioni, consulta the section called “Migrazione dalle finestre di manutenzione di Auto-Tune”.

Console

Per pianificare le azioni di Auto-Tune, la finestra non di punta

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione, in Domini, scegli il nome di dominio per aprire la configurazione del cluster.
3. Vai alla scheda Auto-Tune e scegli Modifica.
4. Scegli Attiva se Auto-Tune non è già abilitato.
5. In Pianifica le ottimizzazioni durante le finestre non di punta, seleziona Finestra non di punta.
6. Scegli Save changes (Salva modifiche).

CLI

Per configurare il tuo dominio in modo da pianificare le azioni di Auto-Tune durante la finestra configurata non di punta, includi nella richiesta: `UseOffPeakWindow` UpdateDomainConfig

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --auto-tune-options
 DesiredState=ENABLED,UseOffPeakWindow=true,MaintenanceSchedules=null
```

Monitoraggio delle modifiche Auto-Tune

È possibile monitorare le statistiche di Auto-Tune in. Amazon CloudWatch Per un elenco completo di parametri, consulta the section called “Metriche Auto-Tune”.

OpenSearch Il servizio invia eventi Auto-Tune ad Amazon EventBridge. Puoi utilizzarlo EventBridge per configurare regole che inviano un'e-mail o eseguono un'azione specifica quando viene ricevuto un evento. Per vedere il formato di ogni evento Auto-Tune inviato a EventBridge, vedithe section called “Eventi di regolazione automatica”.

Taggare i domini Amazon OpenSearch Service

I tag ti consentono di assegnare informazioni arbitrarie a un dominio Amazon OpenSearch Service in modo da poterle classificare e filtrare in base a tali informazioni. Un tag è una coppia chiave-valore che definisci e associ a un dominio di servizio. OpenSearch Puoi utilizzare questi tag per tenere traccia dei costi raggruppando le spese per risorse con tag simili. AWS non applica alcun significato semantico ai tag. I tag sono interpretati prettamente come stringhe di caratteri. Tutti i tag includono gli elementi seguenti:

Elemento del tag	Description	Richiesto
Chiave tag	La chiave di tag corrisponde al nome del tag. La chiave deve essere univoca per il dominio del OpenSearch servizio a cui è associata. Per un elenco di restrizioni di base su chiavi e valori dei tag, consulta la sezione relativa alle restrizioni per i tag definiti dall'utente.	Sì
Valore di tag	Il valore di tag è un valore di stringa del tag. I valori dei tag possono essere `null` e non devono necessariamente essere univoci in un set di tag. Ad esempio, puoi avere una coppia chiave-valore in un set di tag di. project/Trinity and cost-center/Trinity Per un elenco di restrizio ni di base su chiavi e valori dei tag, consulta la sezione relativa alle restrizioni per i tag definiti dall'utente.	No

Ogni dominio OpenSearch di servizio ha un set di tag, che contiene tutti i tag assegnati a quel dominio di OpenSearch servizio. AWS non assegna automaticamente alcun tag ai domini di OpenSearch servizio. Un set di tag può contenere tra 0 e 50 tag. Se si aggiunge un tag a un dominio con la stessa chiave di un tag esistente, il nuovo valore sovrascrive il vecchio valore.

Esempi di assegnazione di tag

È possibile utilizzare una chiave di tag per definire una categoria e il valore del tag potrebbe essere un elemento di tale categoria. Ad esempio, è possibile definire una chiave di tag `project` e un valore di tag di`Salix`, che indicano che il dominio OpenSearch Service è assegnato al progetto Salix. È inoltre possibile utilizzare i tag per designare i domini di OpenSearch servizio da utilizzare per il test o la produzione utilizzando una chiave come o. `environment=test` `environment=production`

Prova a utilizzare un set coerente di chiavi di tag per semplificare il monitoraggio dei metadati associati OpenSearch ai domini di servizio.

Puoi anche utilizzare i tag per organizzare la AWS fattura in modo che rifletta la tua struttura dei costi. A tale scopo, registrati per ricevere la Account AWS fattura con i valori chiave dell'etichetta inclusi. Quindi, per visualizzare il costo delle risorse combinate, puoi organizzare le informazioni di fatturazione in base alle risorse con gli stessi valori di chiave di tag. Ad esempio, puoi taggare diversi domini di OpenSearch servizio con coppie chiave-valore e quindi organizzare i dati di fatturazione in modo da visualizzare il costo totale di ogni dominio per diversi servizi. Per ulteriori informazioni, consultare la sezione relativa all'utilizzo dei tag per l'allocazione dei costi nella documentazione relativa a gestione di costi e fatturazione AWS .

> **Note**
>
> I tag sono memorizzati nella cache a fini di autorizzazione. Per questo motivo, le aggiunte e gli aggiornamenti ai tag sui domini di OpenSearch servizio potrebbero richiedere alcuni minuti prima che siano disponibili.

Etichettatura dei domini (console)

La console è il modo più semplice per assegnare un tag a un dominio.

Per creare un tag (console)

1. Vai sul sito https://aws.amazon.com e scegli Sign In to the Console (Accedi alla console).
2. In Analytics, scegli Amazon OpenSearch Service.
3. Seleziona il dominio a cui aggiungere i tag e passa alla scheda Tag.
4. Scegli Gestisci, quindi seleziona Aggiungi nuovo tag.
5. Inserire una chiave di tag e un valore facoltativo.
6. Scegli Save (Salva).

Per eliminare un tag, esegui la stessa procedura e scegli Rimuovi nella pagina Gestisci tag.

Per ulteriori informazioni sull'utilizzo della console per il funzionamento con i tag, consulta Editor di tag nella Guida alle operazioni di base della Console di gestione AWS .

Taggare i domini ()AWS CLI

Puoi creare tag di risorse usando il --add-tags comando AWS CLI with the.

Sintassi

```
add-tags --arn=<domain_arn> --tag-list Key=<key>,Value=<value>
```

Parametro	Description
`--arn`	Nome della risorsa Amazon per il dominio di OpenSearch servizio a cui è associato il tag.
`--tag-list`	Insieme di coppie chiave-valore separate da spazi nel seguente formato: `Key=<key>,Value=<value>`

Esempio

Nell'esempio seguente vengono creati due tag per il dominio logs:

```
aws opensearch add-tags --arn arn:aws:es:us-east-1:379931976431:domain/logs --tag-list
 Key=service,Value=OpenSearch Key=instances,Value=m3.2xlarge
```

Puoi rimuovere i tag da un dominio OpenSearch di servizio utilizzando il --remove-tags comando.

Sintassi

```
remove-tags --arn=<domain_arn> --tag-keys Key=<key>,Value=<value>
```

Parametro	Description
`--arn`	Amazon Resource Name (ARN) per il dominio OpenSearch di servizio a cui è associato il tag.
`--tag-keys`	Set di coppie chiave-valore separate da spazi che desideri rimuovere dal dominio del servizio. OpenSearch

Esempio

Nell'esempio seguente, i due tag creati nell'esempio precedente vengono rimossi dal dominio logs:

```
aws opensearch remove-tags --arn arn:aws:es:us-east-1:379931976431:domain/logs --tag-
keys service instances
```

È possibile visualizzare i tag esistenti per un dominio di OpenSearch servizio con il comando: --list-tags

Sintassi

```
list-tags --arn=<domain_arn>
```

Parametro	Description
`--arn`	Amazon Resource Name (ARN) per il dominio di OpenSearch servizio a cui sono allegati i tag.

Esempio

Nell'esempio seguente vengono elencati tutti i tag di risorsa per il dominio logs:

```
aws opensearch list-tags --arn arn:aws:es:us-east-1:379931976431:domain/logs
```

Etichettatura dei domini ()AWS SDKs

AWS SDKs (eccetto Android e iOS SDKs) supportano tutte le azioni definite nell'Amazon OpenSearch Service API Reference, tra cui `AddTags`, `ListTags`, e `RemoveTags` le operazioni. Per ulteriori informazioni sull'installazione e l'utilizzo di AWS SDKs, consulta AWS Software Development Kits.

Python

Questo esempio utilizza il client Python di OpenSearchServicebasso livello AWS dell'SDK for Python (Boto) per aggiungere un tag a un dominio, elencare il tag associato al dominio e rimuovere un tag dal dominio. È necessario fornire valori per `DOMAIN_ARN`, `TAG_KEY` e `TAG_VALUE`.

```
import boto3
from botocore.config import Config  # import configuration
```

```
DOMAIN_ARN = ''  # ARN for the domain. i.e "arn:aws:es:us-east-1:123456789012:domain/
my-domain
TAG_KEY = ''  # The name of the tag key. i.e 'Smileyface'
TAG_VALUE = ''  # The value assigned to the tag. i.e 'Practicetag'

# defines the configurations parameters such as region

my_config = Config(region_name='us-east-1')
client = boto3.client('opensearch', config=my_config)


# defines the client variable

def addTags():
    """Adds tags to the domain"""

    response = client.add_tags(ARN=DOMAIN_ARN,
                               TagList=[{'Key': TAG_KEY,
                                         'Value': TAG_VALUE}])

    print(response)


def listTags():
    """List tags that have been added to the domain"""

    response = client.list_tags(ARN=DOMAIN_ARN)
    print(response)


def removeTags():
    """Remove tags that have been added to the domain"""

    response = client.remove_tags(ARN=DOMAIN_ARN, TagKeys=[TAG_KEY])

    print('Tag removed')
    return response
```

Esecuzione di azioni amministrative sui domini Amazon OpenSearch Service

Amazon OpenSearch Service offre diverse opzioni amministrative che forniscono un controllo granulare se devi risolvere problemi con il tuo dominio. Queste opzioni includono la possibilità di riavviare il OpenSearch processo su un nodo di dati e la possibilità di riavviare un nodo di dati.

OpenSearch Il servizio monitora i parametri di integrità dei nodi e, in caso di anomalie, intraprende azioni correttive per mantenere stabili i domini. Con le opzioni amministrative per riavviare il OpenSearch processo su un nodo e riavviare un nodo stesso, hai il controllo su alcune di queste azioni di mitigazione.

È possibile utilizzare Console di gestione AWS AWS CLI, o l' AWS SDK per eseguire queste azioni. Le seguenti sezioni spiegano come eseguire queste azioni con la console.

Riavvio del OpenSearch processo su un nodo in Amazon Service OpenSearch

Per riavviare il OpenSearch processo su un nodo

1. Accedi alla console OpenSearch di servizio all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Domains (Domini). Scegli il nome del dominio con cui vuoi lavorare.
3. Dopo l'apertura della pagina dei dettagli del dominio, vai alla scheda Integrità dell'istanza.
4. In Nodi di dati, seleziona il pulsante accanto al nodo su cui desideri riavviare il processo.
5. Seleziona il menu a discesa Azioni e scegli Riavvia il OpenSearch/Elasticsearch processo.
6. Scegli Conferma nella modalità modale.
7. Per vedere lo stato dell'azione che hai avviato, seleziona il nome del nodo. Dopo l'apertura della pagina dei dettagli del nodo, scegli la scheda Eventi sotto il nome del nodo per visualizzare un elenco di eventi associati a quel nodo.

Riavvio di un nodo di dati in Amazon Service OpenSearch

Per riavviare un nodo di dati

1. Accedi alla console di OpenSearch servizio all'indirizzohttps://console.aws.amazon.com/aos/.

2. Nel riquadro di navigazione a sinistra, scegli Domains (Domini). Scegli il nome del dominio con cui vuoi lavorare.
3. Dopo l'apertura della pagina dei dettagli del dominio, vai alla scheda Integrità dell'istanza.
4. In Nodi di dati, seleziona il pulsante accanto al nodo su cui desideri riavviare il processo.
5. Seleziona il menu a discesa Azioni e scegli il nodo Riavvia.
6. Scegli Conferma nella modalità modale.
7. Per vedere lo stato dell'azione che hai avviato, seleziona il nome del nodo. Dopo l'apertura della pagina dei dettagli del nodo, scegli la scheda Eventi sotto il nome del nodo per visualizzare un elenco di eventi associati a quel nodo.

Riavvio del processo OpenSearch Dashboards su un nodo in Amazon Service OpenSearch

Puoi riavviare il processo OpenSearch Dashboards (precedentemente Kibana) per risolvere problemi come un'interfaccia bloccata, errori di caricamento o visualizzazioni che non rispondono. L'opzione per riavviare OpenSearch Dashboards è disponibile solo per il nodo che esegue attivamente il processo Dashboards. Nella maggior parte dei domini OpenSearch di servizio, questo processo viene eseguito su nodi coordinatori dedicati, non su nodi di dati. Di conseguenza, quando apri il menu a discesa Azioni nella console, l'opzione viene in genere visualizzata solo per i nodi coordinatori. Per ulteriori informazioni, consulta the section called “Nodi coordinatori dedicati”.

Questo comportamento dipende da come è configurato il dominio.

- Domini con nodi coordinatori dedicati: il processo Dashboards viene eseguito esclusivamente su tali nodi e solo in essi viene mostrata l'opzione di riavvio.
- Domini senza nodi coordinatori dedicati: nelle configurazioni più semplici o meno recenti, i dashboard potrebbero invece essere eseguiti su un nodo di dati e lì viene visualizzata l'opzione di riavvio.
- Nodi master: questi nodi servono esclusivamente alla gestione dello stato e delle elezioni del cluster. Non eseguono dashboard e non mostrano mai l'opzione di riavvio.

Per determinare su quale nodo è in esecuzione il processo Dashboards, vai alla sezione Configurazione del cluster del tuo dominio e rivedi i ruoli dei nodi. L'opzione di riavvio è disponibile solo per il nodo che ospita il processo Dashboards.

Per riavviare il processo Dashboards su un nodo

1. Accedi alla console di OpenSearch servizio all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Domains (Domini). Scegli il nome del dominio con cui vuoi lavorare.
3. Dopo l'apertura della pagina dei dettagli del dominio, vai alla scheda Integrità dell'istanza.
4. Nella sezione relativa ai nodi che eseguono il processo Dashboards, seleziona il pulsante accanto al nodo su cui desideri riavviare il processo.
5. Seleziona il menu a discesa Azioni e scegli Riavvia Dashboard/Kibana il processo.
6. Scegli Conferma nella modalità modale.
7. Per vedere lo stato dell'azione che hai avviato, seleziona il nome del nodo. Dopo l'apertura della pagina dei dettagli del nodo, scegli la scheda Eventi sotto il nome del nodo per visualizzare un elenco di eventi associati a quel nodo.

Limitazioni

Le opzioni amministrative presentano le seguenti limitazioni:

- Le opzioni amministrative sono supportate nelle versioni 7.x e successive di Elasticsearch.
- Le opzioni amministrative non supportano i domini con Multi-AZ con Standby abilitato.
- Il OpenSearch riavvio del processo di Elasticsearch e il riavvio del nodo dati sono supportati nei domini con tre o più nodi di dati.
- Il supporto dei processi Dashboards e Kibana è supportato su domini con due o più nodi di dati.
- Per riavviare il OpenSearch processo su un nodo o riavviare un nodo, il dominio non deve essere in rosso e tutti gli indici devono avere delle repliche configurate.

Utilizzo delle query dirette OpenSearch di Amazon Service

Usa la query diretta di Amazon OpenSearch Service per analizzare i dati in Amazon CloudWatch Logs, Amazon S3, Amazon Security Lake e Amazon Managed Service for Prometheus senza creare pipeline di ingestione. Questa integrazione zero-ETL ti consente di interrogare i dati sul posto utilizzando ProMQL, PPL o SQL ed esplorarli in Discover.

Per iniziare a usare Amazon Managed Service for Prometheus CloudWatch , Logs o Security Lake, configura la tua origine dati nella console di gestione.AWS Per Amazon S3, usa connessioni di dominio e crea tabelle con SQL in Query Workbench. CloudWatch Logs e Security Lake utilizzano sorgenti di dati e schemi preconfigurati. Per Amazon S3 e Security Lake, i dati vengono catalogati utilizzando AWS Glue Data Catalog tabelle: Amazon S3 richiede di creare queste tabelle manualmente, mentre Security Lake le configura automaticamente come parte del processo di inserimento.

Interrogazione diretta dei dati di Amazon S3 in Service OpenSearch

Questa sezione ti illustrerà il processo di creazione e configurazione di un'integrazione di origini dati in Amazon OpenSearch Service, consentendoti di interrogare e analizzare in modo efficiente i dati archiviati in Amazon S3.

Nelle pagine seguenti, imparerai come configurare un'origine dati con query diretta di Amazon S3, esplorare i prerequisiti necessari e seguire le step-by-step procedure utilizzando sia l'API che quella del Console di gestione AWS servizio. OpenSearch Descrive anche importanti passaggi successivi, tra cui la mappatura dei AWS Glue Data Catalog ruoli e la configurazione dei controlli di accesso nelle dashboard. OpenSearch

Argomenti

- Creazione di un'integrazione dell'origine dati Amazon S3 in Service OpenSearch
- Configurazione e interrogazione di un'origine dati S3 nei dashboard OpenSearch
- Prezzi
- Limitazioni
- Raccomandazioni
- Quote
- Supportato Regioni AWS

Creazione di un'integrazione dell'origine dati Amazon S3 in Service OpenSearch

Puoi creare una nuova origine dati Amazon S3 a query diretta per OpenSearch Service tramite o l' Console di gestione AWS API. Ogni nuova fonte di dati utilizza AWS Glue Data Catalog per gestire le tabelle che rappresentano i bucket Amazon S3.

Argomenti

- Prerequisiti
- Procedura
- Fasi successive
- Mappa il ruolo AWS Glue Data Catalog
- Risorse aggiuntive

Prerequisiti

Prima di iniziare, assicurati di aver letto la seguente documentazione:

- the section called “Limitazioni”
- the section called “Raccomandazioni”
- the section called “Quote”

Prima di poter creare un'origine dati, devi disporre delle seguenti risorse Account AWS:

- Un OpenSearch dominio con versione 2.13 o successiva. Questa è la base per configurare l'integrazione delle query dirette. Per istruzioni sulla configurazione, consultathe section called “Creazione OpenSearch di domini di servizio”.
- Uno o più bucket S3. Dovrai specificare i bucket contenenti i dati che desideri interrogare e un bucket in cui archiviare i checkpoint delle query. Per istruzioni sulla creazione di un bucket S3, consulta Creare un bucket nella guida per l'utente di Amazon S3.
- (Facoltativo) Una o più tabelle. AWS Glue L'interrogazione dei dati su Amazon S3 richiede che le tabelle siano configurate AWS Glue Data Catalog in modo che puntino ai dati S3. È necessario creare le tabelle utilizzando OpenSearch Query Workbench. Le tabelle Hive esistenti non sono compatibili.

Se è la prima volta che configuri un'origine dati Amazon S3, devi creare un'origine dati di amministrazione per configurare tutte le AWS Glue Data Catalog tabelle. Puoi farlo installando OpenSearch out-of-the-box integrazioni o utilizzando OpenSearch Query Workbench per creare tabelle SQL personalizzate per casi d'uso avanzati. Per esempi sulla creazione di tabelle per VPC e log AWS WAF, consulta la documentazione su GitHub VPC e. CloudTrail CloudTrailAWS WAF Dopo aver creato le tabelle, puoi creare nuove fonti di dati Amazon S3 e limitare l'accesso a tabelle limitate.

- (Facoltativo) Un ruolo IAM creato manualmente.Puoi utilizzare questo ruolo per gestire l'accesso alla tua fonte di dati. In alternativa, puoi fare in modo che OpenSearch Service crei automaticamente un ruolo per te con le autorizzazioni richieste. Se scegli di utilizzare un ruolo IAM creato manualmente, segui le indicazioni riportate inthe section called “Autorizzazioni richieste per i ruoli IAM creati manualmente”.

Procedura

Puoi configurare un'origine dati con interrogazione diretta su un dominio con Console di gestione AWS o l'API OpenSearch Service.

Per configurare una fonte di dati utilizzando il Console di gestione AWS

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Domains (Domini).
3. Seleziona il dominio per cui desideri configurare una nuova fonte di dati. Si apre la pagina dei dettagli del dominio.
4. Scegli la scheda Connessioni sotto i dettagli generali del dominio e trova la sezione Direct Query.
5. Scegli Configura origine dati.
6. Inserisci un nome e una descrizione opzionale per la tua nuova fonte di dati.
7. Scegli Amazon S3 con. AWS Glue Data Catalog
8. In Impostazioni di accesso tramite autorizzazione IAM, scegli come gestire l'accesso.
    a. Se desideri creare automaticamente un ruolo per questa fonte di dati, segui questi passaggi:
        i. Seleziona Crea un nuovo ruolo.
        ii. Inserire un nome per il ruolo IAM.

iii. Seleziona uno o più bucket S3 che contengono i dati da interrogare.

iv. Seleziona un bucket S3 di checkpoint in cui archiviare i checkpoint delle query.

v. Seleziona uno o più AWS Glue database o tabelle per definire quali dati possono essere interrogati. Se le tabelle non sono ancora state create, fornisci l'accesso al database predefinito.

b. Se desideri utilizzare un ruolo esistente che gestisci tu stesso, segui questi passaggi:

i. Seleziona Utilizza un ruolo esistente.

ii. Seleziona un ruolo esistente dal menu a discesa.

> **Note**
>
> Quando utilizzi il tuo ruolo, devi assicurarti che disponga di tutte le autorizzazioni necessarie allegando le policy richieste dalla console IAM. Per ulteriori informazioni, consulta la policy di esempio in. the section called “Autorizzazioni richieste per i ruoli IAM creati manualmente”

9. Scegli Configura. Si apre la schermata dei dettagli dell'origine dati con un URL del OpenSearch pannello di controllo. Puoi accedere a questo URL per completare i passaggi successivi.

OpenSearch API di servizio

Usa l'operazione AddDataSourceAPI per creare una nuova fonte di dati nel tuo dominio.

```
POST https://es.region.amazonaws.com/2021-01-01/opensearch/domain/domain-name/
dataSource

{
   "DataSourceType": {
        "S3GlueDataCatalog": {
            "RoleArn": "arn:aws:iam::account-id:role/role-name"
        }
    }
   "Description": "data-source-description",
   "Name": "my-data-source"
}
```

Fasi successive

Visita OpenSearch Dashboards

Dopo aver creato un'origine dati, OpenSearch Service ti fornisce un link OpenSearch Dashboards. Puoi usarlo per configurare il controllo degli accessi, definire tabelle, installare out-of-the-box integrazioni e interrogare i tuoi dati.

Per ulteriori informazioni, consulta the section called “Configurazione di un'origine dati S3”.

Mappa il ruolo AWS Glue Data Catalog

Se hai abilitato il controllo granulare degli accessi dopo aver creato un'origine dati, devi mappare gli utenti non amministratori a un ruolo IAM con AWS Glue Data Catalog accesso per eseguire query dirette. Per creare manualmente un ruolo di back-end che puoi mappare al `glue_access` ruolo IAM, esegui i seguenti passaggi:

> **Note**
>
> Gli indici vengono utilizzati per qualsiasi interrogazione sulla fonte di dati. Un utente con accesso in lettura all'indice delle richieste per una determinata origine dati può leggere tutte le query relative a tale origine dati. Un utente con accesso in lettura all'indice dei risultati può leggere i risultati di tutte le query eseguite su quell'origine dati.

1. Dal menu principale di OpenSearch Dashboard, scegli Sicurezza, Ruoli e Crea ruoli.
2. Assegna un nome al ruolo glue_access.
3. Per le autorizzazioni del cluster, selezionaindices:data/write/bulk*,. `indices:data/read/scroll` `indices:data/read/scroll/clear`
4. Per Indice, inserisci i seguenti indici a cui desideri concedere all'utente con il ruolo l'accesso:
   - `.query_execution_request_`*`<name of data source>`*
   - `query_execution_result_`*`<name of data source>`*
   - `.async-query-scheduler`
   - `flint_*`
5. Per le autorizzazioni di indicizzazione, seleziona. `indices_all`
6. Scegli Create (Crea).
7. Scegliere Utenti mappati, Gestisci mappatura.

8. In Ruoli di backend, aggiungi l'ARN del ruolo che richiede AWS Glue l'autorizzazione per chiamare il tuo dominio.

```
arn:aws:iam::account-id:role/role-name
```

9. Seleziona Mappa e conferma che il ruolo sia visualizzato in Utenti mappati.

Per ulteriori informazioni sulla mappatura dei ruoli, vedere. the section called “Mappatura dei ruoli agli utenti”

Risorse aggiuntive

Autorizzazioni richieste per i ruoli IAM creati manualmente

Quando crei un'origine dati per il tuo dominio, scegli un ruolo IAM per gestire l'accesso ai tuoi dati. Sono disponibili due opzioni:

1. Crea automaticamente un nuovo ruolo IAM
2. Utilizza un ruolo IAM esistente che hai creato manualmente

Se utilizzi un ruolo creato manualmente, devi assegnare le autorizzazioni corrette al ruolo. Le autorizzazioni devono consentire l'accesso alla fonte di dati specifica e consentire a OpenSearch Service di assumere il ruolo. Ciò è necessario affinché il OpenSearch Servizio possa accedere e interagire in modo sicuro con i dati dell'utente.

La seguente politica di esempio illustra le autorizzazioni con privilegi minimi necessarie per creare e gestire un'origine dati. Se disponi di autorizzazioni più ampie, ad esempio `s3:*` o la policy, queste autorizzazioni comprendono le autorizzazioni con il `AdminstratorAccess` privilegio minimo della politica di esempio.

Nella seguente politica di esempio, sostituiscila con le tue informazioni. *`placeholder text`*

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Sid":"HttpActionsForOpenSearchDomain",
```

```
         "Effect":"Allow",
         "Action":"es:ESHttp*",
"Resource":"arn:aws:es:us-east-1:111122223333:domain/example.com/*"
      },
      {
         "Sid":"AmazonOpenSearchS3GlueDirectQueryReadAllS3Buckets",
         "Effect":"Allow",
         "Action":[
            "s3:GetObject",
            "s3:GetObjectVersion",
            "s3:ListBucket"
         ],
         "Condition":{
            "StringEquals":{
               "aws:ResourceAccount":"111122223333"
            }
         },
         "Resource":"*"
      },
      {
         "Sid":"AmazonOpenSearchDirectQueryGlueCreateAccess",
         "Effect":"Allow",
         "Action":[
            "glue:CreateDatabase",
            "glue:CreatePartition",
            "glue:CreateTable",
            "glue:BatchCreatePartition"
         ],
         "Resource":"*"
      },
      {
         "Sid":"AmazonOpenSearchS3GlueDirectQueryModifyAllGlueResources",
         "Effect":"Allow",
         "Action":[
            "glue:DeleteDatabase",
            "glue:DeletePartition",
            "glue:DeleteTable",
            "glue:GetDatabase",
            "glue:GetDatabases",
            "glue:GetPartition",
            "glue:GetPartitions",
            "glue:GetTable",
            "glue:GetTableVersions",
            "glue:GetTables",
```

```
            "glue:UpdateDatabase",
            "glue:UpdatePartition",
            "glue:UpdateTable",
            "glue:BatchGetPartition",
            "glue:BatchDeletePartition",
            "glue:BatchDeleteTable"
         ],
         "Resource":[
            "arn:aws:glue:us-east-1:111122223333:table/*",
            "arn:aws:glue:us-east-1:111122223333:database/*",
            "arn:aws:glue:us-east-1:111122223333:catalog",
            "arn:aws:es:us-east-1:111122223333:domain/domain_name"
         ],
         "Condition":{
            "StringEquals":{
               "aws:ResourceAccount":"111122223333"
            }
         }
      },
      {
         "Sid":"ReadAndWriteActionsForS3CheckpointBucket",
         "Effect":"Allow",
         "Action":[
            "s3:ListMultipartUploadParts",
            "s3:DeleteObject",
            "s3:GetObject",
            "s3:PutObject",
            "s3:GetBucketLocation",
            "s3:ListBucket"
         ],
         "Condition":{
            "StringEquals":{
               "aws:ResourceAccount":"111122223333"
            }
         },
         "Resource":[
            "arn:aws:s3:::amzn-s3-demo-bucket",
            "arn:aws:s3:::amzn-s3-demo-bucket/*"
         ]
      }
   ]
}
```

Per supportare i bucket Amazon S3 in diversi account, dovrai includere una condizione nella policy di Amazon S3 e aggiungere l'account appropriato.

Nella seguente condizione di esempio, sostituiscila *placeholder text* con le tue informazioni.

```
"Condition": {
                "StringEquals": {
                    "aws:ResourceAccount": "{{accountId}}"
                }
```

Il ruolo deve inoltre avere la seguente politica di fiducia, che specifica l'ID di destinazione.

JSON

```
{
    "Version":"2012-10-17",
    "Statement":[
       {
          "Effect":"Allow",
          "Principal":{
             "Service": "directquery.opensearchservice.amazonaws.com"
          },
          "Action":"sts:AssumeRole"
       }
     ]
}
```

Per istruzioni sulla creazione del ruolo, consulta Creazione di un ruolo utilizzando policy di attendibilità personalizzate.

Se hai abilitato il controllo granulare degli accessi in OpenSearch Service, verrà creato automaticamente un nuovo ruolo di controllo OpenSearch degli accessi granulare per la tua fonte di dati. Il nome del nuovo ruolo di controllo degli accessi a grana fine sarà. `AWS OpenSearchDirectQuery` *<name of data source>*

Per impostazione predefinita, il ruolo ha accesso solo agli indici delle fonti di dati di interrogazione diretta. Sebbene sia possibile configurare il ruolo per limitare o concedere l'accesso all'origine dati, si consiglia di non modificare l'accesso di questo ruolo. Se elimini l'origine dati, questo ruolo verrà eliminato. Ciò rimuoverà l'accesso per tutti gli altri utenti se sono mappati al ruolo.

Configurazione e interrogazione di un'origine dati S3 nei dashboard OpenSearch

Ora che hai creato la tua origine dati, puoi configurare le impostazioni di sicurezza, definire le tabelle Amazon S3 o configurare l'indicizzazione accelerata dei dati. Questa sezione illustra vari casi d'uso con la tua fonte di dati nelle OpenSearch dashboard prima di interrogare i dati.

Per configurare le seguenti sezioni, devi prima accedere alla tua origine dati in OpenSearch Dashboards. Nella barra di navigazione a sinistra, in Gestione, scegli Origini dati. In Gestisci fonti di dati, seleziona il nome dell'origine dati che hai creato nella console.

Crea tabelle Spark utilizzando Query Workbench

Le query dirette dal OpenSearch Servizio ad Amazon S3 utilizzano le tabelle Spark all'interno di. AWS Glue Data Catalog Puoi creare tabelle dall'interno di Query Workbench senza dover uscire dai dashboard. OpenSearch

Per gestire database e tabelle esistenti nella tua origine dati o per creare nuove tabelle su cui desideri utilizzare le query dirette, scegli Query Workbench dalla barra di navigazione a sinistra e seleziona l'origine dati Amazon S3 dal menu a discesa delle origini dati.

Per impostare una tabella per i log di flusso VPC archiviati in S3 in formato Parquet, esegui la seguente query:

```
CREATE TABLE
datasourcename.gluedatabasename.vpclogstable (version INT, account_id STRING,
 interface_id STRING,
srcaddr STRING, dstaddr STRING, srcport INT, dstport INT, protocol INT, packets
 BIGINT,
bytes BIGINT, start BIGINT, end BIGINT, action STRING, log_status STRING,
`aws-account-id` STRING, `aws-service` STRING, `aws-region` STRING, year STRING,
month STRING, day STRING, hour STRING)

USING parquet PARTITIONED BY (aws-account-id, aws-service, aws-region, year, month,
day, hour)

LOCATION "s3://accountnum-vpcflow/AWSLogs"
```

Dopo aver creato la tabella, esegui la seguente query per assicurarti che sia compatibile con le query dirette:

```
MSCK REPAIR TABLE  datasourcename.databasename.vpclogstable
```

Imposta le integrazioni per i tipi di log più diffusi AWS

Puoi integrare i tipi di AWS log archiviati in Amazon S3 con OpenSearch Service. Usa OpenSearch le dashboard per installare integrazioni che creano AWS Glue Data Catalog tabelle, query salvate e dashboard. Queste integrazioni utilizzano viste indicizzate per mantenere aggiornate le dashboard.

Per istruzioni sull'installazione di un'integrazione, consulta Installazione di una risorsa di integrazione nella documentazione. OpenSearch

Quando selezioni un'integrazione, assicurati che abbia il `S3 Glue` tag.

Quando configuri l'integrazione, specifica S3 Connection per il tipo di connessione. Quindi, seleziona l'origine dati per l'integrazione, la posizione dei dati in Amazon S3, il checkpoint per gestire l'indicizzazione dell'accelerazione e gli asset necessari per il tuo caso d'uso.

> **Note**
>
> Assicurati che il bucket S3 per il checkpoint disponga delle autorizzazioni di scrittura per la posizione del checkpoint. Senza queste autorizzazioni, le accelerazioni dell'integrazione falliranno.

Configurazione del controllo degli accessi

Nella pagina dei dettagli della tua origine dati, trova la sezione Controlli di accesso e scegli Modifica. Se il dominio ha abilitato il controllo granulare degli accessi, scegli Restricted e seleziona i ruoli a cui desideri fornire l'accesso alla nuova fonte di dati. Puoi anche scegliere Amministratore solo se desideri che solo l'amministratore abbia accesso all'origine dati.

> **Important**
>
> Gli indici vengono utilizzati per qualsiasi interrogazione sull'origine dati. Un utente con accesso in lettura all'indice delle richieste per una determinata origine dati può leggere tutte le query relative a tale origine dati. Un utente con accesso in lettura all'indice dei risultati può leggere i risultati di tutte le query eseguite su quell'origine dati.

Interrogazione dei dati S3 in Discover OpenSearch

Dopo aver impostato le tabelle e configurato l'accelerazione delle query opzionale desiderata, puoi iniziare ad analizzare i dati. Per interrogare i dati, seleziona la fonte dei dati dal menu a discesa. Se utilizzi Amazon S3 e OpenSearch Dashboards, vai su Discover e seleziona il nome dell'origine dati.

Se utilizzi un indice ignorante o non hai ancora creato un indice, puoi usare SQL o PPL per interrogare i tuoi dati. Se hai configurato una vista materializzata o un indice di copertura, disponi già di un indice e puoi utilizzare Dashboards Query Language (DQL) in tutte le dashboard. Puoi anche usare PPL con il plug-in Observability e SQL con il plug-in Query Workbench. Attualmente, solo i plugin Observability e Query Workbench supportano PPL e SQL. Per interrogare i dati utilizzando l'API di OpenSearch servizio, consulta la documentazione dell'API asincrona.

> **Note**
>
> Non tutte le istruzioni, i comandi e le funzioni SQL e PPL sono supportati. Per un elenco dei comandi supportati, vederethe section called “Comandi SQL e PPL supportati”.
> Se hai creato una vista materializzata o un indice di copertura, puoi usare DQL per interrogare i tuoi dati, purché li abbia indicizzati all'interno.

Risoluzione dei problemi

In alcuni casi i risultati non vengono restituiti come previsto. In caso di problemi, assicurati di seguire ilthe section called “Raccomandazioni”.

Prezzi

Amazon OpenSearch Service offre prezzi OpenSearch Compute Unit (OCU) per le query dirette su Amazon S3. Quando esegui le query dirette, ti vengono addebitati addebiti orari, indicati nella fattura come OCUs DirectQuery tipo di utilizzo dell'OCU. Saranno inoltre addebitati costi separati da Amazon S3 per lo storage dei dati.

Le query dirette sono di due tipi: interattive e con visualizzazione indicizzata.

- Le query interattive vengono utilizzate per compilare il selettore di dati ed eseguire analisi sui dati in Amazon S3. Quando esegui una nuova query da Discover, OpenSearch Service avvia una nuova sessione che dura almeno tre minuti. OpenSearch Il servizio mantiene attiva questa sessione per garantire che le query successive vengano eseguite rapidamente.

- Le query di visualizzazione indicizzata utilizzano il calcolo per mantenere le viste indicizzate nel Servizio. OpenSearch Queste query in genere richiedono più tempo perché inseriscono una quantità variabile di dati in un indice denominato. Per le origini dati Amazon S3, i dati indicizzati vengono archiviati in un dominio in base al tipo di istanza acquistato.

Per ulteriori informazioni, consulta le sezioni Direct Query e Serverless all'interno dei prezzi OpenSearch di Amazon Service.

Limitazioni

Le seguenti limitazioni si applicano alle query dirette in Amazon S3:

- La query diretta per S3 è disponibile solo nei domini di OpenSearch servizio che eseguono la OpenSearch versione 2.13 o successiva e richiede l'accesso a. AWS Glue Data Catalog AWS Glue Data Catalog Le tabelle esistenti devono essere ricreate utilizzando SQL in Query Workbench. OpenSearch
- Direct Query for S3 richiede di specificare un bucket di checkpoint su Amazon S3. Questo bucket mantiene lo stato delle viste indicizzate, inclusa l'ora dell'ultimo aggiornamento e i dati inseriti più di recente.
- Il tuo OpenSearch dominio deve appartenere allo stesso. AWS Glue Data Catalog Account AWS Il tuo bucket S3 può trovarsi in un account diverso (richiede l'aggiunta di una condizione alla tua policy IAM), ma deve appartenere allo Regione AWS stesso dominio.
- OpenSearch Le query Service Direct con S3 supportano solo le tabelle Spark generate da Query Workbench. Le tabelle generate all'interno di AWS Glue Data Catalog o Athena non sono supportate dallo streaming Spark, necessario per mantenere le viste indicizzate.
- OpenSearch i tipi di istanza hanno limiti di payload in rete di 10 MiB o 100 MiB, a seconda del tipo di istanza specifico scelto.
- Alcuni tipi di dati non sono supportati. I tipi di dati supportati sono limitati a Parquet, CSV e JSON.
- Se la struttura dei dati cambia nel tempo, dovrai aggiornare le viste o le out-of-the-box integrazioni indicizzate per tenere conto delle modifiche alla struttura dei dati.
- AWS CloudFormation i modelli non sono ancora supportati.
- OpenSearch Le istruzioni SQL e OpenSearch PPL presentano limitazioni diverse quando si lavora con OpenSearch gli indici rispetto all'utilizzo di query dirette. Direct Query supporta comandi avanzati come JOINs sottoquery e ricerche, mentre il supporto per questi comandi sugli

OpenSearch indici è limitato o inesistente. Per ulteriori informazioni, consulta the section called "Comandi SQL e PPL supportati".

Raccomandazioni

Per l'utilizzo di query dirette in Amazon S3, consigliamo quanto segue:

- Inserisci dati in Amazon S3 utilizzando i formati di partizione di anno, mese, giorno e ora per velocizzare le query.
- Quando crei indici ignoranti, usa i filtri Bloom per i campi con cardinalità elevata e min/max gli indici per i campi con intervalli di valori ampi. Per i campi ad alta cardinalità, prendi in considerazione l'utilizzo di un approccio basato sui valori per migliorare l'efficienza delle query.
- Utilizzate Index State Management per mantenere lo spazio di archiviazione per le viste materializzate e gli indici di copertura.
- Utilizzate la `COALESCE SQL` funzione per gestire le colonne mancanti e assicurarvi che i risultati vengano restituiti.
- Usa i limiti per le tue query per assicurarti di non recuperare troppi dati.

Quote

Ogni volta che si avvia una query su un'origine dati Amazon S3 OpenSearch , Service apre una sessione e la mantiene attiva per almeno tre minuti. Ciò riduce la latenza delle query eliminando il tempo di avvio della sessione nelle query successive.

Description	Massimo	Può sovrascrivere
Connessioni per dominio	10	Sì
Fonti di dati per dominio	20	Sì
Indici per dominio	5	Sì
Sessioni simultanee per origine dati	10	Sì
Numero massimo di OCU per query	60	Sì

Description	Massimo	Può sovrascrivere
Tempo massimo di esecuzione della query (minuti)	30	Sì
Massimo OCUs per acceleraz ione	20	Sì
Archiviazione effimera massima	20	Sì

Supportato Regioni AWS

Quanto segue Regioni AWS è supportato per le query dirette in Amazon S3:

- Asia Pacifico (Hong Kong)
- Asia Pacifico (Mumbai)
- Asia Pacifico (Seoul)
- Asia Pacifico (Singapore)
- Asia Pacifico (Sydney)
- Asia Pacifico (Tokyo)
- Canada (Centrale)
- Europa (Francoforte)
- Europa (Irlanda)
- Europa (Stoccolma)
- Stati Uniti orientali (Virginia settentrionale)
- Stati Uniti orientali (Ohio)
- Stati Uniti occidentali (Oregon)

Interrogazione diretta dei dati di Amazon CloudWatch Logs in Service OpenSearch

Questa sezione ti illustrerà il processo di creazione e configurazione di un'integrazione di fonti di dati in Amazon OpenSearch Service, consentendoti di interrogare e analizzare in modo efficiente i dati archiviati nei CloudWatch log.

Nelle pagine seguenti, imparerai come configurare un'origine dati CloudWatch Logs con interrogazione diretta, esplorare i prerequisiti necessari e seguire le procedure utilizzando. step-by-step Console di gestione AWS

Argomenti

- Creazione di un'integrazione dell'origine dati Amazon CloudWatch Logs in Service OpenSearch
- Configurazione e interrogazione di un'origine dati CloudWatch Logs nei dashboard OpenSearch
- Prezzi
- Limitazioni
- Raccomandazioni
- Quote
- Supportato Regioni AWS

Creazione di un'integrazione dell'origine dati Amazon CloudWatch Logs in Service OpenSearch

Se utilizzi Amazon OpenSearch Serverless per le tue esigenze di osservabilità, ora puoi analizzare i tuoi Amazon CloudWatch Logs senza copiare o importare i dati in Service. OpenSearch Questa funzionalità sfrutta l'interrogazione diretta per l'interrogazione dei dati, in modo simile all'analisi dei dati in Amazon S3 from Service. OpenSearch Puoi iniziare creando una nuova fonte di dati connessa dalla console di gestione. AWS

Puoi creare una nuova fonte di dati per analizzare i dati di CloudWatch Logs senza dover creare Amazon OpenSearch Serverless per interrogare direttamente i log operativi in Logs. CloudWatch Ciò consente di analizzare i dati operativi a cui si accede e che si trovano al di fuori del Servizio. OpenSearch Eseguendo interrogazioni tra OpenSearch Service e CloudWatch Logs, è possibile iniziare ad analizzare i log in CloudWatch Logs e poi tornare al monitoraggio delle fonti di dati senza dover cambiare strumento. OpenSearch

Per utilizzare questa funzionalità, è necessario creare un'origine dati CloudWatch Logs Direct Query per OpenSearch Service tramite la Console di gestione. AWS

Argomenti

- Prerequisiti
- Procedura
- Fasi successive

- Altre risorse

Prerequisiti

Prima di iniziare, assicurati di aver letto la seguente documentazione:

- the section called “Limitazioni”
- the section called “Raccomandazioni”
- the section called “Quote”

Prima di poter creare un'origine dati, devi disporre delle seguenti risorse Account AWS:

- Abilita CloudWatch i registri. Configura CloudWatch Logs per raccogliere i log sulla Account AWS stessa risorsa. OpenSearch Per istruzioni, consulta Getting started with CloudWatch Logs nella guida per l'utente di Amazon CloudWatch Logs.
- Uno o più gruppi di CloudWatch log.È possibile specificare i gruppi di log contenenti i dati che si desidera interrogare. Per istruzioni sulla creazione di un gruppo di log, consulta Creare un gruppo di log in CloudWatch Logs nella guida per l'utente di Amazon CloudWatch Logs.
- (Facoltativo) Un ruolo IAM creato manualmente.Puoi utilizzare questo ruolo per gestire l'accesso alla tua fonte di dati. In alternativa, puoi fare in modo che OpenSearch Service crei automaticamente un ruolo per te con le autorizzazioni richieste. Se scegli di utilizzare un ruolo IAM creato manualmente, segui le indicazioni riportate inthe section called “Autorizzazioni richieste per i ruoli IAM creati manualmente”.

Procedura

Puoi configurare un'origine dati di interrogazione a livello di raccolta con. Console di gestione AWS

Per configurare un'origine dati a livello di raccolta utilizzando il Console di gestione AWS

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, vai a Gestione centrale e scegli Origini dati connesse.
3. Scegli Connetti.
4. Scegli CloudWatchcome tipo di origine dati.
5. Scegli Next (Successivo).

6. In Dettagli della connessione dati, inserisci un nome e una descrizione opzionale.
7. In Ruoli IAM, scegli come gestire l'accesso ai gruppi di log.

   a. Se desideri creare automaticamente un ruolo per questa fonte di dati, segui questi passaggi:

      i. Seleziona Crea un nuovo ruolo.

      ii. Inserire un nome per il ruolo IAM.

      iii. Seleziona uno o più gruppi di log per definire quali dati possono essere interrogati.

   b. Se desideri utilizzare un ruolo esistente che gestisci tu stesso, segui questi passaggi:

      i. Seleziona Utilizza un ruolo esistente.

      ii. Seleziona un ruolo esistente dal menu a discesa.

   > **Note**
   >
   > Quando utilizzi il tuo ruolo, devi assicurarti che disponga di tutte le autorizzazioni necessarie allegando le policy richieste dalla console IAM. Per ulteriori informazioni, consulta the section called “Autorizzazioni richieste per i ruoli IAM creati manualmente”.

8. (Facoltativo) In Politica di accesso, configura una politica di accesso per l'origine dati. Le politiche di accesso controllano se una richiesta all'origine dati OpenSearch di Service Direct Query viene accettata o rifiutata. Se non configuri una politica di accesso, solo il proprietario dell'origine dati può accedervi. Puoi configurare la politica di accesso per abilitare l'accesso tra più account, consentendo ai responsabili di altri di accedere Account AWS alla fonte di dati.

   È possibile creare una politica di accesso utilizzando l'editor visivo o fornendo un documento di policy JSON. Con l'editor visivo, puoi consentire o negare l'accesso specificando un Account AWS ID principale, l'ARN dell'account, l'ARN dell'utente IAM, l'ARN del ruolo IAM, l'indirizzo IP di origine o il blocco CIDR. L'editor visivo supporta fino a 10 elementi. Per definire una policy con più di 10 elementi, usa l'editor JSON.

   Puoi anche scegliere Importa policy per importare una politica di accesso esistente da un'altra fonte di dati.
9. (Facoltativo) In Tag, aggiungi tag alla tua fonte di dati.
10. Scegli Next (Successivo).
11. In Configurazione OpenSearch, scegli come eseguire la configurazione OpenSearch.

a. Usa le impostazioni predefinite:

   - Rivedi i nomi delle risorse predefiniti e le impostazioni di conservazione dei dati. Ti consigliamo di utilizzare nomi personalizzati.

     Quando utilizzi le impostazioni predefinite, vengono create automaticamente una nuova OpenSearch applicazione e un'area di lavoro Essentials senza costi aggiuntivi. OpenSearch consente di analizzare più fonti di dati. Include aree di lavoro, che offrono esperienze personalizzate per i casi d'uso più diffusi. Le aree di lavoro supportano il controllo degli accessi, consentendoti di creare spazi privati per i tuoi casi d'uso e condividerli solo con i tuoi collaboratori.

b. Usa impostazioni personalizzate:

   i. Scegliere Customize (Personalizza).

   ii. Modifica il nome della raccolta e le impostazioni di conservazione dei dati secondo necessità.

   iii. Seleziona l' OpenSearch applicazione e l'area di lavoro che desideri utilizzare.

12. Scegli Next (Successivo).
13. Rivedi le tue scelte e scegli Modifica se devi apportare modifiche.
14. Scegli Connect per configurare l'origine dati. Rimani su questa pagina mentre viene creata la tua fonte di dati. Quando sarà pronta, verrai indirizzato alla pagina dei dettagli dell'origine dati.

Fasi successive

Visita OpenSearch Dashboards

Dopo aver creato un'origine dati, OpenSearch Service ti fornisce un URL per i OpenSearch dashboard. È possibile utilizzarlo per configurare il controllo degli accessi, definire tabelle, configurare dashboard basati sui tipi di log per i tipi di log più diffusi e interrogare i dati utilizzando SQL o PPL.

Per ulteriori informazioni, consulta the section called “Configurazione di un'origine dati Logs CloudWatch ”.

Altre risorse

Autorizzazioni richieste per i ruoli IAM creati manualmente

Quando crei un'origine dati, scegli un ruolo IAM per gestire l'accesso ai tuoi dati. Sono disponibili due opzioni:

1. Crea automaticamente un nuovo ruolo IAM
2. Utilizza un ruolo IAM esistente che hai creato manualmente

Se utilizzi un ruolo creato manualmente, devi assegnare le autorizzazioni corrette al ruolo. Le autorizzazioni devono consentire l'accesso alla fonte di dati specifica e consentire a OpenSearch Service di assumere il ruolo. Ciò è necessario affinché il OpenSearch Servizio possa accedere e interagire in modo sicuro con i dati dell'utente.

La seguente politica di esempio illustra le autorizzazioni con privilegi minimi necessarie per creare e gestire un'origine dati. Se disponi di autorizzazioni più ampie, ad esempio `logs:*` o la policy, queste autorizzazioni comprendono le autorizzazioni con il `AdminstratorAccess` privilegio minimo della politica di esempio.

Nella seguente politica di esempio, sostituiscila con le tue informazioni. *`placeholder text`*

JSON

```
{
"Version":"2012-10-17",
"Statement": [
    {
        "Sid": "AmazonOpenSearchDirectQueryAllLogsAccess",
        "Effect": "Allow",
        "Action": [
            "logs:DescribeLogGroups",
            "logs:StartQuery",
            "logs:GetLogGroupFields"
        ],
        "Condition": {
            "StringEquals": {
                "aws:ResourceAccount": "111122223333"
            }
```

```
            },
            "Resource": [
                "arn:aws:logs:us-east-1:111122223333:log-group:*"
            ]
        }
    ]
}
```

Il ruolo deve inoltre avere la seguente politica di fiducia, che specifica l'ID di destinazione.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "TrustPolicyForAmazonOpenSearchDirectQueryService",
            "Effect": "Allow",
            "Principal": {
                "Service": "directquery.opensearchservice.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "ArnLike": {
                    "aws:SourceArn": "arn:aws:opensearch:us-
east-1:111122223333:datasource/rolename"
                }
            }
        }
    ]
}
```

Per istruzioni sulla creazione del ruolo, consulta Creazione di un ruolo utilizzando policy di attendibilità personalizzate.

Per impostazione predefinita, il ruolo ha accesso solo agli indici delle origini dei dati delle query dirette. Sebbene sia possibile configurare il ruolo per limitare o concedere l'accesso all'origine dati, si consiglia di non modificare l'accesso di questo ruolo. Se elimini l'origine dati, questo ruolo verrà eliminato. Ciò rimuoverà l'accesso per tutti gli altri utenti se sono mappati al ruolo.

Esempio di politica di accesso per una fonte di dati con interrogazione diretta

Le politiche di accesso per le fonti di dati con query dirette seguono la sintassi delle policy IAM. Il documento di policy deve essere in formato JSON valido. La seguente politica di esempio concede un Account AWS accesso specifico all'origine dei dati delle query dirette.

Nella seguente politica di esempio, sostituisci la *placeholder text* con le tue informazioni.

```
{
 "Version": "2012-10-17",
 "Statement": [
   {
     "Effect": "Allow",
     "Principal": {
     "AWS": "arn:aws:iam::account-id:root"
     },
     "Action": [
       "opensearch:StartDirectQuery",
       "opensearch:GetDirectQuery",
       "opensearch:CancelDirectQuery",
       "opensearch:GetDirectQueryResult"
     ],
     "Resource": "arn:aws:opensearch:region:account-id:datasource/data-source-name"
   }
 ]
}
```

Se non configuri una politica di accesso, solo il proprietario dell'origine dati ha accesso all'origine dati.

Configurazione e interrogazione di un'origine dati CloudWatch Logs nei dashboard OpenSearch

Ora che hai creato la tua fonte di dati, puoi iniziare a utilizzarla Dashboards. OpenSearch Questa sezione illustra vari casi d'uso con la tua fonte di dati nelle OpenSearch dashboard.

Interroga i gruppi di log dalla pagina Discover

Nella pagina OpenSearch Discover, puoi utilizzare la nuova fonte di dati di Direct Query configurata per interrogare i gruppi di log di CloudWatch Logs. Per fare ciò, scegli Esplora i registri, quindi usa la barra di ricerca per creare la tua query usando SQL o PPL. Puoi filtrare, ordinare e visualizzare

i dati restituiti dai tuoi gruppi di log. Per comprendere quali istruzioni, comandi e limitazioni sono supportati per l'integrazione con CloudWatch Logs, consulta. the section called “Comandi SQL e PPL supportati”

Crea una visualizzazione del dashboard per la tua fonte di dati

Quando utilizzi OpenSearch Service, puoi analizzare rapidamente i tipi di AWS log più diffusi utilizzando modelli di dashboard predefiniti. Per CloudWatch i registri sono disponibili modelli per i log VPC CloudTrail e Network Firewall. AWS WAF Questi modelli consentono di creare rapidamente una dashboard personalizzata in base ai dati specifici. Includono dashboard personalizzati per quel tipo di registro specifico. Ciò consente di iniziare rapidamente ad analizzare queste fonti di AWS log popolari, senza dover creare tutto da zero.

> **Note**
>
> Le dashboard utilizzano viste indicizzate, che acquisiscono dati dai CloudWatch log utilizzando unità di OpenSearch calcolo a query diretta (OCUs), nonché l'indicizzazione, la ricerca e l'archiviazione delle raccolte senza server. OCUs OCUs

Segui questi passaggi per creare una dashboard utilizzando uno di questi modelli predefiniti, in modo da poter iniziare subito a esplorare e analizzare i tuoi dati.

Per creare una visualizzazione del pannello di controllo

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Dal riquadro di navigazione a sinistra, scegli Gestione centrale, quindi Origini dati connesse.
3. Seleziona l'origine dati per aprire la pagina dei dettagli.
4. Seleziona Crea pannello di controllo.
5. Scegli il tipo di dashboard che desideri creare.
6. Inserisci un nome per la dashboard.
7. Inserisci una descrizione opzionale per la dashboard.
8. Seleziona uno o più gruppi di log da visualizzare nella dashboard.
9. Scegli la frequenza con cui aggiornare i dati nella dashboard.
10. Scegli l' OpenSearch area di lavoro che desideri utilizzare.

a. Per creare un nuovo spazio di lavoro, seleziona Crea nuovo spazio di lavoro e inserisci un nome.
b. Per utilizzare un'area di lavoro esistente, seleziona Seleziona un'area di lavoro esistente.

11. Seleziona Crea pannello di controllo.

Interrogazione dei CloudWatch dati dei log in Discover OpenSearch

Per interrogare i dati, seleziona la fonte dei dati dal menu a discesa. Se utilizzi CloudWatch Logs, accedi a Discover dall'area di lavoro Essentials e inizia a interrogare i dati utilizzando OpenSearch SQL o Piped Processing Language (PPL). Per un elenco dei comandi supportati, consulta. the section called “Comandi SQL e PPL supportati”

> **Note**
>
> Se hai creato una vista materializzata, puoi usare DQL per interrogare i tuoi dati, purché li abbia indicizzati all'interno.

Risoluzione dei problemi

Potrebbero verificarsi casi in cui i risultati non vengono restituiti come previsto. In caso di problemi, assicurati di seguire ilthe section called “Raccomandazioni”.

Prezzi

Amazon OpenSearch Service offre prezzi OpenSearch Compute Unit (OCU) per le query dirette di CloudWatch Logs. Quando esegui query dirette, ti vengono addebitati addebiti orari, indicati nella fattura come OCUs DirectQuery tipo di utilizzo dell'OCU. Inoltre, ti verranno addebitati costi separati da Amazon CloudWatch Logs.

Le query dirette sono di due tipi: interattive e con visualizzazione indicizzata.

- Le query interattive vengono utilizzate per compilare il selettore di dati ed eseguire analisi sui dati nei registri. CloudWatch OpenSearch Il servizio gestisce ogni query con un job separato pre-riscaldato, senza mantenere una sessione prolungata.
- Le query di visualizzazione indicizzata utilizzano il calcolo per mantenere le viste indicizzate nel Servizio. OpenSearch Queste query in genere richiedono più tempo perché inseriscono una

quantità variabile di dati in un indice denominato. Per le origini dati CloudWatch connesse a Logs, i dati indicizzati vengono archiviati in una raccolta OpenSearch Serverless in cui vengono addebitati i dati indicizzati (IndexingOCU), i dati ricercati (SearchOCU) e i dati archiviati in GB.

Per ulteriori informazioni, consulta le sezioni Direct Query e Serverless all'interno dei prezzi OpenSearch di Amazon Service.

Limitazioni

Le seguenti limitazioni si applicano alle query dirette nei CloudWatch log:

- L'integrazione diretta delle query con CloudWatch Logs è disponibile solo nelle raccolte di OpenSearch servizi e nell' OpenSearch interfaccia utente.
- OpenSearch Le raccolte serverless hanno limiti di payload in rete di 100 MiB.
- CloudWatch Logs supporta VPC Flow CloudTrail AWS WAF e integrazioni di dashboard installate dalla console.
- AWS CloudFormation i modelli non sono ancora supportati.
- OpenSearch Le istruzioni SQL e OpenSearch PPL presentano limitazioni diverse quando si lavora con OpenSearch gli indici rispetto all'utilizzo di query dirette. Direct Query supporta comandi avanzati come JOINs sottoquery e ricerche, mentre il supporto per questi comandi sugli OpenSearch indici è limitato o inesistente. Per ulteriori informazioni, consulta the section called “Comandi SQL e PPL supportati”.

Raccomandazioni

Quando si utilizzano interrogazioni dirette nei log, si consiglia quanto segue: CloudWatch

- Quando si cercano più gruppi di log in un'unica query, utilizzare la sintassi appropriata. Per ulteriori informazioni, consulta the section called “Funzioni di gruppo multi-log”.
- Quando usi i comandi SQL o PPL, racchiudi alcuni campi tra i backtick per interrogarli correttamente. I backtick sono necessari per i campi con caratteri speciali (non alfabetici e non numerici). Ad esempio, racchiudi e inserisci i contrassegni. `@message` `Operation.Export, Test::Field` Non è necessario racchiudere colonne con nomi puramente alfabetici nei backtick.

  Query di esempio con campi semplici:

```
SELECT SessionToken, Operation, StartTime  FROM `LogGroup-A`
```

```
LIMIT 1000;
```

Query simile con backtick aggiunti:

```
SELECT `@SessionToken`, `@Operation`, `@StartTime` FROM `LogGroup-A`
LIMIT 1000;
```

- Usa dei limiti per le tue query per assicurarti di non recuperare troppi dati.
- Le query contenenti nomi di campo identici ma che differiscono solo nel caso in cui (ad esempio `field1` e`FIELD1`) non siano supportate.

  Ad esempio, le seguenti query non sono supportate:

  ```
  Select AWSAccountId, AwsAccountId from LogGroup
  Select a.@LogStream, b.@logStream from Table A INNER Join Table B ona.id = b.id
  ```

  Tuttavia, la seguente query è supportata perché il nome del campo (@logStream) è identico in entrambi i gruppi di log:

  ```
  Select a.@logStream, b.@logStream from Table A INNER Join Table B on a.id = b.id
  ```

- Le funzioni e le espressioni devono operare sui nomi dei campi e far parte di un'`SELECT`istruzione con un gruppo di log specificato nella `FROM` clausola.

  Ad esempio, questa query non è supportata:

  ```
  SELECT cos(10) FROM LogGroup
  ```

  Questa interrogazione è supportata:

  ```
  SELECT cos(field1) FROM LogGroup
  ```

Quote

> **Note**
>
> Se stai cercando di eseguire query dirette utilizzando CloudWatch Logs Insights, assicurati di fare riferimento a. the section called “CloudWatch Logs Insights”

Description	Valore	Limite flessibile?	Note
Limite TPS a livello di account per le query dirette APIs	3 TPS	Sì	
Numero massimo di fonti di dati	20	Sì	Il limite è per Account AWS.
Numero massimo di indici o viste materializzate con aggiornamento automatico	30	Sì	Il limite è per fonte di dati.
Numero massimo di interrogazioni simultanee	30	Sì	Il limite è per origine dati e si applica alle query in sospeso o in esecuzione. Include query interattive (ad esempio comandi di recupero dati come`SELECT`) e query di indice (ad esempio, operazioni come/). `CREATE` `ALTER`
Numero massimo di OCU simultanee per query	512	Sì	OpenSearch Unità di calcolo (OCU). Limite basato su 15 executor e 1 driver, ciascuno con 16 vCPU e 32 GB di memoria. Rappresenta la potenza di elaborazione simultanea.

Description	Valore	Limite flessibile?	Note
Tempo massimo di esecuzione delle query in minuti	60	No	Il limite si applica alle query OpenSearch PPL/SQL in Logs Insights. CloudWatch
Periodo di eliminazione delle query obsolete IDs	90 giorni	Sì	Questo è il periodo di tempo dopo il quale il OpenSearch Servizio elimina i metadati delle query dalle voci precedenti. Ad esempio, una chiamata GetDirectQuery o un GetDirectQueryResult errore per le query che risalgono a più di 90 giorni.

Supportato Regioni AWS

Quanto segue Regioni AWS è supportato per le interrogazioni dirette nei CloudWatch registri:

- Asia Pacifico (Mumbai)
- Asia Pacifico (Hong Kong)
- Asia Pacifico (Osaka)
- Asia Pacifico (Seoul)
- Asia Pacifico (Singapore)
- Asia Pacifico (Sydney)
- Asia Pacifico (Tokyo)
- Canada (Centrale)
- Europa (Francoforte)
- Europa (Irlanda)
- Europa (Stoccolma)
- Europa (Milano)
- Europa (Spagna)
- Stati Uniti orientali (Virginia settentrionale)
- Stati Uniti orientali (Ohio)
- Stati Uniti occidentali (Oregon)

- Stati Uniti occidentali (California settentrionale)
- Europa (Parigi)
- Europa (Londra)
- Sud America (San Paolo)

Interrogazione diretta dei dati di Amazon Security Lake nel servizio OpenSearch

Questa sezione ti guiderà attraverso il processo di creazione e configurazione di un'integrazione di fonti di dati in Amazon OpenSearch Service, consentendoti di interrogare e analizzare in modo efficiente i tuoi dati archiviati in Security Lake.

Nelle pagine seguenti, imparerai come configurare un'origine dati con query diretta di Security Lake, esplorare i prerequisiti necessari e seguire le step-by-step procedure utilizzando il. Console di gestione AWS

Argomenti

- Creazione di un'integrazione di origini dati Amazon Security Lake in OpenSearch Service
- Configurazione e interrogazione di un'origine dati Security Lake nelle dashboard OpenSearch
- Prezzi
- Limitazioni
- Raccomandazioni
- Quote
- Supportato Regioni AWS

Creazione di un'integrazione di origini dati Amazon Security Lake in OpenSearch Service

Puoi usare Amazon OpenSearch Serverless per interrogare direttamente i dati di sicurezza in Amazon Security Lake. Per fare ciò, crei una fonte di dati che ti consenta di utilizzare funzionalità OpenSearch zero-ETL sui dati di Security Lake. Quando si crea un'origine dati, è possibile cercare, ottenere informazioni e analizzare direttamente i dati archiviati in Security Lake. È possibile accelerare le prestazioni delle query e utilizzare OpenSearch analisi avanzate su determinati set di dati di Security Lake utilizzando l'indicizzazione su richiesta.

Argomenti

- Prerequisiti
- Procedura
- Fasi successive
- Altre risorse

Prerequisiti

Prima di iniziare, assicurati di aver letto la seguente documentazione:

- the section called “Limitazioni”
- the section called “Raccomandazioni”
- the section called “Quote”

Prima di creare un'origine dati, esegui le seguenti azioni in Security Lake:

- Abilita Security Lake. Configura Security Lake per raccogliere i log sulla Regione AWS stessa OpenSearch risorsa. Per istruzioni, consulta la sezione Guida introduttiva ad Amazon Security Lake nella guida per l'utente di Amazon Security Lake.
- Configura le autorizzazioni di Security Lake. Assicurati di aver accettato le autorizzazioni relative ai ruoli collegati al servizio per la gestione delle risorse e che la console non mostri alcun problema nella pagina Problemi. Per ulteriori informazioni, consulta il ruolo collegato ai servizi per Security Lake nella guida per l'utente di Amazon Security Lake.
- Condividi le fonti di dati di Security Lake. Quando accedi OpenSearch allo stesso account di Security Lake, assicurati che non ci sia alcun messaggio per registrare i tuoi bucket Security Lake con Lake Formation nella console di Security Lake. Per OpenSearch l'accesso su più account, configura un abbonato alle query di Lake Formation nella console Security Lake. Usa l'account associato alla tua OpenSearch risorsa come abbonato. Per ulteriori informazioni, consulta Gestione degli abbonati in Security Lake nella guida per l'utente di Amazon Security Lake.

Inoltre, devi avere a disposizione anche le seguenti risorse: Account AWS

- (Facoltativo) Un ruolo IAM creato manualmente. Puoi utilizzare questo ruolo per gestire l'accesso alla tua fonte di dati. In alternativa, puoi fare in modo che OpenSearch Service crei automaticamente un ruolo per te con le autorizzazioni richieste. Se scegli di utilizzare un ruolo IAM

creato manualmente, segui le indicazioni riportate inthe section called “Autorizzazioni richieste per i ruoli IAM creati manualmente”.

Procedura

È possibile configurare un'origine dati per connettersi a un database di Security Lake dall'interno di Console di gestione AWS.

Per configurare una fonte di dati utilizzando il Console di gestione AWS

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, vai a Gestione centrale e scegli Origini dati connesse.
3. Scegli Connetti.
4. Scegli Security Lake come tipo di origine dati.
5. Scegli Next (Successivo).
6. In Dettagli della connessione dati, inserisci un nome e una descrizione opzionale.
7. In Impostazioni di accesso tramite autorizzazione IAM, scegli come gestire l'accesso alla tua fonte di dati.

    a. Se desideri creare automaticamente un ruolo per questa fonte di dati, segui questi passaggi:

        i. Seleziona Crea un nuovo ruolo.

        ii. Inserire un nome per il ruolo IAM.

        iii. Seleziona una o più AWS Glue tabelle per definire quali dati possono essere interrogati.

    b. Se desideri utilizzare un ruolo esistente che gestisci tu stesso, segui questi passaggi:

        i. Seleziona Utilizza un ruolo esistente.

        ii. Seleziona un ruolo esistente dal menu a discesa.

> **Note**
>
> Quando utilizzi il tuo ruolo, devi assicurarti che disponga di tutte le autorizzazioni necessarie allegando le policy richieste dalla console IAM. Per ulteriori informazioni, consulta the section called “Autorizzazioni richieste per i ruoli IAM creati manualmente”.

8. (Facoltativo) In Tag, aggiungi tag alla tua fonte di dati.
9. Scegli Next (Successivo).
10. In Configurazione OpenSearch, scegli come eseguire la configurazione OpenSearch.

    - Controlla i nomi delle risorse predefiniti e le impostazioni di conservazione dei dati.

      Quando utilizzi le impostazioni predefinite, vengono create automaticamente una nuova OpenSearch applicazione e un'area di lavoro Essentials senza costi aggiuntivi. OpenSearch consente di analizzare più fonti di dati. Include aree di lavoro, che offrono esperienze personalizzate per i casi d'uso più diffusi. Le aree di lavoro supportano il controllo degli accessi, consentendoti di creare spazi privati per i tuoi casi d'uso e condividerli solo con i tuoi collaboratori.

11. Usa impostazioni personalizzate:

    a. Scegliere Customize (Personalizza).

    b. Modifica il nome della raccolta e le impostazioni di conservazione dei dati secondo necessità.

    c. Seleziona l' OpenSearch applicazione e l'area di lavoro che desideri utilizzare.

12. Scegli Next (Successivo).
13. Rivedi le tue scelte e scegli Modifica se devi apportare modifiche.
14. Scegli Connect per configurare l'origine dati. Rimani su questa pagina mentre viene creata la tua fonte di dati. Quando sarà pronta, verrai indirizzato alla pagina dei dettagli dell'origine dati.

Fasi successive

Visita OpenSearch Dashboards e crea una dashboard

Dopo aver creato un'origine dati, OpenSearch Service ti fornisce un URL per i OpenSearch dashboard. Lo usi per interrogare i tuoi dati utilizzando SQL o PPL. L'integrazione con Security Lake include modelli di query preconfezionati per SQL e PPL che consentono di iniziare ad analizzare i log.

Per ulteriori informazioni, consulta the section called “Configurazione di un'origine dati Security Lake”.

Altre risorse

Autorizzazioni richieste per i ruoli IAM creati manualmente

Quando crei un'origine dati, scegli un ruolo IAM per gestire l'accesso ai tuoi dati. Sono disponibili due opzioni:

1. Crea automaticamente un nuovo ruolo IAM
2. Utilizza un ruolo IAM esistente che hai creato manualmente

Se utilizzi un ruolo creato manualmente, devi assegnare le autorizzazioni corrette al ruolo. Le autorizzazioni devono consentire l'accesso alla fonte di dati specifica e consentire al OpenSearch Servizio di assumere il ruolo in modo che il OpenSearch Servizio possa accedere e interagire con i dati in modo sicuro. Inoltre, concedi LakeFormation le autorizzazioni al ruolo per tutti i database e le tabelle su cui desideri eseguire la query. Concedi `DESCRIBE` le autorizzazioni per il ruolo sui SecurityLake database a cui desideri eseguire la query dalla connessione di interrogazione diretta. Concedi almeno `SELECT and DESCRIBE` le autorizzazioni al ruolo di origine dati per le tabelle all'interno del database.

La seguente politica di esempio illustra le autorizzazioni con privilegi minimi necessarie per creare e gestire un'origine dati. Se disponi di autorizzazioni più ampie, come la policy, queste autorizzazioni comprendono le autorizzazioni con il `AdminstratorAccess` privilegio minimo della policy di esempio.

Nella seguente politica di esempio, sostituiscila con le tue informazioni. *`placeholder text`*

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "AmazonOpenSearchDirectQueryServerlessAccess",
            "Effect": "Allow",
            "Action": [
                "aoss:APIAccessAll",
                "aoss:DashboardsAccessAll"
            ],
            "Resource": "arn:aws:aoss:us-
east-1:111122223333:collection/collectionname/*"
```

```
        },
        {
            "Sid": "AmazonOpenSearchDirectQueryGlueAccess",
            "Effect": "Allow",
            "Action": [
                "glue:GetDatabase",
                "glue:GetDatabases",
                "glue:GetPartition",
                "glue:GetPartitions",
                "glue:GetTable",
                "glue:GetTableVersions",
                "glue:GetTables",
                "glue:SearchTables",
                "glue:BatchGetPartition"
            ],
            "Resource": [
                "arn:aws:glue:us-east-1:111122223333:table/databasename/*",
                "arn:aws:glue:us-east-1:111122223333:database/databasename",
                "arn:aws:glue:us-east-1:111122223333:catalog",
                "arn:aws:glue:us-east-1:111122223333:database/default"
            ]
        },
        {
            "Sid": "AmazonOpenSearchDirectQueryLakeFormationAccess",
            "Effect": "Allow",
            "Action": [
                "lakeformation:GetDataAccess"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

Il ruolo deve inoltre avere la seguente politica di fiducia, che specifica l'ID di destinazione.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
```

```
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "directquery.opensearchservice.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

Per istruzioni sulla creazione del ruolo, consulta Creazione di un ruolo utilizzando policy di attendibilità personalizzate.

Per impostazione predefinita, il ruolo ha accesso solo agli indici delle origini dei dati delle query dirette. Sebbene sia possibile configurare il ruolo per limitare o concedere l'accesso all'origine dati, si consiglia di non modificare l'accesso di questo ruolo. Se elimini l'origine dati, questo ruolo verrà eliminato. Ciò rimuoverà l'accesso per tutti gli altri utenti se sono mappati al ruolo.

Interrogazione dei dati di Security Lake crittografati con una chiave gestita dal cliente

Se il bucket Security Lake associato alla connessione dati è crittografato utilizzando la crittografia lato server con gestione dal cliente AWS KMS key, è necessario aggiungere il ruolo di LakeFormation servizio alla policy chiave. Ciò consente al servizio di accedere e leggere i dati per le tue domande.

Nel seguente esempio di politica, sostituisci le tue informazioni *placeholder text* con le tue.

```
{
    "Sid": "Allow LakeFormation to access the key",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::account:role/aws-service-role/lakeformation.amazonaws.com/
AWSServiceRoleForLakeFormationDataAccess"
    },
    "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey"
    ],
    "Resource": "*"
```

```
}
```

Configurazione e interrogazione di un'origine dati Security Lake nelle dashboard OpenSearch

Ora che hai creato la tua fonte di dati, puoi configurarla nelle dashboard. OpenSearch

Questa sezione illustra vari casi d'uso con la tua fonte di dati nelle OpenSearch dashboard prima di interrogare i dati. Per iniziare, devi accedere alla tua fonte di dati in OpenSearch Dashboards. Nel menu a sinistra, in Gestione, scegli Origini dati. Quindi, seleziona il nome dell'origine dati che hai creato in precedenza nella console di OpenSearch servizio.

Interroga le tabelle di Security Lake da Discover

Se hai creato tabelle in base ai log di Security Lake, ora puoi interrogare tali tabelle direttamente da OpenSearch Discover. Ciò consente di accedere e analizzare senza problemi i dati archiviati in Security Lake, direttamente dalla familiare interfaccia Discover. Interrogando Security Lake direttamente da Discover, puoi evitare la necessità di estrarre, trasformare e caricare manualmente i dati in un indice di ricerca separato. Per iniziare rapidamente ad analizzare i log, Discover include una serie di query salvate in PPL e SQL.

Inizia selezionando l'origine dati che hai configurato. Seleziona il database e la tabella associati su cui desideri interrogare, quindi usa la barra di ricerca per scrivere query sulle tue tabelle. Per capire quali istruzioni, comandi e limitazioni sono supportati per l'integrazione di Security Lake, consultathe section called “Comandi SQL e PPL supportati”.

Per sfruttare le query predefinite disponibili per Security Lake, vai a... in alto a destra di Discover, scegli Apri query, quindi scegli Modelli. Sono disponibili molte query predefinite per le fonti di registro supportate in Security Lake. Cerca i modelli che corrispondono al tuo caso d'uso, copia la query da utilizzare nella barra di ricerca e sostituisci i campi basati su modelli (come Regione e azione) con le tue informazioni.

Accelera i dati da Discover

Per migliorare le prestazioni e consentire query e analisi successive più rapide OpenSearch, puoi inserire i risultati della query da Discover in una vista OpenSearch indicizzata.

Per creare una vista indicizzata

1. Da Discover, scegli Crea vista indicizzata.

2. Nell'editor di query, inserisci la query desiderata. Puoi creare una nuova interrogazione qui o usarne una esistente dalle tue ricerche precedenti.
3. Specificate un nome per la nuova vista indicizzata. Scegliete un nome descrittivo che vi aiuti a identificare la vista in un secondo momento.
4. Configura le impostazioni di conservazione dei dati per la visualizzazione indicizzata. È possibile specificare per quanto tempo i dati devono essere conservati nell'indice, in modo da bilanciare le prestazioni con i costi di archiviazione.
5. Crea la vista indicizzata. Dopo la creazione, la vista indicizzata sarà disponibile per interrogazioni e analisi più rapide.

Se hai già creato viste indicizzate, puoi accedervi da Discover.

Per utilizzare una visualizzazione indice esistente

1. Da Discover, scegli Seleziona visualizzazione indicizzata per visualizzare un elenco delle viste indicizzate esistenti per Security Lake.
2. Scegli la vista indicizzata che desideri utilizzare. Ciò applicherà la visualizzazione alla query corrente, velocizzando potenzialmente in modo significativo il recupero e l'analisi dei dati.

Crea una visualizzazione del dashboard per la tua fonte di dati

Quando utilizzi OpenSearch Service, puoi analizzare i tipi di AWS log più diffusi utilizzando modelli di dashboard predefiniti. Per Security Lake sono disponibili modelli per i log VPC e WAF. CloudTrail Questi modelli consentono di creare una dashboard personalizzata in base ai dati specifici. Includono query e dashboard predefiniti personalizzati per quel tipo di registro specifico. Ciò consente di iniziare rapidamente ad analizzare queste fonti di AWS log popolari, senza dover creare tutto da zero.

> **Note**
>
> Le dashboard utilizzano viste indicizzate, che acquisiscono dati da Security Lake e contribuiscono al calcolo diretto delle interrogazioni e della raccolta.

Segui questi passaggi per creare una dashboard utilizzando uno di questi modelli predefiniti, in modo da poter iniziare subito a esplorare e analizzare i tuoi dati.

Per creare una visualizzazione del pannello di controllo

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Dal riquadro di navigazione a sinistra, scegli Gestione centrale, quindi Origini dati connesse.
3. Seleziona l'origine dati per aprire la pagina dei dettagli.
4. Seleziona Crea pannello di controllo.
5. Scegli il tipo di dashboard che desideri creare.
6. Inserisci un nome per la dashboard.
7. Inserisci una descrizione opzionale per la dashboard.
8. Seleziona una o più tabelle AWS Glue da visualizzare sulla dashboard.
9. Scegli la frequenza con cui aggiornare i dati nella dashboard.
10. Scegli l' OpenSearch area di lavoro che desideri utilizzare.
    a. Per creare un nuovo spazio di lavoro, seleziona Crea nuovo spazio di lavoro.
    b. Per utilizzare un'area di lavoro esistente, seleziona Seleziona un'area di lavoro esistente.
11. Inserisci un nome per il tuo spazio di lavoro.
12. Seleziona Crea pannello di controllo.

Risoluzione dei problemi

In alcuni casi i risultati non vengono restituiti come previsto. In caso di problemi, assicurati di seguire ilthe section called “Raccomandazioni”.

Prezzi

Amazon OpenSearch Service offre prezzi OpenSearch Compute Unit (OCU) per le query dirette di Security Lake. Quando esegui le query dirette, ti vengono addebitati addebiti orari, indicati nella fattura OCUs come tipo di utilizzo dell' DirectQuery OCU. Ti verranno inoltre addebitati addebiti separati da Amazon Security Lake.

Le interrogazioni dirette sono di due tipi: interattive e con visualizzazione indicizzata.

- Le query interattive vengono utilizzate per compilare il selettore di dati ed eseguire analisi sui dati in Security Lake. OpenSearch Il servizio gestisce ogni query con un job separato pre-riscaldato, senza mantenere una sessione prolungata.

- Le query di visualizzazione indicizzata utilizzano il calcolo per mantenere le viste indicizzate nel Servizio. OpenSearch Queste query in genere richiedono più tempo perché inseriscono una quantità variabile di dati in un indice denominato. Per le sorgenti dati connesse a Security Lake, i dati indicizzati vengono archiviati in una raccolta OpenSearch Serverless in cui vengono addebitati i dati indicizzati (IndexingOCU), i dati ricercati (SearchOCU) e i dati archiviati in GB.

Per ulteriori informazioni, consulta le sezioni Direct Query e Serverless all'interno dei prezzi OpenSearch di Amazon Service.

Limitazioni

Le seguenti limitazioni si applicano alle query dirette in Security Lake:

- L'integrazione diretta delle query con Security Lake è disponibile solo nelle raccolte OpenSearch di servizi e nell'interfaccia OpenSearch utente.
- OpenSearch Le raccolte serverless hanno limiti di payload in rete di 100 MiB.
- La gestione delle tabelle per Security Lake viene eseguita in Lake Formation.
- Security Lake supporta solo le viste materializzate come viste indicizzate. Gli indici di copertura non sono supportati.
- AWS CloudFormation i modelli non sono ancora supportati.
- OpenSearch Le istruzioni SQL e OpenSearch PPL presentano limitazioni diverse quando si lavora con OpenSearch gli indici rispetto all'utilizzo di query dirette. Direct Query supporta comandi avanzati come JOINs sottoquery e ricerche, mentre il supporto per questi comandi sugli OpenSearch indici è limitato o inesistente. Per ulteriori informazioni, consulta the section called “Comandi SQL e PPL supportati”.

Raccomandazioni

Per l'utilizzo di query dirette in Security Lake, consigliamo quanto segue:

- Controlla lo stato del tuo Security Lake e assicurati che funzioni senza problemi. Per una procedura dettagliata di risoluzione dei problemi, consulta Risoluzione dei problemi dello stato del data lake nella Guida per l'utente di Amazon Security Lake.
- Verifica l'accesso alla tua query:
  - Se stai interrogando Security Lake da un account diverso dall'account amministratore delegato di Security Lake, configura un abbonato con accesso alle query in Security Lake.

  - Se stai interrogando Security Lake dallo stesso account, controlla la presenza di messaggi in Security Lake sulla registrazione dei bucket S3 gestiti con. LakeFormation
- Esplora i modelli di query e le dashboard predefinite per iniziare subito l'analisi.
- Acquisisci familiarità con Open Cybersecurity Schema Framework (OCSF) e Security Lake:
  - Esamina gli esempi di mappatura degli schemi per le fonti nel repository OCSF AWS GitHub
  - Scopri come interrogare Security Lake in modo efficace visitando Security Lake queries for AWS source version 2 (OCSF 1.1.0)
  - Migliora le prestazioni delle query utilizzando le partizioni:,, e `accountid` `region` `time_dt`
- Prendi confidenza con la sintassi SQL, supportata da Security Lake per le interrogazioni. Per ulteriori informazioni, consulta the section called “Comandi SQL supportati”.
- Usa dei limiti per le tue query per assicurarti di non recuperare troppi dati.

Quote

Description	Valore	Limite flessibile?	Note
Limite TPS a livello di account per le query dirette APIs	3 TPS	Sì	
Numero massimo di fonti di dati	20	Sì	Il limite è per Account AWS.
Numero massimo di indici o viste materializzate con aggiornamento automatico	30	Sì	Il limite si applica per fonte di dati. Include solo indici e viste materializzate (MVs) con aggiornamento automatico impostato su true.
Numero massimo di interrogazioni simultanee	30	Sì	Il limite si applica alle query in sospeso o in esecuzione. Include query interattive (ad esempio comandi per il recupero dei dati come`SELECT`) e query sull'indi

Description	Valore	Limite flessibile?	Note
			ce (ad esempio, operazioni come//). `CREATE` `ALTER` `DROP`
Numero massimo di OCU simultanee per query	512	Sì	OpenSearch Unità di calcolo (OCU). Limite basato su 15 executor e 1 driver, ciascuno con 16 vCPU e 32 GB di memoria. Rappresenta la potenza di elaborazione simultanea.
Tempo massimo di esecuzione delle query in minuti	30	No	Si applica solo alle query interattive (ad esempio, comandi di recupero dati come). `SELECT` Per le `REFRESH` interrogazioni, il limite è di 6 ore.
Periodo di eliminazione delle interrogazioni obsolete IDs	90 giorni	Sì	Questo è il periodo di tempo dopo il quale il OpenSearch Servizio elimina i metadati delle query dalle voci precedenti. Ad esempio, una chiamata GetDirectQuery o un GetDirectQueryResult errore per le query che risalgono a più di 90 giorni.

Supportato Regioni AWS

Le seguenti opzioni Regioni AWS sono supportate per le interrogazioni dirette in Security Lake:

- Asia Pacifico (Mumbai)
- Asia Pacifico (Singapore)
- Asia Pacifico (Sydney)
- Asia Pacifico (Tokyo)
- Canada (Centrale)
- Europa (Francoforte)
- Europa (Irlanda)
- Europa (Stoccolma)
- Stati Uniti orientali (Virginia settentrionale)

- Stati Uniti orientali (Ohio)
- Stati Uniti occidentali (Oregon)
- Europa (Parigi)
- Europa (Londra)
- Sud America (San Paolo)

Interrogazioni dirette in Amazon Managed Service per Prometheus

Puoi utilizzare Amazon OpenSearch Service per interrogare direttamente i parametri operativi archiviati in Amazon Managed Service for Prometheus. Questa integrazione consente di analizzare e visualizzare i dati delle serie temporali di Prometheus insieme ai log e alle tracce OpenSearch all'interno dell'interfaccia utente, abilitando un'esperienza di osservabilità unificata.

A differenza delle query dirette basate sullo storage (come Amazon S3 o Logs CloudWatch), l'integrazione con Prometheus utilizza un'architettura di chiamate live. OpenSearch Service funge da cliente diretto, traduce le tue domande ed effettua chiamate API in tempo reale verso il tuo spazio di lavoro Prometheus. Poiché il OpenSearch Servizio non fornisce elaborazione temporanea per la scansione dei dati, non sono previsti costi di Compute Unit (OCU) per OpenSearch queste query.

Per configurare e utilizzare questa integrazione, devi prima creare l'origine dati e quindi configurare le aree di lavoro per interrogare i dati.

Argomenti

- Creazione di un'origine dati Amazon Managed Service per Prometheus
- Interrogazione delle metriche di Prometheus
- Prezzi
- Limitazioni
- Raccomandazioni
- Quote
- Supportato Regioni AWS

Creazione di un'origine dati Amazon Managed Service per Prometheus

Per creare un'origine dati Amazon Managed Service for Prometheus, hai bisogno di uno spazio di lavoro attivo e di un ruolo IAM che OpenSearch conceda a Service le autorizzazioni necessarie per interrogare le tue metriche.

Prerequisiti

Prima di connettere la fonte di dati, assicurati di disporre di quanto segue:

- Prometheus workspace: un Amazon Managed Service attivo per l'area di lavoro Prometheus. Annota il tuo ID Workspace e il suo indirizzo di residenza. Regione AWS
- Ruolo IAM: AWS Identity and Access Management ruolo con una politica di fiducia che consente al responsabile del `directquery.opensearchservice.amazonaws.com` servizio di assumerlo.

Connessione della fonte di dati

Una volta soddisfatti i prerequisiti, puoi connettere l'origine dati utilizzando la console OpenSearch di servizio.

Per configurare un'origine dati Amazon Managed Service for Prometheus

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, vai a Gestione centrale e scegli Origini dati connesse.
3. Scegli Connect new data source.
4. Scegli Amazon Managed Service for Prometheus come tipo di origine dati.
5. Scegli Next (Successivo).
6. In Dettagli della connessione dati, inserisci un nome e una descrizione opzionale.
7. In Ruoli IAM, scegli come gestire l'accesso:
    - Per creare automaticamente un ruolo per questa fonte di dati:
        a. Seleziona Crea un nuovo ruolo.
        b. Inserire un nome per il ruolo IAM.
        c. Seleziona uno o più spazi di lavoro per definire quali dati possono essere interrogati.
    - Per utilizzare un ruolo esistente che gestisci tu stesso:

a. Seleziona Utilizza un ruolo esistente.

b. Seleziona un ruolo esistente dal menu a discesa.

> **Note**
>
> Quando usi il tuo ruolo, assicurati che disponga di tutte le autorizzazioni necessarie allegando le policy richieste dalla console IAM. Per ulteriori informazioni, consulta the section called “Autorizzazioni richieste per i ruoli IAM creati manualmente”.

8. (Facoltativo) In Politica di accesso, configura una politica di accesso per l'origine dati. Le politiche di accesso controllano se una richiesta all'origine dati OpenSearch di Service Direct Query viene accettata o rifiutata. Se non configuri una politica di accesso, solo il proprietario dell'origine dati può accedervi. Puoi configurare la politica di accesso per abilitare l'accesso tra più account, consentendo ai responsabili di altri di accedere Account AWS alla fonte di dati.

   È possibile creare una politica di accesso utilizzando l'editor visivo o fornendo un documento di policy JSON. Con l'editor visivo, puoi consentire o negare l'accesso specificando un Account AWS ID principale, l'ARN dell'account, l'ARN dell'utente IAM, l'ARN del ruolo IAM, l'indirizzo IP di origine o il blocco CIDR. L'editor visivo supporta fino a 10 elementi. Per definire una policy con più di 10 elementi, usa l'editor JSON.

   Puoi anche scegliere Importa policy per importare una politica di accesso esistente da un'altra fonte di dati.

9. (Facoltativo) In Tag, aggiungi tag alla tua fonte di dati.
10. Scegli Next (Successivo).
11. In Configurazione OpenSearch, scegli come configurare l' OpenSearch interfaccia utente:

    a. Se nel tuo account non esiste alcuna applicazione OpenSearch UI, crea una nuova OpenSearch applicazione. Se esiste un' OpenSearch applicazione esistente, selezionala.

    b. Se crei una nuova applicazione, crea un nuovo spazio di lavoro di osservabilità. Se hai selezionato un'applicazione esistente, crea un nuovo spazio di lavoro di osservabilità o selezionane uno esistente. Amazon Managed Service for Prometheus è disponibile solo nell'area di lavoro di osservabilità.

12. Scegli Next (Successivo).
13. Rivedi le tue scelte e scegli Modifica se devi apportare modifiche.

14. Scegli Connect per configurare l'origine dati. Rimani su questa pagina mentre viene creata la tua fonte di dati. Quando è pronta, verrai indirizzato alla pagina dei dettagli dell'origine dati.

Fasi successive

Visita l' OpenSearch interfaccia utente

Dopo aver creato un'origine dati, OpenSearch Service ti fornisce un URL dell'applicazione OpenSearch UI. Lo usi per configurare chi ha accesso all' OpenSearch interfaccia utente e analizzare i dati di Amazon Managed Service for Prometheus utilizzando Discover Metrics with PromQL.

Risorse aggiuntive

Autorizzazioni richieste per i ruoli IAM creati manualmente

Quando crei un'origine dati, scegli un ruolo IAM per gestire l'accesso ai tuoi dati. Sono disponibili due opzioni:

- Crea automaticamente un nuovo ruolo IAM
- Utilizza un ruolo IAM esistente che hai creato manualmente

Se utilizzi un ruolo creato manualmente, devi assegnare le autorizzazioni corrette al ruolo. Le autorizzazioni devono consentire l'accesso alla fonte di dati specifica e consentire a OpenSearch Service di assumere il ruolo. Ciò è necessario per consentire OpenSearch al Servizio di accedere e interagire con i dati in modo sicuro.

La seguente politica di esempio illustra le autorizzazioni con privilegi minimi necessarie per creare e gestire un'origine dati. Se disponi di autorizzazioni più ampie, ad esempio `aps:*` o la policy, queste autorizzazioni comprendono le autorizzazioni con il `AdministratorAccess` privilegio minimo della politica di esempio.

Nella seguente politica di esempio, sostituisci il testo con le tue informazioni. *`placeholder`*

Policy IAM di esempio

Assegna le seguenti autorizzazioni al tuo ruolo IAM per consentire al OpenSearch servizio di recuperare i metadati delle metriche ed eseguire le query:

```
{
    "Version": "2012-10-17",
    "Statement": [
```

```
        {
            "Sid": "AmazonOpenSearchDirectQueryPrometheusAccess",
            "Effect": "Allow",
            "Action": [
                "aps:DeleteAlertManagerSilence",
                "aps:GetAlertManagerSilence",
                "aps:GetAlertManagerStatus",
                "aps:GetLabels",
                "aps:GetMetricMetadata",
                "aps:GetSeries",
                "aps:ListAlertManagerAlertGroups",
                "aps:ListAlertManagerAlerts",
                "aps:ListAlertManagerReceivers",
                "aps:ListAlertManagerSilences",
                "aps:ListAlerts",
                "aps:QueryMetrics",
                "aps:PutAlertManagerSilences",
                "aps:DescribeAlertManagerDefinition",
                "aps:CreateRuleGroupsNamespace",
                "aps:DeleteRuleGroupsNamespace",
                "aps:ListRuleGroupsNamespaces",
                "aps:DescribeRuleGroupsNamespace",
                "aps:PutRuleGroupsNamespace"
            ],
            "Resource": "arn:aws:aps:region:account-id:workspace/workspace-id",
            "Condition": {
                "ForAnyValue:StringEquals": {
                    "aws:CalledVia": [
                        "directquery.opensearchservice.amazonaws.com"
                    ]
                }
            }
        },
        {
            "Sid": "AmazonOpenSearchDirectQueryPrometheusListAccess",
            "Effect": "Allow",
            "Action": [
                "aps:ListWorkspaces"
            ],
            "Resource": "*",
            "Condition": {
                "ForAnyValue:StringEquals": {
                    "aws:CalledVia": [
                        "directquery.opensearchservice.amazonaws.com"
```

```
                    ]
                }
            }
        }
    ]
}
```

Esempio di politica di fiducia

Allega la seguente politica di fiducia al tuo ruolo IAM:

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "TrustPolicyForAmazonOpenSearchDirectQueryService",
            "Effect": "Allow",
            "Principal": {
                "Service": "directquery.opensearchservice.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "ArnEquals": {
                    "aws:SourceArn": "arn:aws:opensearch:region:account-
id:datasource/data-source-name"
                },
                "StringEquals": {
                    "aws:SourceAccount": "account-id"
                }
            }
        }
    ]
}
```

Interrogazione delle metriche di Prometheus

Amazon OpenSearch Service ti consente di interrogare i tuoi dati Prometheus utilizzando PromQL (Prometheus Query Language) direttamente dall'interfaccia Observability. Quando esegui una query PromQL sull'origine dati Prometheus OpenSearch , Service passa la query direttamente all'API del tuo spazio di lavoro per l'esecuzione.

Esecuzione di una query PromQL

Per eseguire una query:

1. Apri l'applicazione OpenSearch dell'interfaccia utente e l'area di lavoro di osservabilità.
2. Vai su Observability e seleziona Discover Metrics.
3. Nel menu a discesa delle origini dati, seleziona la tua fonte dati Prometheus.
4. Inserisci la tua query PromQL nella barra delle interrogazioni.

Ad esempio, per calcolare l'utilizzo medio della CPU al secondo in una finestra di 5 minuti per un pod specifico:

```
avg(rate(container_cpu_usage_seconds_total{pod="payment-service-pod"}[5m])) by (pod)
```

> **Note**
>
> Imposta il time picker su una finestra ristretta e pertinente (ad esempio, l'ultima ora) per ottimizzare le prestazioni dell'API ed evitare i timeout.

Visualizzazione delle metriche nelle dashboard

Puoi aggiungere visualizzazioni metriche basate su ProMQL ai dashboard di osservabilità esistenti per correlarle ai dati di log e traccia.

1. Vai a Discover Metrics, seleziona l'area di lavoro Prometheus dal menu a discesa delle fonti di dati ed esegui la query PromQL.
2. Utilizza la scheda di visualizzazione in Discover Metrics per creare una visualizzazione e definire il tipo di visualizzazione.
3. Salva la visualizzazione nella dashboard.

> **Note**
>
> Le visualizzazioni delle metriche possono essere aggiunte solo da Discover Metrics. Le visualizzazioni disponibili nella scheda Visualizzazioni sono ottimizzate solo per i log.

Prezzi

Per Amazon Managed Service for Prometheus OpenSearch , Service effettua chiamate in tempo reale direttamente all'area di lavoro Prometheus per recuperare i dati. Poiché OpenSearch Service non fornisce risorse di elaborazione separate per eseguire queste query, non sono previsti costi per la Compute Unit (OCU). OpenSearch Sei responsabile solo dei costi di interrogazione standard associati ad Amazon Managed Service for Prometheus.

Per ulteriori informazioni, consulta le sezioni Direct Query e Serverless all'interno dei prezzi OpenSearch di Amazon Service.

Limitazioni

Le seguenti limitazioni si applicano alle query dirette in Amazon Managed Service for Prometheus:

- Limiti relativi all'intervallo di tempo: le query in tempo reale sono ottimizzate per i dashboard operativi. L'interrogazione di metriche altamente granulari e non sottoposte a downsampling su orizzonti temporali molto lunghi (ad esempio, su più mesi) potrebbe raggiungere i limiti di dimensione del payload o causare timeout dall'API Prometheus.
- Quote API Prometheus: le tue domande sono soggette alle quote di servizio standard di Amazon Managed Service per Prometheus, inclusi i limiti per l'esecuzione simultanea di query Samples Processed (QSP) e l'esecuzione simultanea delle query.
- Timeout delle query: il timeout delle query a esecuzione prolungata scade dopo 30 secondi.

Raccomandazioni

Per l'utilizzo di query dirette in Amazon Managed Service for Prometheus, consigliamo quanto segue:

- Usa regole di registrazione per lunghi intervalli di tempo: poiché le query dirette effettuano chiamate API in tempo reale al tuo spazio di lavoro, l'esecuzione di query su dati altamente granulari per lunghi periodi (ad esempio più mesi) può comportare timeout delle API o limiti di payload. Usa le regole di registrazione di Amazon Managed Service for Prometheus per creare metriche sottocampionate per l'analisi storica.
- Applica filtri temporali ristretti: specifica sempre un intervallo di tempo mirato nelle query OpenSearch dell'interfaccia utente per ridurre al minimo il volume di dati recuperati dinamicamente dall'API Prometheus.

- Monitora le tue quote Prometheus: poiché queste query in tempo reale non comportano costi di Compute Unit (OCU), monitora OpenSearch invece il tuo Amazon Managed Service per l'utilizzo di Prometheus. Monitora i limiti delle query Query Samples Processed (QSP) e delle query simultanee per evitare limitazioni a livello di area di lavoro.

Quote

A differenza di altre fonti di dati di interrogazione diretta, ogni volta che avvii una query su Amazon Managed Service for Prometheus OpenSearch , Service effettua una chiamata in tempo reale ad Amazon Managed Service for Prometheus. Non viene creata alcuna sessione. Le seguenti quote sono applicabili per account e regione. Ad esempio, a un account vengono assegnate 20 fonti di dati in us-east-1 e 20 in us-east-2.

Description	Massimo	Può sovrascrivere
Origini dati	20	Sì
Esegui query: query istantanee e a intervallo (TPS)	500	Sì
Leggi le risorse: etichette, metriche, avvisi, regole, gestore degli avvisi (TPS)	50	Sì
Risorse di scrittura: create/update silence, create/update rule (TPS)	50	Sì
Tempo massimo di esecuzione delle query (secondi)	30	No

Supportato Regioni AWS

Quanto segue è supportato per Regioni AWS le query dirette in Amazon Managed Service for Prometheus:

- Asia Pacifico (Hong Kong)
- Asia Pacific (Mumbai)
- Asia Pacifico (Osaka)

- Asia Pacifico (Seoul)
- Asia Pacifico (Singapore)
- Asia Pacifico (Sydney)
- Asia Pacifico (Tokyo)
- Canada (Centrale)
- Europa (Francoforte)
- Europa (Irlanda)
- Europa (Londra)
- Europa (Milano)
- Europa (Parigi)
- Europa (Spagna)
- Europa (Stoccolma)
- Sud America (San Paolo)
- Stati Uniti orientali (Virginia settentrionale)
- Stati Uniti orientali (Ohio)
- Stati Uniti occidentali (California settentrionale)
- Stati Uniti occidentali (Oregon)

Gestione di un'origine dati in Amazon OpenSearch Service

La gestione della fonte di dati è una parte importante per mantenere l'affidabilità, la disponibilità e le prestazioni delle fonti di dati con query dirette e delle altre AWS soluzioni. AWS fornisce i seguenti strumenti per monitorare, segnalare quando qualcosa non va e intraprendere azioni automatiche quando necessario.

Argomenti

- Monitoraggio con CloudWatch fonti di dati metriche
- Abilitazione e disabilitazione delle fonti di dati
- Monitoraggio con budget AWS
- Eliminazione di un'origine dati

Monitoraggio con CloudWatch fonti di dati metriche

È possibile monitorare le interrogazioni dirette utilizzando CloudWatch. CloudWatch raccoglie dati grezzi e li elabora in metriche leggibili e quasi in tempo reale. Queste statistiche vengono conservate per un periodo di 15 mesi, per permettere l'accesso alle informazioni storiche e offrire una prospettiva migliore sulle prestazioni del servizio o dell'applicazione web.

Puoi anche impostare allarmi per monitorare determinate soglie e inviare notifiche o intraprendere azioni quando tali soglie vengono raggiunte. Per ulteriori informazioni, consulta What is Amazon CloudWatch.

Amazon S3 riporta i seguenti parametri:

Metrica	Description
AsyncQueryCreateAPI	Il numero totale di richieste inviate all'API per la creazione di query asincrone. Statistiche rilevanti: Average (Media), Maximum (Massimo) , Sum (Somma) Dimensioni:, `ClientId` `DomainName` Frequenza: 60 secondi
AsyncQueryGetApiRequestCount	Il numero totale di richieste inviate all'API per il recupero dei risultati delle query asincrone. Statistiche rilevanti: Average (Media), Maximum (Massimo) , Sum (Somma) Dimensioni:, `ClientId` `DomainName` Frequenza: 60 secondi
AsyncQueryCancelApiRequestC ount	Il numero totale di richieste inviate all'API per l'annulla mento delle query asincrone. Statistiche rilevanti: Average (Media), Maximum (Massimo) , Sum (Somma)

Metrica	Description
	Dimensioni:, `ClientId` `DomainName` Frequenza: 60 secondi
AsyncQueryGetApiFailedRequestCusErrCount	Il numero di richieste non riuscite durante il recupero dei risultati delle query asincrone a causa di errori relativi al cliente (ad esempio, ID di query non valido). Statistiche rilevanti: Average (Media), Maximum (Massimo), Sum (Somma) Dimensioni:, `ClientId` `DomainName` Frequenza: 60 secondi
AsyncQueryCancelApiFailedRequestCusErrCount	Il numero di richieste non riuscite durante il recupero dei risultati delle query asincrone a causa di errori relativi al cliente (ad esempio, ID di query non valido). Statistiche rilevanti: Average (Media), Maximum (Massimo), Sum (Somma) Dimensioni:, `ClientId` `DomainName` Frequenza: 60 secondi
AsyncQueryCancelApiFailedRequestSysErrCount	Il numero di richieste non riuscite durante la creazione di query asincrone a causa di errori relativi al cliente. Statistiche rilevanti: Average (Media), Maximum (Massimo), Sum (Somma) Dimensioni:, `ClientId` `DomainName` Frequenza: 60 secondi

Metrica	Description
AsyncQueryGetApiFailedRequestSysErrCount	Il numero di richieste non riuscite durante il recupero dei risultati delle query asincrone a causa di errori relativi al sistema. Statistiche rilevanti: Average (Media), Maximum (Massimo), Sum (Somma) Dimensioni:, `ClientId` `DomainName` Frequenza: 60 secondi

CloudWatch Logs e Security Lake riportano le seguenti metriche:

Metrica	Description
DirectQueryRate	La frequenza delle richieste effettuate rispetto alle fonti di dati. Statistiche pertinenti: somma, massimo, minimo, media Dimensioni: `DataSourceName` Frequenza: 60 secondi
DirectQueryLatency	La latenza osservata per l'esecuzione di query sulle fonti di dati. Statistiche pertinenti: media, P90, P99, somma, minima, massima Dimensioni: `DataSourceName` Frequenza: 60 secondi
FailedDirectQueries	Il numero totale di errori di query osservati nelle query relative all'origine dati. Statistiche pertinenti: somma, massimo, minimo, media

Metrica	Description
	Dimensioni: `DataSourceName` Frequenza: 60 secondi
DirectQueryConsumedOCU	Il numero di OCUs quelli utilizzati per eseguire le query sulle fonti di dati. Statistiche pertinenti: media, P90, P99, somma, minima, massima Dimensioni: `DataSourceName` Frequenza: 60 secondi

Abilitazione e disabilitazione delle fonti di dati

> **Note**
>
> Le seguenti informazioni sono applicabili solo alle fonti di dati Amazon S3.

Nei casi in cui desideri interrompere l'utilizzo diretto delle query per un'origine dati, puoi scegliere di disabilitare l'origine dati. La disabilitazione di un'origine dati terminerà l'esecuzione delle query esistenti e interromperà l'esecuzione di tutte le nuove query.

L'impostazione delle accelerazioni per migliorare le prestazioni delle query, ad esempio saltare gli indici, visualizzare le viste materializzate e coprire gli indici, verrà impostata su manuale una volta disattivata l'origine dati. Una volta che un'origine dati è impostata come attiva dopo essere stata disabilitata, le query degli utenti verranno eseguite come previsto. Le accelerazioni precedentemente configurate e impostate come manuali dovranno essere configurate manualmente per essere nuovamente eseguite secondo una pianificazione.

Monitoraggio con budget AWS

Amazon OpenSearch Service sta inserendo i dati sull'utilizzo dell'OCU a livello di account nel Cost Explorer di Billing and Cost Management. Puoi tenere conto dell'utilizzo dell'OCU a livello di account e impostare soglie e avvisi quando le soglie vengono superate.

Il formato del tipo di utilizzo su cui filtrare in Cost Explorer è simile a RegionCode: DirectQuery OCU (OCU-Hours). Se desideri ricevere una notifica quando l'utilizzo di DirectQuery OCU (OCU-Hours) raggiunge la tua soglia, puoi creare un account AWS Budgets e configurare un avviso in base alla soglia impostata. Facoltativamente, per Amazon S3, puoi impostare un argomento Amazon SNS, che disattiverà un'origine dati nel caso in cui venga soddisfatto un criterio di soglia.

> **Note**
>
> I dati di utilizzo in AWS Budgets non sono in tempo reale e possono subire ritardi fino a 8 ore.

Eliminazione di un'origine dati

Quando elimini un'origine dati, Amazon OpenSearch Service la rimuove dal tuo dominio o dalla tua raccolta. OpenSearch Il servizio rimuove anche gli indici associati alla fonte di dati. I tuoi dati transazionali non vengono eliminati dall'altro Servizio AWS, ma l'altro Servizio AWS non invia nuovi dati al Servizio. OpenSearch

Puoi eliminare l'integrazione di un'origine dati utilizzando l'API Console di gestione AWS o il OpenSearch Service.

Console di gestione AWS

Per eliminare un'origine dati Amazon S3

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Dal riquadro di navigazione a sinistra, scegli Domini.
3. Seleziona il dominio per il quale desideri eliminare un'origine dati. Si apre la pagina dei dettagli del dominio. Scegli la scheda Connessioni sotto le informazioni generali e trova la sezione Direct Query.
4. Seleziona l'origine dati che desideri eliminare, scegli Elimina e conferma l'eliminazione.

Per eliminare un'origine dati CloudWatch Logs o Security Lake

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.

2. Dal riquadro di navigazione a sinistra, scegli Gestione centrale, quindi Origini dati connesse.
3. Seleziona l'origine dati che desideri eliminare, scegli Elimina e conferma l'eliminazione.

OpenSearch API di servizio

Per eliminare un'origine dati Amazon S3, utilizza l'operazione DeleteDataSourceAPI.

```
POST https://es.region.amazonaws.com/2021-01-01/opensearch/domain/domain-name/
dataSource/data-source-name
```

Per eliminare un'origine dati CloudWatch Logs o Security Lake, utilizza l'operazione DeleteDirectQueryDataSourceAPI.

Comandi SQL e PPL supportati

OpenSearch SQL e OpenSearch Pipeline Processing Language (PPL) sono linguaggi per l'interrogazione, l'analisi e l'elaborazione dei dati in OpenSearch Logs Insights e Security Lake. CloudWatch Puoi usare OpenSearch SQL e OpenSearch PPL in OpenSearch Discover per interrogare i dati all'interno di CloudWatch Logs, Amazon S3 o Security Lake. CloudWatch Logs Insights supporta anche i linguaggi di query OpenSearch PPL e OpenSearch SQL, oltre a Logs Insights QL, un linguaggio di query creato appositamente per l'analisi dei log. CloudWatch

- OpenSearch SQL: OpenSearch SQL offre un'opzione familiare se sei abituato a lavorare con database relazionali. OpenSearch SQL offre un sottoinsieme di funzionalità SQL, il che lo rende una buona scelta per eseguire query ad hoc e attività di analisi dei dati. Con OpenSearch SQL, è possibile utilizzare comandi come SELECT, FROM, WHERE, GROUP BY, HAVING e vari altri comandi e funzioni SQL disponibili in SQL. È possibile eseguire JOINs più tabelle (o gruppi di log), correlare i dati tra tabelle (o gruppi di log) utilizzando sottoquery e utilizzare il ricco set di funzioni SQL JSON, matematiche, di stringa, condizionali e di altro tipo per eseguire analisi approfondite sui dati di registro e di sicurezza.
- OpenSearch PPL (Piped Processing Language): con OpenSearch PPL, è possibile recuperare, interrogare e analizzare i dati utilizzando comandi concatenati, semplificando la comprensione e la composizione di query complesse. La sua sintassi è basata sulle pipe Unix e consente il concatenamento di comandi per trasformare ed elaborare i dati. Con PPL, è possibile filtrare e aggregare i dati e utilizzare comandi come JOINs sottoquery, LOOKUP e un ricco set di funzioni matematiche, di stringhe, di data, condizionali e di altro tipo per l'analisi.

Sebbene la maggior parte dei comandi nei linguaggi di query OpenSearch PPL e OpenSearch SQL siano comuni a CloudWatch Logs and OpenSearch, esistono alcune differenze nel set di comandi e funzioni supportato in ciascuno di questi servizi. Per ulteriori dettagli, consultate le tabelle nelle pagine seguenti.

- the section called “Comandi SQL supportati”
  - the section called “CloudWatch Logs Insights”
  - the section called “Restrizioni SQL generali”
- the section called “Comandi PPL supportati”
  - the section called “Informazioni aggiuntive per gli utenti di CloudWatch Logs Insights che utilizzano PPL OpenSearch”

Comandi e funzioni OpenSearch SQL supportati

Le seguenti tabelle di riferimento mostrano quali comandi SQL sono supportati in OpenSearch Discover per l'interrogazione dei dati in Amazon S3, Security Lake CloudWatch o Logs e quali comandi SQL sono supportati CloudWatch in Logs Insights. La sintassi SQL supportata in CloudWatch Logs Insights e quella supportata in OpenSearch Discover per l'interrogazione dei log sono le stesse e viene citata CloudWatch come Logs nelle tabelle seguenti. CloudWatch

> **Note**
>
> OpenSearch include anche il supporto SQL per l'interrogazione dei dati che vengono inseriti e archiviati negli indici. OpenSearch Questo dialetto SQL è diverso dall'SQL utilizzato nelle query dirette e viene denominato SQL sugli indici. OpenSearch

Argomenti

- Comandi
- Funzioni
- Restrizioni SQL generali
- Informazioni aggiuntive per gli utenti di CloudWatch Logs Insights che utilizzano SQL OpenSearch

Comandi

> **Note**
>
> Nella colonna dei comandi di esempio, sostituisci se necessario *<tableName/logGroup>* a seconda della fonte di dati su cui stai interrogando.
>
> - Comando di esempio: `SELECT Body , Operation FROM` **`<tableName/logGroup>`**
> - Se stai eseguendo una query su Amazon S3 o Security Lake, usa: `SELECT Body , Operation FROM` **`table_name`**
> - Se stai interrogando i CloudWatch log, usa: `SELECT Body , Operation FROM` **`` `LogGroupA` ``**

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called "Clausola SELECT"	Visualizza i valori proiettati.	Supporto	Supporto	Supporto	`SELECT method, status FROM <tableName/logGroup>`
the section called "Clausola WHERE"	Filtra gli eventi di registro in base ai criteri di	Supporto	Supporto	Supporto	`SELECT * FROM <tableName/logGroup> WHERE status = 100`

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
	campo forniti.				
the section called “Clausola GROUP BY”	I gruppi registrano gli eventi in base alla categoria e trovano la media in base alle statistiche.	Supporto	Supporto	Supporto	`SELECT method, status, COUNT(*) AS request_count, SUM(bytes) AS total_bytes FROM <tableName/logGroup> GROUP BY method, status`

Comando	Descripti on	CloudWa h Registri	Simple Storage Service (Amazor S3)	Lago di sicurezz	Esempio di comando
the section called “Clausola HAVING”	Filtra i risultati in base alle condizio i di raggrupp mento.	Suppor o	Suppor o	Suppor o	`SELECT method, status, COUNT(*) AS request_count, SUM(bytes) AS total_bytes FROM <tableName/logGroup> GROUP BY method, status HAVING COUNT(*) > 5`

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “Clausola ORDER BY”	Ordina i risultati in base ai campi della clausola order. È possibile ordinare in ordine decrescente o crescente.	Supporto	Supporto	Supporto	`SELECT * FROM <tableName/logGroup> ORDER BY status DESC`

Comando	Descripti on	CloudWa h Registri	Simple Storage Service (Amazor S3)	Lago di sicurezz	Esempio di comando

the section called “Clausola JOIN”
 	
(INNER | CROSS | LEFT OUTER)
	Unisce i risultati di due tabelle basate su campi comuni.	Suppor o (è necessa o utilizzar e Inner e Left Outer parole chiave per il join; in un'istruz ione SELECT è supporta a una sola operazio e JOIN)	Suppor a (è necessa o utilizzar e le parole chiave Inner, Left Outer e Cross per l'unione)	Suppor o (è necessa o utilizzar e le parole chiave Inner, Left Outer e Cross per l'unione)	`SELECT`
					`  A.Body,`
					`  B.Timestamp`
					`FROM`
					`  <tableNameA/logGroupA>  AS A`
					`INNER JOIN`
					`  <tableNameB/logGroupB>  AS B`
					`  ON A.`requestId` = B.`requestId``

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “Clausola LIMIT”	Limita i risultati alle prime N righe.	Supporto	Supporto	Supporto	`SELECT * FROM <tableName/logGroup> LIMIT 10`

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “Clausola CASE”	Valuta le condizioni e restituisce un valore quando viene soddisfatta la prima condizione.	Supporto	Supporto	Supporto	`SELECT method, status, CASE WHEN status BETWEEN 100 AND 199 THEN 'Informational' WHEN status BETWEEN 200 AND 299 THEN 'Success' WHEN status BETWEEN 300 AND 399 THEN 'Redirection' WHEN status BETWEEN 400 AND 499 THEN 'Client Error' WHEN status BETWEEN 500 AND 599 THEN 'Server Error' ELSE 'Unknown Status' END AS status_category, CASE method WHEN 'GET' THEN 'Read Operation' WHEN 'POST' THEN 'Create Operation' WHEN 'PUT' THEN 'Update Operation' WHEN 'PATCH' THEN 'Partial Update Operation' WHEN 'DELETE' THEN 'Delete Operation' ELSE 'Other Operation' END AS operation_type, bytes, datetime FROM <tableName/logGroup>`

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando

the section called “Espressione di tabella comune”	Crea un set di risultati temporaneo denominato all'interno di un'istruzione SELECT, INSERT, UPDATE, DELETE o MERGE.	Non supportato	Supporto	Supporto	`WITH RequestStats AS (`
					`  SELECT`
					`  method,`
					`  status,`
					`  bytes,`
					`  COUNT(*) AS request_count`
					`  FROM`
					`  tableName`
					`  GROUP BY`
					`  method,`
					`  status,`
					`  bytes`
					`)`
					`SELECT`
					`  method,`
					`  status,`
					`  bytes,`
					`  request_count`
					`FROM`
					`  RequestStats`
					`WHERE`
					`  bytes > 1000`

Comando	Descripti on	CloudWa h Registri	Simple Storage Service (Amazor S3)	Lago di sicurezz	Esempio di comando
the section called “EXPLAIN	Visualizz a il piano di esecuzic e di un'istruz ione SQL senza eseguirla effettiva mente.	Non supporta o	Suppor o	Suppor o	`EXPLAIN SELECT k, SUM(v) FROM VALUES (1, 2), (1, 3) AS t(k, v) GROUP BY k`

Comando	Descripti on	CloudWa h Registri	Simple Storage Service (Amazor S3)	Lago di sicurezz	Esempio di comando
the section called “Clausola LATERAL SUBQUEF	Consent a una sottoque y nella clausola FROM di fare riferimen to alle colonne degli elementi precede i nella stessa clausola FROM.	Non supporta a	Suppor o	Suppor o	`SELECT * FROM tableName LATERAL ( SELECT * FROM t2 WHERE t1.c1 = t2.c1 )`

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “Clausola LATERAL VIEW”	Genera una tabella virtuale applicando una funzione di generazione di tabelle a ciascuna riga di una tabella base.	Non supportata	Supporto	Supporto	`SELECT * FROM tableName LATERAL VIEW EXPLODE(ARRAY(30, 60)) tableName AS c_age LATERAL VIEW EXPLODE(ARRAY(40, 80)) AS d_age`
the section called “Predicato LIKE”	Corrisponde a una stringa in base a uno schema utilizzando caratteri jolly.	Supporto	Supporto	Supporto	`SELECT method, status, request, host FROM <tableName/logGroup> WHERE method LIKE 'D%'`

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “OFFSET”	Specifica te il numero di righe da saltare prima di iniziare a restituire le righe della query.	Supporto se utilizzato insieme a una `LIMIT` clausola in una query. Esempio • Supporto: `SELECT * FROM Table LIMIT 100 OFFSET 10` • Non supporto: `SELECT * FROM Table`	Supporto	Supporto	`SELECT method, status, bytes, datetime FROM <tableName/logGroup> ORDER BY datetime OFFSET 10`

Comando	Descriptio	CloudWa h Registri	Simple Storage Service (Amazor S3)	Lago di sicurezz	Esempio di comando
		OFFSE 10			
the section called “Clausola PIVOT”	Trasform le righe in colonne, ruotando i dati da un formato basato su righe a un formato basato su colonne.	Non supporta o	Suppor o	Suppor o	`SELECT * FROM ( SELECT method, status, bytes FROM <tableName/logGroup> ) AS SourceTable PIVOT ( SUM(bytes) FOR method IN ('GET', 'POST', 'PATCH', 'PUT', 'DELETE') ) AS PivotTable`

Comando	Descripti on	CloudWa h Registri	Simple Storage Service (Amazor S3)	Lago di sicurezz	Esempio di comando
the section called “Operator i su set”	Combina i risultati di due o più istruzion i SELECT (ad esempio UNION, INTERSI , EXCEPT	Suppor o	Suppor o	Suppor o	`SELECT method, status, bytes FROM <tableName/logGroup> WHERE status = '416' UNION SELECT method, status, bytes FROM <tableName/logGroup> WHERE bytes > 20000`
the section called “Clausola SORT BY”	Specifica l'ordine in cui restituir e i risultati della query.	Suppor o	Suppor o	Suppor o	`SELECT method, status, bytes FROM <tableName/logGroup> SORT BY bytes DESC`

Comando	Descrizione	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called "UNPIVOT	Trasforma le colonne in righe, ruotando i dati da un formato basato su colonne a un formato basato su righe.	Non supportato	Supporto	Supporto	`SELECT status, REPLACE(method, '_bytes', '') AS request_method, bytes, datetime FROM PivotedData UNPIVOT ( bytes FOR method IN ( GET_bytes, POST_bytes, PATCH_bytes, PUT_bytes, DELETE_bytes ) ) AS UnpivotedData`

Funzioni

> **Note**
>
> Nella colonna dei comandi di esempio, sostituisci *<tableName/logGroup>* se necessario a seconda della fonte di dati su cui stai interrogando.
>
> - Comando di esempio: `SELECT Body , Operation FROM` **`<tableName/logGroup>`**
> - Se stai eseguendo una query su Amazon S3 o Security Lake, usa: `SELECT Body , Operation FROM` **`table_name`**

- Se stai interrogando i CloudWatch log, usa: `SELECT Body , Operation FROM` **`` `LogGroupA` ``**

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic stringa	Funzioni integrate in grado di manipolare e trasforma re stringhe e dati di testo all'interno di query SQL. Ad esempio, convertire maiuscole e minuscole, combinare stringhe, estrarre parti e pulire il testo.	Suppo o	Suppo o	Suppo o	`SELECT UPPER(method) AS upper_method, LOWER(host) AS lower_hos t FROM <tableName/logGroup>`
the section called “Funzic di data	Funzioni integrate per la gestione e la trasforma zione dei dati di data e ora nelle	Suppo o	Suppo o	Suppo o	`SELECT TO_TIMESTAMP(datetime) AS timestamp, TIMESTAMP_SECONDS( UNIX_TIMESTAMP(datetime)) AS from_seconds, UNIX_TIMESTAMP(datetime) AS to_unix,`

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
e ora”	query. Ad esempio, date_add, date_form at, datediff e current_d ate.				FROM_UTC_TIMESTAMP (datetime, 'PST') AS to_pst, TO_UTC_TIMESTAMP(d atetime, 'EST') AS from_est FROM <tableName/logGroup>
the section called “Funzic di aggreg one”	Funzioni integrate che eseguono calcoli su più righe per produrre un unico valore riepilogato. Ad esempio, sum, count, avg, max e min.	Suppo o	Suppo o	Suppo o	SELECT COUNT(*) AS total_rec ords, COUNT(DISTINCT method) AS unique_methods, SUM(bytes) AS total_byt es, AVG(bytes) AS avg_bytes, MIN(bytes) AS min_bytes, MAX(bytes) AS max_bytes FROM <tableName/logGroup>

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic condizi ali”	Funzioni integrate che eseguono azioni in base a condizioni specifiche o che valutano le espressio ni in modo condizion ale. Ad esempio, CASE e IF.	Suppo o	Suppo o	Suppo o	`SELECT CASE WHEN method = 'GET' AND bytes < 1000 THEN 'Small Read' WHEN method = 'POST' AND bytes > 10000 THEN 'Large Write' WHEN status >= 400 OR bytes = 0 THEN 'Problem' ELSE 'Normal' END AS request_type FROM <tableName/logGroup>`

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic JSON”	Funzioni integrate per l'analisi, l'estrazione, la modifica e l'interro gazione di dati in formato JSON all'interno delle query SQL (ad esempio, from_json , to_json, get_json_ object, json_tupl e) che consentono la manipolaz ione delle strutture JSON nei set di dati.	Suppo o	Suppo o	Suppo o	`SELECT FROM_JSON( @message, 'STRUCT< host: STRING, user-identifier: STRING, datetime: STRING, method: STRING, status: INT, bytes: INT >' ) AS parsed_json FROM <tableName/logGroup>`

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic di array”	Funzioni integrate per lavorare con colonne di tipo array nelle query SQL, che consenton o operazion i come l'accesso, la modifica e l'analisi dei dati dell'array (ad esempio, size, explode, array_con tains).	Suppo o	Suppo o	Suppo o	`SELECT scores, size(scores) AS length, array_contains(scores, 90) AS has_90 FROM <tableName/logGroup>`

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic finestra	Funzioni integrate che eseguono calcoli su un insieme specifico di righe relative alla riga (finestra) corrente, che consenton o operazion i come la classific azione, i totali correnti e le medie mobili (ad esempio, ROW_NUMBE R, RANK, LAG, LEAD)	Suppo o	Suppo o	Suppo o	`SELECT field1, field2, RANK() OVER (ORDER BY field2 DESC) AS field2Rank FROM <tableName/logGroup>`

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic di conver ne”	Funzioni integrate per la conversio ne dei dati da un tipo all'altro all'interno delle query SQL, che consenton o trasforma zioni dei tipi di dati e conversioni di formato (ad esempio, CAST, TO_DATE, TO_TIMEST AMP, BINARY)	Suppo o	Suppo o	Suppo o	`SELECT CAST('123' AS INT) AS converted_number, CAST(123 AS STRING) AS converted_string FROM <tableName/logGroup>`

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic di predica ”	Funzioni integrate che valutano le condizion i e restituis cono valori booleani (true/fal se) in base a criteri o modelli specificati (ad esempio, IN, LIKE, BETWEEN, IS NULL, EXISTS)	Suppo o	Suppo o	Suppo o	`SELECT * FROM <tableName/logGroup> WHERE id BETWEEN 50000 AND 75000`

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic della mappa	Applica una funzione specifica ta a ogni elemento di una raccolta, trasforma ndo i dati in un nuovo set di valori.	Non suppor a	Supp o	Supp o	`SELECT MAP_FILTER( MAP( 'method', method, 'status', CAST(status AS STRING), 'bytes', CAST(byte s AS STRING) ), (k, v) -> k IN ('method', 'status') AND v ! = 'null' ) AS filtered_map FROM <tableName/logGroup> WHERE status = 100`
the section called “Funzic matem he”	Esegue operazioni matematic he su dati numerici, ad esempio il calcolo di medie, somme o valori trigonome trici.	Supp o	Supp o	Supp o	`SELECT bytes, bytes + 1000 AS added, bytes - 1000 AS subtracte d, bytes * 2 AS doubled, bytes / 1024 AS kilobytes , bytes % 1000 AS remainder FROM <tableName/logGroup>`

Gramm SQL disponi le	Description	CloudV h Registr	Simple Storag Service (Amaz S3)	Lago di sicurez	Esempio di comando
the section called “Funzic di gruppo multi-log”	Consente agli utenti di specificare più gruppi di log in un'istruz ione SQL SELECT	Supp o	Non applica le	Non applica le	`SELECT lg1.Column1, lg1.Column2 FROM `logGroups(logGrou pIdentifier: ['LogGroup1', 'LogGroup2'])` AS lg1 WHERE lg1.Column3 = "Success"`
the section called “Funzic del genera e”	Crea un oggetto iteratore che produce una sequenza di valori, che consente un utilizzo efficient e della memoria in set di dati di grandi dimensioni.	Non suppor o	Supp o	Supp o	`SELECT explode(array(10, 20))`

Restrizioni SQL generali

Le seguenti restrizioni si applicano all'utilizzo di OpenSearch SQL with CloudWatch Logs, Amazon S3 e Security Lake.

1. È possibile utilizzare una sola operazione JOIN in un'istruzione SELECT.

2. È supportato solo un livello di sottoquery annidate.
3. Le query di istruzioni multiple separate da punto e virgola non sono supportate.
4. Le query contenenti nomi di campo identici ma che differiscono solo per maiuscole e minuscole (ad esempio field1 e) non sono supportate. FIELD1

   Ad esempio, le seguenti query non sono supportate:

   ```
   Select AWSAccountId, awsaccountid from LogGroup
   ```

   Tuttavia, la seguente query è dovuta al fatto che il nome del campo (@logStream) è identico in entrambi i gruppi di log:

   ```
   Select a.`@logStream`, b.`@logStream` from Table A INNER Join Table B on a.id = b.id
   ```
5. Le funzioni e le espressioni devono operare sui nomi dei campi e far parte di un'istruzione SELECT con un gruppo di log specificato nella clausola FROM.

   Ad esempio, questa query non è supportata:

   ```
   SELECT cos(10) FROM LogGroup
   ```

   Questa interrogazione è supportata:

   ```
   SELECT cos(field1) FROM LogGroup
   ```

Informazioni aggiuntive per gli utenti di CloudWatch Logs Insights che utilizzano SQL OpenSearch

CloudWatch Logs supporta le query OpenSearch SQL nella console, nell'API e nella CLI di Logs Insights. Supporta la maggior parte dei comandi, tra cui SELECT, FROM, WHERE, GROUP BY, HAVING, JOINS e le query annidate, oltre a funzioni JSON, matematiche, stringhe e condizionali. Tuttavia, CloudWatch Logs supporta solo operazioni di lettura, quindi non consente istruzioni DDL o DML. Consulta le tabelle nelle sezioni precedenti per un elenco completo dei comandi e delle funzioni supportati.

Funzioni di gruppo multi-log

CloudWatch Logs Insights supporta la possibilità di interrogare più gruppi di log. Per risolvere questo caso d'uso in SQL, è possibile utilizzare il `logGroups` comando. Questo comando è specifico per l'interrogazione dei dati in CloudWatch Logs Insights che coinvolge uno o più gruppi di log. Utilizzate questa sintassi per interrogare più gruppi di log specificandoli nel comando, invece di scrivere una query per ciascuno dei gruppi di log e combinarli con un comando. `UNION`

Sintassi:

```
`logGroups(
    logGroupIdentifier: ['LogGroup1','LogGroup2', ...'LogGroupn']
)
```

In questa sintassi, è possibile specificare fino a 50 gruppi di log nel parametro. `logGroupIndentifier` Per fare riferimento ai gruppi di log in un account di monitoraggio, usa ARNs al posto dei `LogGroup` nomi.

Query di esempio:

```
SELECT LG1.Column1, LG1.Column2 from `logGroups(
    logGroupIdentifier: ['LogGroup1', 'LogGroup2']
)` as LG1
WHERE LG1.Column1 = 'ABC'
```

La seguente sintassi che coinvolge più gruppi di log dopo l'`FROM`istruzione non è supportata durante l' CloudWatch interrogazione dei log:

```
SELECT Column1, Column2 FROM 'LogGroup1', 'LogGroup2', ...'LogGroupn'
WHERE Column1 = 'ABC'
```

Restrizioni

Quando usi i comandi SQL o PPL, racchiudi alcuni campi tra i backtick per interrogarli. I backtick sono obbligatori per i campi con caratteri speciali (non alfabetici e non numerici). Ad esempio, racchiudi e inserisci i contrassegni. `@message` `Operation.Export, Test::Field` Non è necessario racchiudere colonne con nomi puramente alfabetici nei backtick.

Query di esempio con campi semplici:

```
SELECT SessionToken, Operation, StartTime  FROM `LogGroup-A`
```

```
LIMIT 1000;
```

Stessa query con i backtick aggiunti:

```
SELECT `SessionToken`, `Operation`, `StartTime`  FROM `LogGroup-A`
LIMIT 1000;
```

Per ulteriori restrizioni generali non specifiche dei CloudWatch registri, consulta. the section called "Restrizioni SQL generali"

Esempi di domande e quote

> **Note**
>
> Quanto segue si applica sia agli utenti di CloudWatch Logs Insights che agli utenti che effettuano query sui dati. OpenSearch CloudWatch

Per esempi di query SQL che puoi utilizzare nei CloudWatch log, consulta la sezione Query salvate e di esempio nella console Amazon CloudWatch Logs Insights per alcuni esempi.

Per informazioni sui limiti che si applicano all'interrogazione di CloudWatch Logs from OpenSearch Service, consulta CloudWatch Logs quotas nella Amazon CloudWatch Logs User Guide. I limiti riguardano il numero di gruppi di CloudWatch log su cui è possibile eseguire query, il numero massimo di query simultanee che è possibile eseguire, il tempo massimo di esecuzione delle query e il numero massimo di righe restituite nei risultati. I limiti sono gli stessi indipendentemente dalla lingua utilizzata per l'interrogazione dei log (vale a dire OpenSearch PPL, SQL e CloudWatch Logs Insights).

Comandi SQL

Argomenti

- Funzioni stringa
- Funzioni di data e ora
- Funzioni di aggregazione
- Funzioni condizionali
- Funzioni JSON
- Funzioni di array

- Funzioni finestra
- Funzioni di conversione
- Funzioni di predicato
- Funzioni della mappa
- Funzioni matematiche
- Funzioni del generatore
- Clausola SELECT
- Clausola WHERE
- Clausola GROUP BY
- Clausola HAVING
- Clausola ORDER BY
- Clausola JOIN
- Clausola LIMIT
- Clausola CASE
- Espressione di tabella comune
- EXPLAIN
- Clausola LATERAL SUBQUERY
- Clausola LATERAL VIEW
- Predicato LIKE
- OFFSET
- Clausola PIVOT
- Operatori su set
- Clausola SORT BY
- UNPIVOT

Funzioni stringa

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Funzione	Description
ascii (str)	Restituisce il valore numerico del primo carattere di. `str`
base64 (bin)	Converte l'argomento da un file binario `bin` a una stringa in base 64.
lunghezza_bit (expr)	Restituisce la lunghezza in bit dei dati di stringa o il numero di bit dei dati binari.
btrim (str)	Rimuove i caratteri spaziali iniziali e finali da. `str`
btrim (str, trimStr)	Rimuovi i caratteri iniziali e finali `trimStr` da. `str`
char (expr)	Restituisce il carattere ASCII con l'equival ente binario di. `expr` Se n è maggiore di 256 il risultato è equivalente a chr (n% 256)
lunghezza del carattere (expr)	Restituisce la lunghezza dei caratteri dei dati della stringa o il numero di byte dei dati binari. La lunghezza dei dati della stringa include gli spazi finali. La lunghezza dei dati binari include gli zeri binari.
carattere_lunghezza (expr)	Restituisce la lunghezza dei caratteri dei dati della stringa o il numero di byte dei dati binari. La lunghezza dei dati della stringa include gli spazi finali. La lunghezza dei dati binari include gli zeri binari.
chr (expr)	Restituisce il carattere ASCII con l'equival ente binario di. `expr` Se n è maggiore di 256 il risultato è equivalente a chr (n% 256)

Funzione	Description
concat_ws (sep [, str | array (str)] +)	Restituisce la concatenazione delle stringhe separate da, saltando i valori nulli. `sep`
contiene (sinistra, destra)	Restituisce un valore booleano. Il valore è True se la parte destra si trova all'interno della sinistra. Restituisce NULL se una delle espressioni di input è NULL. Altrimenti, restituis ce False. Sia la sinistra che la destra devono essere di tipo STRING o BINARY.
decode (bin, set di caratteri)	Decodifica il primo argomento utilizzando il set di caratteri del secondo argomento.
decode (expr, search, result [, search, result]... [, predefinito])	Confronta expr con ogni valore di ricerca in ordine. Se expr è uguale a un valore di ricerca, decode restituisce il risultato corrispondente. Se non viene trovata alcuna corrispondenza, restituisce il valore predefinito. Se il valore predefinito viene omesso, restituisce null.
elt (n, input1, input2,...)	Restituisce il n -esimo input, ad esempio restituisce quando è 2. `input2` `n`
encode (str, charset)	Codifica il primo argomento utilizzando il set di caratteri del secondo argomento.
termina con (sinistra, destra)	Restituisce un valore booleano. Il valore è True se la parte sinistra termina con la destra. Restituisce NULL se una delle espressioni di input è NULL. Altrimenti, restituisce False. Sia la sinistra che la destra devono essere di tipo STRING o BINARY.

Funzione	Description
find_in_set (str, str_array)	Restituisce l'indice (basato su 1) della stringa specificata () nell'elenco delimitato da virgole (`str`). `str_array` Restituisce 0, se la stringa non è stata trovata o se la stringa specificata () `str` contiene una virgola.
format_number (espr 1, espr 2)	Formatta il numero `expr1` come '#, ###, ###.##', arrotondato al numero decimale. `expr2` Se `expr2` è 0, il risultato non ha punti decimali o parti frazionarie. `expr2`accetta anche un formato specificato dall'utente. Questo dovrebbe funzionare come FORMAT di MySQL.
format_string (strfmt, obj,...)	Restituisce una stringa formattata da stringhe di formato in stile printf.
initcap (str)	Restituisce `str` con la prima lettera di ogni parola in maiuscolo. Tutte le altre lettere sono in minuscolo. Le parole sono delimitate da spazi bianchi.
instr (str, substr)	Restituisce l'indice (a base 1) della prima occorrenza di in. `substr` `str`
lcase (str)	Restituisce `str` con tutti i caratteri cambiati in minuscolo.
sinistra (str, len)	Restituisce i caratteri più a sinistra `len` (`len`può essere di tipo stringa) dalla stringa`str`, se `len` è minore o uguale a 0 il risultato è una stringa vuota.

Funzione	Description
len (espr)	Restituisce la lunghezza dei caratteri dei dati della stringa o il numero di byte dei dati binari. La lunghezza dei dati della stringa include gli spazi finali. La lunghezza dei dati binari include gli zeri binari.
lunghezza (expr)	Restituisce la lunghezza dei caratteri dei dati della stringa o il numero di byte dei dati binari. La lunghezza dei dati della stringa include gli spazi finali. La lunghezza dei dati binari include gli zeri binari.
levenshtein (str1, str2 [, soglia])	Restituisce la distanza di Levenshtein tra le due stringhe date. Se la soglia è impostata e la distanza è maggiore, restituisce -1.
localizza (substr, str [, pos])	Restituisce la posizione della prima occorrenz a di `substr` in `str` dopo la posizione`pos`. Il valore dato `pos` e quello restituito sono basati su 1.
inferiore (str)	Restituisce `str` con tutti i caratteri cambiati in minuscolo.
lpad (str, len [, pad])	Restituisce`str`, imbottito a sinistra con `pad` una lunghezza di. `len` Se `str` è più lungo di`len`, il valore restituito viene abbreviato in caratteri o byte. `len` Se non `pad` è specifica to, `str` verrà aggiunto a sinistra con caratteri spaziali se si tratta di una stringa di caratteri e con zeri se si tratta di una sequenza di byte.
ltrim (str)	Rimuove i caratteri spaziali iniziali da. `str`

Funzione	Description
lun_check (str)	Verifica che una stringa di cifre sia valida secondo l'algoritmo di Luhn. Questa funzione di checksum è ampiamente utilizzata sui numeri di carte di credito e sui numeri di identificazione governativi per distinguere i numeri validi da quelli digitati erroneamente.
maschera (input [, upperChar, lowerChar, DigitChar, OtherChar])	maschera il valore di stringa dato. La funzione sostituisce i caratteri con 'X' o 'x' e i numeri con 'n'. Ciò può essere utile per creare copie di tabelle con informazioni riservate rimosse.
lunghezza_ottetto (expr)	Restituisce la lunghezza in byte dei dati di stringa o il numero di byte dei dati binari.
overlay (input, replace, pos [, len])	Sostituisci `input` con `replace` quello inizia da `pos` ed è di lunghezza. `len`
posizione (substr, str [, pos])	Restituisce la posizione della prima occorrenz a di `substr` in `str` dopo la posizione`pos`. Il valore dato `pos` e quello restituito sono basati su 1.
printf (strfmt, obj,...)	Restituisce una stringa formattata da stringhe di formato in stile printf.
regexp_count (str, regexp)	Restituisce un conteggio del numero di volte in cui il modello di espressione regolare `regexp` viene trovato nella stringa. `str`
regexp_extract (str, regexp [, idx])	Estrai la prima stringa che corrisponde all'`regexp`espressione e `str` che corrisponde all'indice del gruppo regex.

Funzione	Description
regexp_extract_all (str, regexp [, idx])	Estrai tutte le stringhe `str` che corrispondono all'espressione e corrispondono all'indice del gruppo regex. `regexp`
regexp_instr (str, regexp)	Cerca un'espressione regolare in una stringa e restituisce un numero intero che indica la posizione iniziale della sottostringa corrispon dente. Le posizioni sono basate su 1, non su 0. Se non viene trovata alcuna corrispondenza, restituisce 0.
regexp_replace (str, regexp, rep [, position])	Sostituisce tutte le sottostringhe di quella corrispondenza con. `str` `regexp` `rep`
regexp_substr (str, regexp)	Restituisce la sottostringa che corrispon de all'espressione regolare all'interno della stringa. `regexp` `str` Se l'espressione regolare non viene trovata, il risultato è nullo.
ripetere (str, n)	Restituisce la stringa che ripete il valore di stringa dato n volte.
sostituisci (str, cerca [, sostituisci])	Sostituisce tutte le occorrenze di with. `search` `replace`
destra (str, len)	Restituisce i caratteri più a destra `len` (`len`possono essere di tipo stringa) dalla string`str`, se `len` è minore o uguale a 0 il risultato è una stringa vuota.

Funzione	Description
rpad (str, len [, pad])	Restituisce`str`, riempito a destra con una lunghezza `pad` di. `len` Se `str` è più lungo di`len`, il valore restituito viene ridotto in caratteri. `len` Se non `pad` è specificato, `str` verrà aggiunto a destra con caratteri spaziali se si tratta di una stringa di caratteri e con zeri se si tratta di una stringa binaria.
rtrim (str)	Rimuove i caratteri spaziali finali da. `str`
frasi (str [, lang, country])	Si `str` divide in una serie di parole.
soundex (str)	Restituisce il codice Soundex della stringa.
spazio (n)	Restituisce una stringa composta da n spazi.
split (str, regex, limit)	Divide in base `str` alle occorrenze che corrispondono `regex` e restituisce un array con una lunghezza massima di `limit`
split_part (str, delimiter, partNum)	Divide `str` per delimitatore e restituisce la parte della divisione richiesta (in base 1). Se un input è nullo, restituisce nullo. se non rientra nell'intervallo delle parti `partNum` suddivise, restituisce una stringa vuota. Se `partNum` è 0, genera un errore. Se `partNum` è negativo, le parti vengono contate all'indietro a partire dalla fine della stringa. Se `delimiter` è una stringa vuota, non `str` viene divisa.
inizia con (sinistra, destra)	Restituisce un valore booleano. Il valore è True se la sinistra inizia con la destra. Restituis ce NULL se una delle espressioni di input è NULL. Altrimenti, restituisce False. Sia la sinistra che la destra devono essere di tipo STRING o BINARY.

Funzione	Description
substr (str, pos [, len])	Restituisce la sottostringa `str` che inizia da `pos` ed è lunga`len`, oppure la porzione di matrice di byte che inizia da ed è lunga. `pos` `len`
substr (str FROM pos [FOR len]])	Restituisce la sottostringa `str` che inizia da `pos` ed è lunga`len`, oppure la porzione di matrice di byte che inizia da ed è lunga. `pos` `len`
sottostringa (str, pos [, len])	Restituisce la sottostringa `str` che inizia da `pos` ed è lunga`len`, oppure la porzione di matrice di byte che inizia da ed è lunga. `pos` `len`
sottostringa (str FROM pos [FOR len]])	Restituisce la sottostringa `str` che inizia da `pos` ed è lunga`len`, oppure la porzione di matrice di byte che inizia da ed è lunga. `pos` `len`
substring_index (str, delim, count)	Restituisce la sottostringa `str` precedente `count` alle occorrenze del delimitatore. `delim` Se `count` è positivo, viene restituito tutto ciò che si trova a sinistra del delimitatore finale (contando da sinistra). Se `count` è negativo, viene restituito tutto ciò che si trova a destra del delimitatore finale (contando da destra). La funzione substring_index esegue una corrispondenza con distinzione tra maiuscole e minuscole durante la ricerca di. `delim`

Funzione	Description
to_binary (str [, fmt])	Converte l'input in un valore binario in base `str` al valore fornito. `fmt` `fmt`può essere una stringa letterale senza distinzione tra maiuscole e minuscole di «hex», «utf-8", «utf8" o «base64". Per impostazione predefinita, il formato binario per la conversione è «hex» se viene omesso. `fmt` La funzione restituisce NULL se almeno uno dei parametri di input è NULL.
to_char (NumberExpr, FormatExpr)	Converti in una stringa basata su. `numberExpr formatExpr` Genera un'eccezione se la conversione fallisce. Il formato può essere composto dai seguenti caratteri, senza distinzione tra maiuscole e minuscole: '0' o '9': specifica una cifra prevista compresa tra 0 e 9. Una sequenza di 0 o 9 nella stringa di formato corrisponde a una sequenza di cifre nel valore di input, generando una stringa di risultato della stessa lunghezza della sequenza corrispondente nella stringa di formato. La stringa risultante viene riempita con zeri a sinistra se la sequenza 0/9 comprende più cifre della parte corrispondente del valore decimale, inizia con 0 ed è prima della virgola decimale. Altrimenti, viene riempito con spazi. '.' o 'D': specifica la posizione del punto decimale (facoltativo, consentito solo una volta). ',' o 'G': specifica la posizione del separatore di raggruppamento (migliaia) (,). Deve esserci uno 0 o 9 a sinistra e a destra di ogni separator e di raggruppamento. '

Funzione	Description
to_number (expr, fmt)	Converti la stringa 'expr' in un numero in base al formato di stringa 'fmt'. Genera un'eccezi one se la conversione fallisce. Il formato può essere composto dai seguenti caratteri, senza distinzione tra maiuscole e minuscole: '0' o '9': specifica una cifra prevista compresa tra 0 e 9. Una sequenza di 0 o 9 nella stringa di formato corrisponde a una sequenza di cifre nella stringa di input. Se la sequenza 0/9 inizia con 0 e precede la virgola decimale, può corrispondere solo a una sequenza di cifre della stessa dimensione. Altrimenti, se la sequenza inizia con 9 o è dopo la virgola decimale, può corrispondere a una sequenza di cifre di dimensioni uguali o inferiori. '.' o 'D': specifica la posizione della virgola decimale (opzionale, consentita solo una volta). ',' o 'G': specifica la posizione del separatore di raggruppamento (migliaia) (,). Deve esserci uno 0 o 9 a sinistra e a destra di ogni separator e di raggruppamento. 'expr' deve corrispondere al separatore di raggruppamento pertinente alla dimensione del numero. '

Funzione	Description
to_varchar (NumberExpr, FormatExpr)	Converti in `numberExpr` una stringa basata su. `formatExpr` Genera un'eccezione se la conversione fallisce. Il formato può essere composto dai seguenti caratteri, senza distinzione tra maiuscole e minuscole: '0' o '9': specifica una cifra prevista compresa tra 0 e 9. Una sequenza di 0 o 9 nella stringa di formato corrisponde a una sequenza di cifre nel valore di input, generando una stringa di risultato della stessa lunghezza della sequenza corrispondente nella stringa di formato. La stringa risultante viene riempita con zeri a sinistra se la sequenza 0/9 comprende più cifre della parte corrispondente del valore decimale, inizia con 0 ed è prima della virgola decimale. Altrimenti, viene riempito con spazi. '.' o 'D': specifica la posizione del punto decimale (facoltativo, consentito solo una volta). ',' o 'G': specifica la posizione del separatore di raggruppamento (migliaia) (,). Deve esserci uno 0 o 9 a sinistra e a destra di ogni separator e di raggruppamento. '
tradurre (input, da, a)	Traduce la `input` stringa sostituendo i caratteri presenti nella `from` stringa con i caratteri corrispondenti nella `to` stringa.
trim (str)	Rimuove i caratteri spaziali iniziali e finali da. `str`
trim (ENTRAMBI DA str)	Rimuove i caratteri spaziali iniziali e finali da. `str`
trim (LEADING FROM str)	Rimuove i caratteri spaziali iniziali da`str`.
trim (TRAILING FROM str)	Rimuove i caratteri spaziali finali da. `str`

Funzione	Description
trim (trimStr FROM str)	Rimuovi i caratteri iniziali e finali da. `trimStr` `str`
trim (SIA trimStr che str)	Rimuovi i caratteri iniziali e finali da. `trimStr` `str`
trim (LEADING TrimStr FROM str)	Rimuovi i `trimStr` personaggi principali da. `str`
trim (TRAILING TrimStr FROM str)	Rimuovi i caratteri finali `trimStr` da. `str`
try_to_binary (str [, fmt])	Questa è una versione speciale `to_binary` che esegue la stessa operazione, ma restituis ce un valore NULL invece di generare un errore se la conversione non può essere eseguita.
try_to_number (expr, fmt)	Converti la stringa 'expr' in un numero in base al formato della stringa. `fmt` Restituisce NULL se la stringa 'expr' non corrisponde al formato previsto. Il formato segue la stessa semantica della funzione to_number.
o case (str)	Ritorna `str` con tutti i caratteri cambiati in maiuscolo.
unbase64 (str)	Converte l'argomento da una stringa `str` in base 64 a un file binario.
upper (str)	Ritorna `str` con tutti i caratteri cambiati in maiuscolo.

Esempi

```
-- ascii
SELECT ascii('222');
```

```
+----------+
|ascii(222)|
+----------+
|        50|
+----------+
SELECT ascii(2);
+--------+
|ascii(2)|
+--------+
|      50|
+--------+
-- base64
SELECT base64('Feathers');
+-----------------+
|base64(Feathers)|
+-----------------+
|     RmVhdGhlcnM=|
+-----------------+
SELECT base64(x'537061726b2053514c');
+-----------------------------+
|base64(X'537061726B2053514C')|
+-----------------------------+
|                 U3BhcmsgU1FM|
+-----------------------------+
-- bit_length
SELECT bit_length('Feathers');
+---------------------+
|bit_length(Feathers)|
+---------------------+
|                   64|
+---------------------+
SELECT bit_length(x'537061726b2053514c');
+---------------------------------+
|bit_length(X'537061726B2053514C')|
+---------------------------------+
|                               72|
+---------------------------------+
-- btrim
SELECT btrim('    Feathers   ');
+----------------------+
|btrim(    Feathers   )|
+----------------------+
|              Feathers|
+----------------------+
```

```
SELECT btrim(encode('    Feathers   ', 'utf-8'));
+-----------------------------------+
|btrim(encode(    Feathers   , utf-8))|
+-----------------------------------+
|                           Feathers|
+-----------------------------------+
SELECT btrim('Feathers', 'Fe');
+---------------------+
|btrim(Alphabet, Al)|
+---------------------+
|               athers|
+---------------------+
SELECT btrim(encode('Feathers', 'utf-8'), encode('Al', 'utf-8'));
+-----------------------------------------------------+
|btrim(encode(Feathers, utf-8), encode(Al, utf-8))|
+-----------------------------------------------------+
|                                              athers|
+-----------------------------------------------------+
-- char
SELECT char(65);
+--------+
|char(65)|
+--------+
|       A|
+--------+
-- char_length
SELECT char_length('Feathers ');
+-----------------------+
|char_length(Feathers )|
+-----------------------+
|                      9 |
+-----------------------+
SELECT char_length(x'537061726b2053514c');
+---------------------------------+
|char_length(X'537061726B2053514C')|
+---------------------------------+
|                                9|
+---------------------------------+
SELECT CHAR_LENGTH('Feathers ');
+-----------------------+
|char_length(Feathers )|
+-----------------------+
|                      9|
+-----------------------+
```

```
SELECT CHARACTER_LENGTH('Feathers ');
+---------------------------+
|character_length(Feathers )|
+---------------------------+
|                          9|
+---------------------------+
-- character_length
SELECT character_length('Feathers ');
+---------------------------+
|character_length(Feathers )|
+---------------------------+
|                          9|
+---------------------------+
SELECT character_length(x'537061726b2053514c');
+--------------------------------------+
|character_length(X'537061726B2053514C')|
+--------------------------------------+
|                                     9|
+--------------------------------------+
SELECT CHAR_LENGTH('Feathers ');
+-----------------------+
|char_length(Feathers )|
+-----------------------+
|                      9|
+-----------------------+
SELECT CHARACTER_LENGTH('Feathers ');
+---------------------------+
|character_length(Feathers )|
+---------------------------+
|                          9|
+---------------------------+
-- chr
SELECT chr(65);
+-------+
|chr(65)|
+-------+
|      A|
+-------+
-- concat_ws
SELECT concat_ws(' ', 'Fea', 'thers');
+-----------------------+
|concat_ws( , Fea, thers)|
+-----------------------+
|               Feathers|
```

```
+------------------------+
SELECT concat_ws('s');
+------------+
|concat_ws(s)|
+------------+
|            |
+------------+
SELECT concat_ws('/', 'foo', null, 'bar');
+----------------------------+
|concat_ws(/, foo, NULL, bar)|
+----------------------------+
|                     foo/bar|
+----------------------------+
SELECT concat_ws(null, 'Fea', 'thers');
+---------------------------+
|concat_ws(NULL, Fea, thers)|
+---------------------------+
|                       NULL|
+---------------------------+
-- contains
SELECT contains('Feathers', 'Fea');
+--------------------------+
|contains(Feathers, Fea)|
+--------------------------+
|                      true|
+--------------------------+
SELECT contains('Feathers', 'SQL');
+--------------------------+
|contains(Feathers, SQL)|
+--------------------------+
|                     false|
+--------------------------+
SELECT contains('Feathers', null);
+-------------------------+
|contains(Feathers, NULL)|
+-------------------------+
|                     NULL|
+-------------------------+
SELECT contains(x'537061726b2053514c', x'537061726b');
+-----------------------------------------------+
|contains(X'537061726B2053514C', X'537061726B')|
+-----------------------------------------------+
|                                           true|
+-----------------------------------------------+
```

```
-- decode
SELECT decode(encode('abc', 'utf-8'), 'utf-8');
+---------------------------------+
|decode(encode(abc, utf-8), utf-8)|
+---------------------------------+
|                              abc|
+---------------------------------+
SELECT decode(2, 1, 'Southlake', 2, 'San Francisco', 3, 'New Jersey', 4, 'Seattle',
 'Non domestic');
+-------------------------------------------------------------------------------+
|decode(2, 1, Southlake, 2, San Francisco, 3, New Jersey, 4, Seattle, Non domestic)|
+-------------------------------------------------------------------------------+
|                                                                  San Francisco|
+-------------------------------------------------------------------------------+
SELECT decode(6, 1, 'Southlake', 2, 'San Francisco', 3, 'New Jersey', 4, 'Seattle',
 'Non domestic');
+-------------------------------------------------------------------------------+
|decode(6, 1, Southlake, 2, San Francisco, 3, New Jersey, 4, Seattle, Non domestic)|
+-------------------------------------------------------------------------------+
|                                                                   Non domestic|
+-------------------------------------------------------------------------------+
SELECT decode(6, 1, 'Southlake', 2, 'San Francisco', 3, 'New Jersey', 4, 'Seattle');
+-----------------------------------------------------------------+
|decode(6, 1, Southlake, 2, San Francisco, 3, New Jersey, 4, Seattle)|
+-----------------------------------------------------------------+
|                                                             NULL|
+-----------------------------------------------------------------+
SELECT decode(null, 6, 'Fea', NULL, 'thers', 4, 'rock');
+-----------------------------------------+
|decode(NULL, 6, Fea, NULL, thers, 4, rock)|
+-----------------------------------------+
|                                    thers|
+-----------------------------------------+
-- elt
SELECT elt(1, 'scala', 'java');
+-------------------+
|elt(1, scala, java)|
+-------------------+
|              scala|
+-------------------+
SELECT elt(2, 'a', 1);
+------------+
|elt(2, a, 1)|
+------------+
```

```
|           1|
+------------+
-- encode
SELECT encode('abc', 'utf-8');
+------------------+
|encode(abc, utf-8)|
+------------------+
|        [61 62 63]|
+------------------+
-- endswith
SELECT endswith('Feathers', 'ers');
+------------------------+
|endswith(Feathers, ers)|
+------------------------+
|                    true|
+------------------------+
SELECT endswith('Feathers', 'SQL');
+-------------------------+
|endswith(Feathers, SQL)|
+-------------------------+
|                     false|
+-------------------------+
SELECT endswith('Feathers', null);
+-------------------------+
|endswith(Feathers, NULL)|
+-------------------------+
|                     NULL|
+-------------------------+
SELECT endswith(x'537061726b2053514c', x'537061726b');
+-----------------------------------------------+
|endswith(X'537061726B2053514C', X'537061726B')|
+-----------------------------------------------+
|                                          false|
+-----------------------------------------------+
SELECT endswith(x'537061726b2053514c', x'53514c');
+---------------------------------------+
|endswith(X'537061726B2053514C', X'53514C')|
+---------------------------------------+
|                                   true|
+---------------------------------------+
-- find_in_set
SELECT find_in_set('ab','abc,b,ab,c,def');
+-------------------------------+
|find_in_set(ab, abc,b,ab,c,def)|
```

```
+------------------------------+
|                             3|
+------------------------------+
-- format_number
SELECT format_number(12332.123456, 4);
+------------------------------+
|format_number(12332.123456, 4)|
+------------------------------+
|                  12,332.1235|
+------------------------------+
SELECT format_number(12332.123456, '##################.###');
+---------------------------------------------------+
|format_number(12332.123456, ##################.###)|
+---------------------------------------------------+
|                                           12332.123|
+---------------------------------------------------+
-- format_string
SELECT format_string("Hello World %d %s", 100, "days");
+-------------------------------------------+
|format_string(Hello World %d %s, 100, days)|
+-------------------------------------------+
|                       Hello World 100 days|
+-------------------------------------------+
-- initcap
SELECT initcap('Feathers');
+------------------+
|initcap(Feathers)|
+------------------+
|         Feathers|
+------------------+
-- instr
SELECT instr('Feathers', 'ers');
+--------------------+
|instr(Feathers, ers)|
+--------------------+
|                   6|
+--------------------+
-- lcase
SELECT lcase('Feathers');
+---------------+
|lcase(Feathers)|
+---------------+
|       feathers|
+---------------+
```

```
-- left
SELECT left('Feathers', 3);
+------------------+
|left(Feathers, 3)|
+------------------+
|               Fea|
+------------------+
SELECT left(encode('Feathers', 'utf-8'), 3);
+---------------------------------+
|left(encode(Feathers, utf-8), 3)|
+---------------------------------+
|                       [RmVh]|
+---------------------------------+
-- len
SELECT len('Feathers ');
+---------------+
|len(Feathers )|
+---------------+
|             9|
+---------------+
SELECT len(x'537061726b2053514c');
+--------------------------+
|len(X'537061726B2053514C')|
+--------------------------+
|                         9|
+--------------------------+
SELECT CHAR_LENGTH('Feathers ');
+-----------------------+
|char_length(Feathers )|
+-----------------------+
|                      9|
+-----------------------+
SELECT CHARACTER_LENGTH('Feathers ');
+----------------------------+
|character_length(Feathers )|
+----------------------------+
|                           9|
+----------------------------+
-- length
SELECT length('Feathers ');
+------------------+
|length(Feathers )|
+------------------+
|                9|
```

```
+------------------+
SELECT length(x'537061726b2053514c');
+----------------------------+
|length(X'537061726B2053514C')|
+----------------------------+
|                          9|
+----------------------------+
SELECT CHAR_LENGTH('Feathers ');
+----------------------+
|char_length(Feathers )|
+----------------------+
|                     9|
+----------------------+
SELECT CHARACTER_LENGTH('Feathers ');
+---------------------------+
|character_length(Feathers )|
+---------------------------+
|                         9|
+---------------------------+
-- levenshtein
SELECT levenshtein('kitten', 'sitting');
+---------------------------+
|levenshtein(kitten, sitting)|
+---------------------------+
|                          3|
+---------------------------+
SELECT levenshtein('kitten', 'sitting', 2);
+-----------------------------+
|levenshtein(kitten, sitting, 2)|
+-----------------------------+
|                           -1|
+-----------------------------+
-- locate
SELECT locate('bar', 'foobarbar');
+-------------------------+
|locate(bar, foobarbar, 1)|
+-------------------------+
|                        4|
+-------------------------+
SELECT locate('bar', 'foobarbar', 5);
+-------------------------+
|locate(bar, foobarbar, 5)|
+-------------------------+
|                        7|
```

```
+-------------------------+
SELECT POSITION('bar' IN 'foobarbar');
+-------------------------+
|locate(bar, foobarbar, 1)|
+-------------------------+
|                        4|
+-------------------------+
-- lower
SELECT lower('Feathers');
+---------------+
|lower(Feathers)|
+---------------+
|       feathers|
+---------------+
-- lpad
SELECT lpad('hi', 5, '??');
+---------------+
|lpad(hi, 5, ??)|
+---------------+
|          ???hi|
+---------------+
SELECT lpad('hi', 1, '??');
+---------------+
|lpad(hi, 1, ??)|
+---------------+
|              h|
+---------------+
SELECT lpad('hi', 5);
+--------------+
|lpad(hi, 5,  )|
+--------------+
|            hi|
+--------------+
SELECT hex(lpad(unhex('aabb'), 5));
+--------------------------------+
|hex(lpad(unhex(aabb), 5, X'00'))|
+--------------------------------+
|                     000000AABB|
+--------------------------------+
SELECT hex(lpad(unhex('aabb'), 5, unhex('1122')));
+-------------------------------------+
|hex(lpad(unhex(aabb), 5, unhex(1122)))|
+-------------------------------------+
|                           112211AABB|
```

```
+--------------------------------------+
-- ltrim
SELECT ltrim('    Feathers   ');
+----------------------+
|ltrim(    Feathers   )|
+----------------------+
|           Feathers   |
+----------------------+
-- luhn_check
SELECT luhn_check('8112189876');
+----------------------+
|luhn_check(8112189876)|
+----------------------+
|                  true|
+----------------------+
SELECT luhn_check('79927398713');
+-----------------------+
|luhn_check(79927398713)|
+-----------------------+
|                   true|
+-----------------------+
SELECT luhn_check('79927398714');
+-----------------------+
|luhn_check(79927398714)|
+-----------------------+
|                  false|
+-----------------------+
-- mask
SELECT mask('abcd-EFGH-8765-4321');
+---------------------------------------+
|mask(abcd-EFGH-8765-4321, X, x, n, NULL)|
+---------------------------------------+
|                    xxxx-XXXX-nnnn-nnnn|
+---------------------------------------+
SELECT mask('abcd-EFGH-8765-4321', 'Q');
+---------------------------------------+
|mask(abcd-EFGH-8765-4321, Q, x, n, NULL)|
+---------------------------------------+
|                    xxxx-QQQQ-nnnn-nnnn|
+---------------------------------------+
SELECT mask('AbCD123-@$#', 'Q', 'q');
+--------------------------------+
|mask(AbCD123-@$#, Q, q, n, NULL)|
+--------------------------------+
```

```
|                     QqQQnnn-@$#|
+--------------------------------+
SELECT mask('AbCD123-@$#');
+--------------------------------+
|mask(AbCD123-@$#, X, x, n, NULL)|
+--------------------------------+
|                     XxXXnnn-@$#|
+--------------------------------+
SELECT mask('AbCD123-@$#', 'Q');
+--------------------------------+
|mask(AbCD123-@$#, Q, x, n, NULL)|
+--------------------------------+
|                     QxQQnnn-@$#|
+--------------------------------+
SELECT mask('AbCD123-@$#', 'Q', 'q');
+--------------------------------+
|mask(AbCD123-@$#, Q, q, n, NULL)|
+--------------------------------+
|                     QqQQnnn-@$#|
+--------------------------------+
SELECT mask('AbCD123-@$#', 'Q', 'q', 'd');
+--------------------------------+
|mask(AbCD123-@$#, Q, q, d, NULL)|
+--------------------------------+
|                     QqQQddd-@$#|
+--------------------------------+
SELECT mask('AbCD123-@$#', 'Q', 'q', 'd', 'o');
+-----------------------------+
|mask(AbCD123-@$#, Q, q, d, o)|
+-----------------------------+
|                  QqQQdddoooo|
+-----------------------------+
SELECT mask('AbCD123-@$#', NULL, 'q', 'd', 'o');
+--------------------------------+
|mask(AbCD123-@$#, NULL, q, d, o)|
+--------------------------------+
|                     AqCDdddoooo|
+--------------------------------+
SELECT mask('AbCD123-@$#', NULL, NULL, 'd', 'o');
+-----------------------------------+
|mask(AbCD123-@$#, NULL, NULL, d, o)|
+-----------------------------------+
|                        AbCDdddoooo|
+-----------------------------------+
```

```
SELECT mask('AbCD123-@$#', NULL, NULL, NULL, 'o');
+-------------------------------------+
|mask(AbCD123-@$#, NULL, NULL, NULL, o)|
+-------------------------------------+
|                           AbCD123oooo|
+-------------------------------------+
SELECT mask(NULL, NULL, NULL, NULL, 'o');
+-------------------------------+
|mask(NULL, NULL, NULL, NULL, o)|
+-------------------------------+
|                           NULL|
+-------------------------------+
SELECT mask(NULL);
+-------------------------+
|mask(NULL, X, x, n, NULL)|
+-------------------------+
|                     NULL|
+-------------------------+
SELECT mask('AbCD123-@$#', NULL, NULL, NULL, NULL);
+----------------------------------------+
|mask(AbCD123-@$#, NULL, NULL, NULL, NULL)|
+----------------------------------------+
|                             AbCD123-@$#|
+----------------------------------------+
-- octet_length
SELECT octet_length('Feathers');
+-----------------------+
|octet_length(Feathers)|
+-----------------------+
|                      8|
+-----------------------+
SELECT octet_length(x'537061726b2053514c');
+-----------------------------------+
|octet_length(X'537061726B2053514C')|
+-----------------------------------+
|                                  9|
+-----------------------------------+
-- overlay
SELECT overlay('Feathers' PLACING '_' FROM 6);
+----------------------------+
|overlay(Feathers, _, 6, -1)|
+----------------------------+
|                  Feathe_ers|
+----------------------------+
```

```
SELECT overlay('Feathers' PLACING 'ures' FROM 5);
+------------------------------+
|overlay(Feathers, ures, 5, -1)|
+------------------------------+
|                     Features  |
+------------------------------+
-- position
SELECT position('bar', 'foobarbar');
+--------------------------+
|position(bar, foobarbar, 1)|
+--------------------------+
|                         4|
+--------------------------+
SELECT position('bar', 'foobarbar', 5);
+--------------------------+
|position(bar, foobarbar, 5)|
+--------------------------+
|                         7|
+--------------------------+
SELECT POSITION('bar' IN 'foobarbar');
+-------------------------+
|locate(bar, foobarbar, 1)|
+-------------------------+
|                        4|
+-------------------------+
-- printf
SELECT printf("Hello World %d %s", 100, "days");
+-----------------------------------+
|printf(Hello World %d %s, 100, days)|
+-----------------------------------+
|                Hello World 100 days|
+-----------------------------------+
-- regexp_count
SELECT regexp_count('Steven Jones and Stephen Smith are the best players', 'Ste(v|
ph)en');
+-----------------------------------------------------------------------------+
|regexp_count(Steven Jones and Stephen Smith are the best players, Ste(v|ph)en)|
+-----------------------------------------------------------------------------+
|                                                                            2|
+-----------------------------------------------------------------------------+
SELECT regexp_count('abcdefghijklmnopqrstuvwxyz', '[a-z]{3}');
+-------------------------------------------------+
|regexp_count(abcdefghijklmnopqrstuvwxyz, [a-z]{3})|
+-------------------------------------------------+
```

```
|                                                  8|
+---------------------------------------------------+
-- regexp_extract
SELECT regexp_extract('100-200', '(\\d+)-(\\d+)', 1);
+--------------------------------------+
|regexp_extract(100-200, (\d+)-(\d+), 1)|
+--------------------------------------+
|                                   100|
+--------------------------------------+
-- regexp_extract_all
SELECT regexp_extract_all('100-200, 300-400', '(\\d+)-(\\d+)', 1);
+-----------------------------------------------------+
|regexp_extract_all(100-200, 300-400, (\d+)-(\d+), 1)|
+-----------------------------------------------------+
|                                          [100, 300]|
+-----------------------------------------------------+
-- regexp_instr
SELECT regexp_instr('user@opensearch.org', '@[^.]*');
+-----------------------------------------------+
|regexp_instr(user@opensearch.org, @[^.]*, 0)|
+-----------------------------------------------+
|                                              5|
+-----------------------------------------------+
-- regexp_replace
SELECT regexp_replace('100-200', '(\\d+)', 'num');
+--------------------------------------+
|regexp_replace(100-200, (\d+), num, 1)|
+--------------------------------------+
|                               num-num|
+--------------------------------------+
-- regexp_substr
SELECT regexp_substr('Steven Jones and Stephen Smith are the best players', 'Ste(v|
ph)en');
+-----------------------------------------------------------------------------------+
|regexp_substr(Steven Jones and Stephen Smith are the best players, Ste(v|ph)en)|
+-----------------------------------------------------------------------------------+
|                                                                             Steven|
+-----------------------------------------------------------------------------------+
SELECT regexp_substr('Steven Jones and Stephen Smith are the best players', 'Jeck');
+-----------------------------------------------------------------------+
|regexp_substr(Steven Jones and Stephen Smith are the best players, Jeck)|
+-----------------------------------------------------------------------+
|                                                                   NULL|
+-----------------------------------------------------------------------+
```

```
-- repeat
SELECT repeat('123', 2);
+--------------+
|repeat(123, 2)|
+--------------+
|        123123|
+--------------+
-- replace
SELECT replace('ABCabc', 'abc', 'DEF');
+-------------------------+
|replace(ABCabc, abc, DEF)|
+-------------------------+
|                   ABCDEF|
+-------------------------+
-- right
SELECT right('Feathers', 3);
+-------------------+
|right(Feathers, 3)|
+-------------------+
|                ers|
+-------------------+
-- rpad
SELECT rpad('hi', 5, '??');
+---------------+
|rpad(hi, 5, ??)|
+---------------+
|          hi???|
+---------------+
SELECT rpad('hi', 1, '??');
+---------------+
|rpad(hi, 1, ??)|
+---------------+
|              h|
+---------------+
SELECT rpad('hi', 5);
+--------------+
|rpad(hi, 5,  )|
+--------------+
|         hi   |
+--------------+
SELECT hex(rpad(unhex('aabb'), 5));
+--------------------------------+
|hex(rpad(unhex(aabb), 5, X'00'))|
+--------------------------------+
```

```
|                      AABB000000|
+--------------------------------+
SELECT hex(rpad(unhex('aabb'), 5, unhex('1122')));
+-------------------------------------+
|hex(rpad(unhex(aabb), 5, unhex(1122)))|
+-------------------------------------+
|                           AABB112211|
+-------------------------------------+
-- rtrim
SELECT rtrim('    Feathers   ');
+----------------------+
|rtrim(    Feathers   )|
+----------------------+
|              Feathers|
+----------------------+
-- sentences
SELECT sentences('Hi there! Good morning.');
+-------------------------------------+
|sentences(Hi there! Good morning., , )|
+-------------------------------------+
|                  [[Hi, there], [Go...|
+-------------------------------------+
-- soundex
SELECT soundex('Miller');
+---------------+
|soundex(Miller)|
+---------------+
|           M460|
+---------------+
-- space
SELECT concat(space(2), '1');
+-------------------+
|concat(space(2), 1)|
+-------------------+
|                  1|
+-------------------+
-- split
SELECT split('oneAtwoBthreeC', '[ABC]');
+--------------------------------+
|split(oneAtwoBthreeC, [ABC], -1)|
+--------------------------------+
|             [one, two, three, ]|
+--------------------------------+
SELECT split('oneAtwoBthreeC', '[ABC]', -1);
```

```
+-------------------------------+
|split(oneAtwoBthreeC, [ABC], -1)|
+-------------------------------+
|             [one, two, three, ]|
+-------------------------------+
SELECT split('oneAtwoBthreeC', '[ABC]', 2);
+-------------------------------+
|split(oneAtwoBthreeC, [ABC], 2)|
+-------------------------------+
|              [one, twoBthreeC]|
+-------------------------------+
-- split_part
SELECT split_part('11.12.13', '.', 3);
+--------------------------+
|split_part(11.12.13, ., 3)|
+--------------------------+
|                        13|
+--------------------------+
-- startswith
SELECT startswith('Feathers', 'Fea');
+---------------------------+
|startswith(Feathers, Fea)|
+---------------------------+
|                       true|
+---------------------------+
SELECT startswith('Feathers', 'SQL');
+-------------------------+
|startswith(Feathers, SQL)|
+-------------------------+
|                     false|
+-------------------------+
SELECT startswith('Feathers', null);
+-------------------------+
|startswith(Feathers, NULL)|
+-------------------------+
|                      NULL|
+-------------------------+
SELECT startswith(x'537061726b2053514c', x'537061726b');
+-------------------------------------------------+
|startswith(X'537061726B2053514C', X'537061726B')|
+-------------------------------------------------+
|                                             true|
+-------------------------------------------------+
SELECT startswith(x'537061726b2053514c', x'53514c');
```

```
+-----------------------------------------+
|startswith(X'537061726B2053514C', X'53514C')|
+-----------------------------------------+
|                                    false|
+-----------------------------------------+
-- substr
SELECT substr('Feathers', 5);
+--------------------------------+
|substr(Feathers, 5, 2147483647)|
+--------------------------------+
|                           hers |
+--------------------------------+
SELECT substr('Feathers', -3);
+---------------------------------+
|substr(Feathers, -3, 2147483647)|
+---------------------------------+
|                              ers|
+---------------------------------+
SELECT substr('Feathers', 5, 1);
+-----------------------+
|substr(Feathers, 5, 1)|
+-----------------------+
|                      h|
+-----------------------+
SELECT substr('Feathers' FROM 5);
+----------------------------------+
|substring(Feathers, 5, 2147483647)|
+----------------------------------+
|                              hers |
+----------------------------------+
SELECT substr('Feathers' FROM -3);
+-----------------------------------+
|substring(Feathers, -3, 2147483647)|
+-----------------------------------+
|                                ers|
+-----------------------------------+
SELECT substr('Feathers' FROM 5 FOR 1);
+-------------------------+
|substring(Feathers, 5, 1)|
+-------------------------+
|                        h|
+-------------------------+
-- substring
SELECT substring('Feathers', 5);
```

```
+-----------------------------------+
|substring(Feathers, 5, 2147483647)|
+-----------------------------------+
|                              hers |
+-----------------------------------+
SELECT substring('Feathers', -3);
+-----------------------------------+
|substring(Feathers, -3, 2147483647)|
+-----------------------------------+
|                                ers|
+-----------------------------------+
SELECT substring('Feathers', 5, 1);
+-------------------------+
|substring(Feathers, 5, 1)|
+-------------------------+
|                        h|
+-------------------------+
SELECT substring('Feathers' FROM 5);
+-----------------------------------+
|substring(Feathers, 5, 2147483647)|
+-----------------------------------+
|                              hers |
+-----------------------------------+
SELECT substring('Feathers' FROM -3);
+-----------------------------------+
|substring(Feathers, -3, 2147483647)|
+-----------------------------------+
|                                ers|
+-----------------------------------+
SELECT substring('Feathers' FROM 5 FOR 1);
+-------------------------+
|substring(Feathers, 5, 1)|
+-------------------------+
|                        h|
+-------------------------+
-- substring_index
SELECT substring_index('www.apache.org', '.', 2);
+-----------------------------------+
|substring_index(www.apache.org, ., 2)|
+-----------------------------------+
|                          www.apache|
+-----------------------------------+
-- to_binary
SELECT to_binary('abc', 'utf-8');
```

```
+---------------------+
|to_binary(abc, utf-8)|
+---------------------+
|           [61 62 63]|
+---------------------+
-- to_char
SELECT to_char(454, '999');
+-----------------+
|to_char(454, 999)|
+-----------------+
|              454|
+-----------------+
SELECT to_char(454.00, '000D00');
+-----------------------+
|to_char(454.00, 000D00)|
+-----------------------+
|                 454.00|
+-----------------------+
SELECT to_char(12454, '99G999');
+----------------------+
|to_char(12454, 99G999)|
+----------------------+
|                12,454|
+----------------------+
SELECT to_char(78.12, '$99.99');
+----------------------+
|to_char(78.12, $99.99)|
+----------------------+
|                $78.12|
+----------------------+
SELECT to_char(-12454.8, '99G999D9S');
+----------------------------+
|to_char(-12454.8, 99G999D9S)|
+----------------------------+
|                   12,454.8-|
+----------------------------+
-- to_number
SELECT to_number('454', '999');
+-------------------+
|to_number(454, 999)|
+-------------------+
|                454|
+-------------------+
SELECT to_number('454.00', '000.00');
```

```
+-------------------------+
|to_number(454.00, 000.00)|
+-------------------------+
|                   454.00|
+-------------------------+
SELECT to_number('12,454', '99,999');
+-------------------------+
|to_number(12,454, 99,999)|
+-------------------------+
|                    12454|
+-------------------------+
SELECT to_number('$78.12', '$99.99');
+-------------------------+
|to_number($78.12, $99.99)|
+-------------------------+
|                    78.12|
+-------------------------+
SELECT to_number('12,454.8-', '99,999.9S');
+-------------------------------+
|to_number(12,454.8-, 99,999.9S)|
+-------------------------------+
|                       -12454.8|
+-------------------------------+
-- to_varchar
SELECT to_varchar(454, '999');
+-----------------+
|to_char(454, 999)|
+-----------------+
|              454|
+-----------------+
SELECT to_varchar(454.00, '000D00');
+-----------------------+
|to_char(454.00, 000D00)|
+-----------------------+
|                 454.00|
+-----------------------+
SELECT to_varchar(12454, '99G999');
+----------------------+
|to_char(12454, 99G999)|
+----------------------+
|                12,454|
+----------------------+
SELECT to_varchar(78.12, '$99.99');
+----------------------+
```

```
|to_char(78.12, $99.99)|
+----------------------+
|               $78.12|
+----------------------+
SELECT to_varchar(-12454.8, '99G999D9S');
+---------------------------+
|to_char(-12454.8, 99G999D9S)|
+---------------------------+
|                  12,454.8-|
+---------------------------+
-- translate
SELECT translate('AaBbCc', 'abc', '123');
+--------------------------+
|translate(AaBbCc, abc, 123)|
+--------------------------+
|                    A1B2C3|
+--------------------------+
-- try_to_binary
SELECT try_to_binary('abc', 'utf-8');
+-------------------------+
|try_to_binary(abc, utf-8)|
+-------------------------+
|              [61 62 63]|
+-------------------------+
select try_to_binary('a!', 'base64');
+-------------------------+
|try_to_binary(a!, base64)|
+-------------------------+
|                     NULL|
+-------------------------+
select try_to_binary('abc', 'invalidFormat');
+---------------------------------+
|try_to_binary(abc, invalidFormat)|
+---------------------------------+
|                             NULL|
+---------------------------------+
-- try_to_number
SELECT try_to_number('454', '999');
+----------------------+
|try_to_number(454, 999)|
+----------------------+
|                   454|
+----------------------+
SELECT try_to_number('454.00', '000.00');
```

```
+-----------------------------+
|try_to_number(454.00, 000.00)|
+-----------------------------+
|                       454.00|
+-----------------------------+
SELECT try_to_number('12,454', '99,999');
+-----------------------------+
|try_to_number(12,454, 99,999)|
+-----------------------------+
|                        12454|
+-----------------------------+
SELECT try_to_number('$78.12', '$99.99');
+-----------------------------+
|try_to_number($78.12, $99.99)|
+-----------------------------+
|                        78.12|
+-----------------------------+
SELECT try_to_number('12,454.8-', '99,999.9S');
+-----------------------------------+
|try_to_number(12,454.8-, 99,999.9S)|
+-----------------------------------+
|                           -12454.8|
+-----------------------------------+
-- ucase
SELECT ucase('Feathers');
+---------------+
|ucase(Feathers)|
+---------------+
|       FEATHERS|
+---------------+
-- unbase64
SELECT unbase64('U3BhcmsgU1FM');
+----------------------+
|unbase64(U3BhcmsgU1FM)|
+----------------------+
|  [53 70 61 72 6B 2...|
+----------------------+
-- upper
SELECT upper('Feathers');
+---------------+
|upper(Feathers)|
+---------------+
|       FEATHERS|
```

```
+---------------+
```

Funzioni di data e ora

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Funzione	Description
add_months (data_iniziale, num_months)	Restituisce la data successiva. `num_months` `start_date`
convert_timezone ([SourceTz,] targetTZ, SourceTs)	Converte il timestamp senza fuso orario dal fuso orario a. `sourceTs` `sourceTz` `targetTz`
curdata ()	Restituisce la data corrente all'inizio della valutazione della query. Tutte le chiamate di curdate all'interno della stessa query restituis cono lo stesso valore.
current_date ()	Restituisce la data corrente all'inizio della valutazione della query. Tutte le chiamate di current_date all'interno della stessa query restituiscono lo stesso valore.
data_corrente	Restituisce la data corrente all'inizio della valutazione della query.
current_timestamp ()	Restituisce il timestamp corrente all'inizio della valutazione della query. Tutte le chiamate di current_timestamp all'interno della stessa query restituiscono lo stesso valore.

Funzione	Description
current_timestamp	Restituisce il timestamp corrente all'inizio della valutazione della query.
current_timezone ()	Restituisce il fuso orario locale della sessione corrente.
date_add (data_iniziale, num_giorni)	Restituisce la data successiva. `num_days` `start_date`
date_diff (DataFine, DataInizio)	Restituisce il numero di giorni compreso tra a. `startDate endDate`
date_format (timestamp, fmt)	Converte in un valore `timestamp` di stringa nel formato specificato dal formato della data. `fmt`
date_from_unix_date (giorni)	Crea la data a partire dal numero di giorni dal 01/01/1970.
date_part (campo, fonte)	Estrae una parte della sorgente o dell' date/timestamp intervallo.
date_sub (data_iniziale, num_giorni)	Restituisce la data precedente`num_days`. `start_date`
date_trunc (fmt, ts)	Restituisce il timestamp `ts` troncato all'unità specificata dal modello di formato. `fmt`
dateadd (data_iniziale, num_days)	Restituisce la data successiva. `num_days` `start_date`
datediff (DataFine, DataAvvio)	Restituisce il numero di giorni da a. `startDate endDate`
datepart (campo, fonte)	Estrae una parte della sorgente date/time stamp o dell'intervallo.
giorno (data)	Restituisce il giorno del mese della data/ora.

Funzione	Description
giornodel mese (data)	Restituisce il giorno del mese della data/ora.
giornodella settimana (data)	Restituisce il giorno della settimana per date/timestamp (1 = domenica, 2 = lunedì,..., 7 = sabato).
giornodel'anno (data)	Restituisce il giorno dell'anno della data/ora.
estratto (campo DALLA fonte)	Estrae una parte della sorgente date/time stamp o dell'intervallo.
from_unixtime (unix_time [, fmt])	`unix_time` Restituisce nel formato `fmt` specificato.
from_utc_timestamp (timestamp, fuso orario)	Dato un timestamp come '2017-07-14 02:40:00.0 ', lo interpreta come un orario in UTC e visualizza quell'ora come un timestamp nel fuso orario specificato. Ad esempio, 'GMT +1' restituirebbe '2017-07-14 03:40:00.0 '.
ora (timestamp)	Restituisce il componente orario della stringa/timestamp.
last_day (data)	Restituisce l'ultimo giorno del mese a cui appartiene la data.
timestamp locale ()	Restituisce il timestamp corrente senza fuso orario all'inizio della valutazione della query. Tutte le chiamate di localtimestamp all'interno della stessa query restituiscono lo stesso valore.
timestamp locale	Restituisce la data-ora locale corrente nel fuso orario della sessione all'inizio della valutazione della query.

Funzione	Description
make_date (anno, mese, giorno)	Crea data dai campi relativi all'anno, al mese e al giorno.
make_dt_interval ([giorni [, ore [, minuti [, secondi]]]])	Calcola la DayTimeIntervalType durata in giorni, ore, minuti e secondi.
make_interval ([anni [, mesi [, settimane [, giorni [, ore [, minuti [, secondi]]]]]]])	Crea un intervallo tra anni, mesi, settimane, giorni, ore, minuti e secondi.
make_timestamp (anno, mese, giorno, ora, min, sec [, fuso orario])	Crea timestamp dai campi anno, mese, giorno, ora, min, sec e fuso orario.
make_timestamp_ltz (anno, mese, giorno, ora, min, sec [, fuso orario])	Crea il timestamp corrente con il fuso orario locale utilizzando i campi anno, mese, giorno, ora, min, sec e fuso orario.
make_timestamp_ntz (anno, mese, giorno, ora, min, sec)	Crea data-ora locale dai campi anno, mese, giorno, ora, min, sec.
make_ym_interval ([anni [, mesi]])	Crea un intervallo anno-mese tra anni, mesi.
minuto (timestamp)	Restituisce il componente minuto della stringa/timestamp.
mese (data)	Restituisce il componente mensile della data/ora.
months_between (timestamp1, timestamp2 [, RoundOff])	Se è `timestamp1` successivo a, `timestamp2` allora il risultato è positivo. Se `timestamp1` e `timestamp2` sono nello stesso giorno del mese, o entrambi sono l'ultimo giorno del mese, l'ora del giorno verrà ignorata. Altrimenti, la differenza viene calcolata in base a 31 giorni al mese e arrotondata a 8 cifre, a meno che RoundOff=False.

Funzione	Description
giorno_successivo (data_inizio, giorno_della settimana)	Restituisce la prima data successiva e denominata come indicato. `start_date` La funzione restituisce NULL se almeno uno dei parametri di input è NULL.
adesso ()	Restituisce il timestamp corrente all'inizio della valutazione della query.
trimestre (data)	Restituisce il trimestre dell'anno per la data, nell'intervallo da 1 a 4.
secondo (timestamp)	Restituisce il secondo componente della stringa/timestamp.
finestra_sessione (time_column, gap_duration)	Genera una finestra di sessione con un timestamp che specifica la durata della colonna e del gap. Vedi «Tipi di finestre temporali» nel documento della guida Structure d Streaming per spiegazioni dettagliate ed esempi.
timestamp_micros (microsecondi)	Crea un timestamp a partire dal numero di microsecondi dall'epoca UTC.
timestamp_millis (millisecondi)	Crea un timestamp in base al numero di millisecondi trascorsi dall'epoca UTC.
timestamp_seconds (secondi)	Crea il timestamp in base al numero di secondi (può essere frazionario) dall'epoca UTC.
to_date (date_str [, fmt])	Analizza l'espressione con l'`date_str`espressione fino a una data. `fmt` Restituisce null con un input non valido. Per impostazione predefinita, segue le regole di casting fino a una data se `fmt` viene omesso.

Funzione	Description
to_timestamp (timestamp_str [, fmt])	Analizza l'espressione con l'espressione in un timestamp. `timestamp_str` `fmt` Restituisce null con un input non valido. Per impostazione predefinita, segue le regole di casting in base a un timestamp se viene omesso. `fmt`
to_timestamp_ltz (timestamp_str [, fmt])	Analizza l'espressione con l'espressione in un timestamp con fuso orario locale. `timestamp_str` `fmt` Restituisce null con un input non valido. Per impostazione predefinita, segue le regole di casting in base a un timestamp se viene omesso. `fmt`
to_timestamp_ntz (timestamp_str [, fmt])	Analizza l'espressione con l'espressione in un timestamp senza fuso orario. `timestamp_str` `fmt` Restituisce null con un input non valido. Per impostazione predefinita, segue le regole di casting in base a un timestamp se viene omesso. `fmt`
to_unix_timestamp (timeExp [, fmt])	Restituisce il timestamp UNIX dell'ora specifica ta.
to_utc_timestamp (timestamp, timezone)	Dato un timestamp come '2017-07-14 02:40:00.0 ', lo interpreta come un orario nel fuso orario specificato e visualizza quell'ora come un timestamp in UTC. Ad esempio, 'GMT +1' restituirebbe '2017-07-14 01:40:00.0 '.
trunc (data, fmt)	Restituisce `date` con la parte temporale del giorno troncata all'unità specificata dal modello di formato. `fmt`
try_to_timestamp (timestamp_str [, fmt])	Analizza l'espressione con l'espressione in `timestamp_str` un timestamp. `fmt`

Funzione	Description
unix_date (data)	Restituisce il numero di giorni dal 01/01/1970.
unix_micros (timestamp)	Restituisce il numero di microsecondi dal 01/01/1970 00:00:00 UTC.
unix_millis (timestamp)	Restituisce il numero di millisecondi dal 01/01/1970 alle 00:00:00 UTC. Tronca i livelli di precisione più elevati.
unix_seconds (timestamp)	Restituisce il numero di secondi dal 01/01/1970 alle 00:00:00 UTC. Tronca i livelli di precisione più elevati.
unix_timestamp ([timeExp [, fmt]])	Restituisce il timestamp UNIX dell'ora corrente o specificata.
giorno della settimana (data)	Restituisce il giorno della settimana per date/timestamp (0 = lunedì, 1 = martedì,..., 6 = domenica).
settimana dell'anno (data)	Restituisce la settimana dell'anno della data specificata. Si considera che una settimana inizi di lunedì e la settimana 1 è la prima settimana con più di 3 giorni.
finestra (time_column, window_duration [, slide_duration [, start_time]])	Suddividi le righe in una o più finestre temporali con un timestamp che specifica una colonna. Gli inizi delle finestre sono inclusi ma le estremità delle finestre sono esclusive, ad esempio 12:05 sarà nella finestra [12:05,12 :10) ma non in [12:00,12:05). Windows può supportare la precisione in microsecondi. Windows nell'ordine dei mesi non è supportato. Vedi «Window Operations on Event Time» nel documento della guida Structured Streaming per spiegazioni dettagliate ed esempi.

Funzione	Description
window_time (window_column)	Estrai il valore time/session temporale dalla colonna della finestra che può essere utilizzat a per il valore temporale dell'evento della finestra. L'ora estratta è (window.end - 1) e riflette il fatto che le finestre di aggregazione hanno un limite superiore esclusivo - [start, end) Vedi 'Window Operations on Event Time' nel documento guida Structured Streaming per spiegazioni dettagliate ed esempi.
anno (data)	Restituisce il componente annuale del data/ timestamp.

Esempi

```
-- add_months
SELECT add_months('2016-08-31', 1);
+-------------------------+
|add_months(2016-08-31, 1)|
+-------------------------+
|               2016-09-30|
+-------------------------+
-- convert_timezone
SELECT convert_timezone('Europe/Brussels', 'America/Los_Angeles',
 timestamp_ntz'2021-12-06 00:00:00');
+-----------------------------------------------------------------------------------------------
+
|convert_timezone(Europe/Brussels, America/Los_Angeles, TIMESTAMP_NTZ '2021-12-06
 00:00:00')|
+-----------------------------------------------------------------------------------------------
+
|                                                                                2021-12-05
 15:00:00|
+-----------------------------------------------------------------------------------------------
+
SELECT convert_timezone('Europe/Brussels', timestamp_ntz'2021-12-05 15:00:00');
+-----------------------------------------------------------------------------------------------
+
```

```
|convert_timezone(current_timezone(), Europe/Brussels, TIMESTAMP_NTZ '2021-12-05
 15:00:00')|
+-----------------------------------------------------------------------------------------------
+
|                                                                                2021-12-05
 07:00:00|
+-----------------------------------------------------------------------------------------------
+
-- curdate
SELECT curdate();
+--------------+
|current_date()|
+--------------+
|    2024-02-24|
+--------------+
-- current_date
SELECT current_date();
+--------------+
|current_date()|
+--------------+
|    2024-02-24|
+--------------+
SELECT current_date;
+--------------+
|current_date()|
+--------------+
|    2024-02-24|
+--------------+
-- current_timestamp
SELECT current_timestamp();
+--------------------+
| current_timestamp()|
+--------------------+
|2024-02-24 16:36:...|
+--------------------+
SELECT current_timestamp;
+--------------------+
| current_timestamp()|
+--------------------+
|2024-02-24 16:36:...|
+--------------------+
-- current_timezone
SELECT current_timezone();
+------------------+
```

```
|current_timezone()|
+------------------+
|        Asia/Seoul|
+------------------+
-- date_add
SELECT date_add('2016-07-30', 1);
+-----------------------+
|date_add(2016-07-30, 1)|
+-----------------------+
|             2016-07-31|
+-----------------------+
-- date_diff
SELECT date_diff('2009-07-31', '2009-07-30');
+---------------------------------+
|date_diff(2009-07-31, 2009-07-30)|
+---------------------------------+
|                                1|
+---------------------------------+
SELECT date_diff('2009-07-30', '2009-07-31');
+---------------------------------+
|date_diff(2009-07-30, 2009-07-31)|
+---------------------------------+
|                               -1|
+---------------------------------+
-- date_format
SELECT date_format('2016-04-08', 'y');
+--------------------------+
|date_format(2016-04-08, y)|
+--------------------------+
|                      2016|
+--------------------------+
-- date_from_unix_date
SELECT date_from_unix_date(1);
+----------------------+
|date_from_unix_date(1)|
+----------------------+
|            1970-01-02|
+----------------------+
-- date_part
SELECT date_part('YEAR', TIMESTAMP '2019-08-12 01:00:00.123456');
+-------------------------------------------------------+
|date_part(YEAR, TIMESTAMP '2019-08-12 01:00:00.123456')|
+-------------------------------------------------------+
|                                                   2019|
```

```
+-----------------------------------------------------+
SELECT date_part('week', timestamp'2019-08-12 01:00:00.123456');
+-----------------------------------------------------+
|date_part(week, TIMESTAMP '2019-08-12 01:00:00.123456')|
+-----------------------------------------------------+
|                                                   33|
+-----------------------------------------------------+
SELECT date_part('doy', DATE'2019-08-12');
+--------------------------------+
|date_part(doy, DATE '2019-08-12')|
+--------------------------------+
|                             224|
+--------------------------------+
SELECT date_part('SECONDS', timestamp'2019-10-01 00:00:01.000001');
+------------------------------------------------------------+
|date_part(SECONDS, TIMESTAMP '2019-10-01 00:00:01.000001')|
+------------------------------------------------------------+
|                                                  1.000001|
+------------------------------------------------------------+
SELECT date_part('days', interval 5 days 3 hours 7 minutes);
+-------------------------------------------------+
|date_part(days, INTERVAL '5 03:07' DAY TO MINUTE)|
+-------------------------------------------------+
|                                                5|
+-------------------------------------------------+
SELECT date_part('seconds', interval 5 hours 30 seconds 1 milliseconds 1 microseconds);
+-------------------------------------------------------------+
|date_part(seconds, INTERVAL '05:00:30.001001' HOUR TO SECOND)|
+-------------------------------------------------------------+
|                                                    30.001001|
+-------------------------------------------------------------+
SELECT date_part('MONTH', INTERVAL '2021-11' YEAR TO MONTH);
+-------------------------------------------------+
|date_part(MONTH, INTERVAL '2021-11' YEAR TO MONTH)|
+-------------------------------------------------+
|                                                11|
+-------------------------------------------------+
SELECT date_part('MINUTE', INTERVAL '123 23:55:59.002001' DAY TO SECOND);
+---------------------------------------------------------------+
|date_part(MINUTE, INTERVAL '123 23:55:59.002001' DAY TO SECOND)|
+---------------------------------------------------------------+
|                                                             55|
+---------------------------------------------------------------+
-- date_sub
```

```
SELECT date_sub('2016-07-30', 1);
+-----------------------+
|date_sub(2016-07-30, 1)|
+-----------------------+
|             2016-07-29|
+-----------------------+
-- date_trunc
SELECT date_trunc('YEAR', '2015-03-05T09:32:05.359');
+----------------------------------------+
|date_trunc(YEAR, 2015-03-05T09:32:05.359)|
+----------------------------------------+
|                     2015-01-01 00:00:00|
+----------------------------------------+
SELECT date_trunc('MM', '2015-03-05T09:32:05.359');
+---------------------------------------+
|date_trunc(MM, 2015-03-05T09:32:05.359)|
+---------------------------------------+
|                    2015-03-01 00:00:00|
+---------------------------------------+
SELECT date_trunc('DD', '2015-03-05T09:32:05.359');
+---------------------------------------+
|date_trunc(DD, 2015-03-05T09:32:05.359)|
+---------------------------------------+
|                    2015-03-05 00:00:00|
+---------------------------------------+
SELECT date_trunc('HOUR', '2015-03-05T09:32:05.359');
+----------------------------------------+
|date_trunc(HOUR, 2015-03-05T09:32:05.359)|
+----------------------------------------+
|                     2015-03-05 09:00:00|
+----------------------------------------+
SELECT date_trunc('MILLISECOND', '2015-03-05T09:32:05.123456');
+-----------------------------------------------------+
|date_trunc(MILLISECOND, 2015-03-05T09:32:05.123456)|
+-----------------------------------------------------+
|                              2015-03-05 09:32:...|
+-----------------------------------------------------+
-- dateadd
SELECT dateadd('2016-07-30', 1);
+-----------------------+
|date_add(2016-07-30, 1)|
+-----------------------+
|             2016-07-31|
+-----------------------+
```

```
-- datediff
SELECT datediff('2009-07-31', '2009-07-30');
+--------------------------------+
|datediff(2009-07-31, 2009-07-30)|
+--------------------------------+
|                               1|
+--------------------------------+
SELECT datediff('2009-07-30', '2009-07-31');
+--------------------------------+
|datediff(2009-07-30, 2009-07-31)|
+--------------------------------+
|                              -1|
+--------------------------------+
-- datepart
SELECT datepart('YEAR', TIMESTAMP '2019-08-12 01:00:00.123456');
+-----------------------------------------------------------+
|datepart(YEAR FROM TIMESTAMP '2019-08-12 01:00:00.123456')|
+-----------------------------------------------------------+
|                                                       2019|
+-----------------------------------------------------------+
SELECT datepart('week', timestamp'2019-08-12 01:00:00.123456');
+-----------------------------------------------------------+
|datepart(week FROM TIMESTAMP '2019-08-12 01:00:00.123456')|
+-----------------------------------------------------------+
|                                                         33|
+-----------------------------------------------------------+
SELECT datepart('doy', DATE'2019-08-12');
+-----------------------------------+
|datepart(doy FROM DATE '2019-08-12')|
+-----------------------------------+
|                                224|
+-----------------------------------+
SELECT datepart('SECONDS', timestamp'2019-10-01 00:00:01.000001');
+-------------------------------------------------------------+
|datepart(SECONDS FROM TIMESTAMP '2019-10-01 00:00:01.000001')|
+-------------------------------------------------------------+
|                                                     1.000001|
+-------------------------------------------------------------+
SELECT datepart('days', interval 5 days 3 hours 7 minutes);
+---------------------------------------------------+
|datepart(days FROM INTERVAL '5 03:07' DAY TO MINUTE)|
+---------------------------------------------------+
|                                                  5|
+---------------------------------------------------+
```

```
SELECT datepart('seconds', interval 5 hours 30 seconds 1 milliseconds 1 microseconds);
+-------------------------------------------------------------------+
|datepart(seconds FROM INTERVAL '05:00:30.001001' HOUR TO SECOND)|
+-------------------------------------------------------------------+
|                                                          30.001001|
+-------------------------------------------------------------------+
SELECT datepart('MONTH', INTERVAL '2021-11' YEAR TO MONTH);
+-----------------------------------------------------+
|datepart(MONTH FROM INTERVAL '2021-11' YEAR TO MONTH)|
+-----------------------------------------------------+
|                                                   11|
+-----------------------------------------------------+
SELECT datepart('MINUTE', INTERVAL '123 23:55:59.002001' DAY TO SECOND);
+-------------------------------------------------------------------+
|datepart(MINUTE FROM INTERVAL '123 23:55:59.002001' DAY TO SECOND)|
+-------------------------------------------------------------------+
|                                                                 55|
+-------------------------------------------------------------------+
-- day
SELECT day('2009-07-30');
+---------------+
|day(2009-07-30)|
+---------------+
|             30|
+---------------+
-- dayofmonth
SELECT dayofmonth('2009-07-30');
+----------------------+
|dayofmonth(2009-07-30)|
+----------------------+
|                    30|
+----------------------+
-- dayofweek
SELECT dayofweek('2009-07-30');
+---------------------+
|dayofweek(2009-07-30)|
+---------------------+
|                    5|
+---------------------+
-- dayofyear
SELECT dayofyear('2016-04-09');
+---------------------+
|dayofyear(2016-04-09)|
+---------------------+
```

```
|                  100|
+---------------------+
-- extract
SELECT extract(YEAR FROM TIMESTAMP '2019-08-12 01:00:00.123456');
+-----------------------------------------------------------+
|extract(YEAR FROM TIMESTAMP '2019-08-12 01:00:00.123456')|
+-----------------------------------------------------------+
|                                                       2019|
+-----------------------------------------------------------+
SELECT extract(week FROM timestamp'2019-08-12 01:00:00.123456');
+-----------------------------------------------------------+
|extract(week FROM TIMESTAMP '2019-08-12 01:00:00.123456')|
+-----------------------------------------------------------+
|                                                         33|
+-----------------------------------------------------------+
SELECT extract(doy FROM DATE'2019-08-12');
+----------------------------------+
|extract(doy FROM DATE '2019-08-12')|
+----------------------------------+
|                               224|
+----------------------------------+
SELECT extract(SECONDS FROM timestamp'2019-10-01 00:00:01.000001');
+-------------------------------------------------------------+
|extract(SECONDS FROM TIMESTAMP '2019-10-01 00:00:01.000001')|
+-------------------------------------------------------------+
|                                                     1.000001|
+-------------------------------------------------------------+
SELECT extract(days FROM interval 5 days 3 hours 7 minutes);
+---------------------------------------------------+
|extract(days FROM INTERVAL '5 03:07' DAY TO MINUTE)|
+---------------------------------------------------+
|                                                  5|
+---------------------------------------------------+
SELECT extract(seconds FROM interval 5 hours 30 seconds 1 milliseconds 1 microseconds);
+-----------------------------------------------------------------+
|extract(seconds FROM INTERVAL '05:00:30.001001' HOUR TO SECOND)|
+-----------------------------------------------------------------+
|                                                        30.001001|
+-----------------------------------------------------------------+
SELECT extract(MONTH FROM INTERVAL '2021-11' YEAR TO MONTH);
+---------------------------------------------------+
|extract(MONTH FROM INTERVAL '2021-11' YEAR TO MONTH)|
+---------------------------------------------------+
|                                                  11|
```

```
+-----------------------------------------------------+
SELECT extract(MINUTE FROM INTERVAL '123 23:55:59.002001' DAY TO SECOND);
+-------------------------------------------------------------------+
|extract(MINUTE FROM INTERVAL '123 23:55:59.002001' DAY TO SECOND)|
+-------------------------------------------------------------------+
|                                                                 55|
+-------------------------------------------------------------------+
-- from_unixtime
SELECT from_unixtime(0, 'yyyy-MM-dd HH:mm:ss');
+-------------------------------------+
|from_unixtime(0, yyyy-MM-dd HH:mm:ss)|
+-------------------------------------+
|                  1970-01-01 09:00:00|
+-------------------------------------+
SELECT from_unixtime(0);
+-------------------------------------+
|from_unixtime(0, yyyy-MM-dd HH:mm:ss)|
+-------------------------------------+
|                  1970-01-01 09:00:00|
+-------------------------------------+
-- from_utc_timestamp
SELECT from_utc_timestamp('2016-08-31', 'Asia/Seoul');
+-----------------------------------------+
|from_utc_timestamp(2016-08-31, Asia/Seoul)|
+-----------------------------------------+
|                      2016-08-31 09:00:00|
+-----------------------------------------+
-- hour
SELECT hour('2009-07-30 12:58:59');
+-------------------------+
|hour(2009-07-30 12:58:59)|
+-------------------------+
|                       12|
+-------------------------+
-- last_day
SELECT last_day('2009-01-12');
+--------------------+
|last_day(2009-01-12)|
+--------------------+
|          2009-01-31|
+--------------------+
-- localtimestamp
SELECT localtimestamp();
+--------------------+
```

```
|    localtimestamp()|
+--------------------+
|2024-02-24 16:36:...|
+--------------------+
-- make_date
SELECT make_date(2013, 7, 15);
+----------------------+
|make_date(2013, 7, 15)|
+----------------------+
|            2013-07-15|
+----------------------+
SELECT make_date(2019, 7, NULL);
+------------------------+
|make_date(2019, 7, NULL)|
+------------------------+
|                    NULL|
+------------------------+
-- make_dt_interval
SELECT make_dt_interval(1, 12, 30, 01.001001);
+-----------------------------------+
|make_dt_interval(1, 12, 30, 1.001001)|
+-----------------------------------+
|                 INTERVAL '1 12:30...|
+-----------------------------------+
SELECT make_dt_interval(2);
+---------------------------------+
|make_dt_interval(2, 0, 0, 0.000000)|
+---------------------------------+
|               INTERVAL '2 00:00...|
+---------------------------------+
SELECT make_dt_interval(100, null, 3);
+---------------------------------------+
|make_dt_interval(100, NULL, 3, 0.000000)|
+---------------------------------------+
|                                   NULL|
+---------------------------------------+
-- make_interval
SELECT make_interval(100, 11, 1, 1, 12, 30, 01.001001);
+-----------------------------------------------+
|make_interval(100, 11, 1, 1, 12, 30, 1.001001)|
+-----------------------------------------------+
|                         100 years 11 mont...|
+-----------------------------------------------+
SELECT make_interval(100, null, 3);
```

```
+-----------------------------------------------+
|make_interval(100, NULL, 3, 0, 0, 0, 0.000000)|
+-----------------------------------------------+
|                                           NULL|
+-----------------------------------------------+
SELECT make_interval(0, 1, 0, 1, 0, 0, 100.000001);
+-----------------------------------------------+
|make_interval(0, 1, 0, 1, 0, 0, 100.000001)|
+-----------------------------------------------+
|                       1 months 1 days 1...|
+-----------------------------------------------+
-- make_timestamp
SELECT make_timestamp(2014, 12, 28, 6, 30, 45.887);
+-----------------------------------------------+
|make_timestamp(2014, 12, 28, 6, 30, 45.887)|
+-----------------------------------------------+
|                       2014-12-28 06:30:...|
+-----------------------------------------------+
SELECT make_timestamp(2014, 12, 28, 6, 30, 45.887, 'CET');
+-----------------------------------------------------+
|make_timestamp(2014, 12, 28, 6, 30, 45.887, CET)|
+-----------------------------------------------------+
|                             2014-12-28 14:30:...|
+-----------------------------------------------------+
SELECT make_timestamp(2019, 6, 30, 23, 59, 60);
+---------------------------------------+
|make_timestamp(2019, 6, 30, 23, 59, 60)|
+---------------------------------------+
|                     2019-07-01 00:00:00|
+---------------------------------------+
SELECT make_timestamp(2019, 6, 30, 23, 59, 1);
+--------------------------------------+
|make_timestamp(2019, 6, 30, 23, 59, 1)|
+--------------------------------------+
|                    2019-06-30 23:59:01|
+--------------------------------------+
SELECT make_timestamp(null, 7, 22, 15, 30, 0);
+-------------------------------------+
|make_timestamp(NULL, 7, 22, 15, 30, 0)|
+-------------------------------------+
|                                 NULL|
+-------------------------------------+
-- make_timestamp_ltz
SELECT make_timestamp_ltz(2014, 12, 28, 6, 30, 45.887);
```

```
+-----------------------------------------------+
|make_timestamp_ltz(2014, 12, 28, 6, 30, 45.887)|
+-----------------------------------------------+
|                           2014-12-28 06:30:...|
+-----------------------------------------------+
SELECT make_timestamp_ltz(2014, 12, 28, 6, 30, 45.887, 'CET');
+-------------------------------------------------------+
|make_timestamp_ltz(2014, 12, 28, 6, 30, 45.887, CET)|
+-------------------------------------------------------+
|                                2014-12-28 14:30:...|
+-------------------------------------------------------+
SELECT make_timestamp_ltz(2019, 6, 30, 23, 59, 60);
+-----------------------------------------+
|make_timestamp_ltz(2019, 6, 30, 23, 59, 60)|
+-----------------------------------------+
|                      2019-07-01 00:00:00|
+-----------------------------------------+
SELECT make_timestamp_ltz(null, 7, 22, 15, 30, 0);
+---------------------------------------+
|make_timestamp_ltz(NULL, 7, 22, 15, 30, 0)|
+---------------------------------------+
|                                   NULL|
+---------------------------------------+
-- make_timestamp_ntz
SELECT make_timestamp_ntz(2014, 12, 28, 6, 30, 45.887);
+-----------------------------------------------+
|make_timestamp_ntz(2014, 12, 28, 6, 30, 45.887)|
+-----------------------------------------------+
|                           2014-12-28 06:30:...|
+-----------------------------------------------+
SELECT make_timestamp_ntz(2019, 6, 30, 23, 59, 60);
+-----------------------------------------+
|make_timestamp_ntz(2019, 6, 30, 23, 59, 60)|
+-----------------------------------------+
|                      2019-07-01 00:00:00|
+-----------------------------------------+
SELECT make_timestamp_ntz(null, 7, 22, 15, 30, 0);
+---------------------------------------+
|make_timestamp_ntz(NULL, 7, 22, 15, 30, 0)|
+---------------------------------------+
|                                   NULL|
+---------------------------------------+
-- make_ym_interval
SELECT make_ym_interval(1, 2);
```

```
+----------------------+
|make_ym_interval(1, 2)|
+----------------------+
|  INTERVAL '1-2' YE...|
+----------------------+
SELECT make_ym_interval(1, 0);
+----------------------+
|make_ym_interval(1, 0)|
+----------------------+
|  INTERVAL '1-0' YE...|
+----------------------+
SELECT make_ym_interval(-1, 1);
+----------------------+
|make_ym_interval(-1, 1)|
+----------------------+
|   INTERVAL '-0-11' ...|
+----------------------+
SELECT make_ym_interval(2);
+----------------------+
|make_ym_interval(2, 0)|
+----------------------+
|  INTERVAL '2-0' YE...|
+----------------------+
-- minute
SELECT minute('2009-07-30 12:58:59');
+--------------------------+
|minute(2009-07-30 12:58:59)|
+--------------------------+
|                        58|
+--------------------------+
-- month
SELECT month('2016-07-30');
+-----------------+
|month(2016-07-30)|
+-----------------+
|                7|
+-----------------+
-- months_between
SELECT months_between('1997-02-28 10:30:00', '1996-10-30');
+-------------------------------------------------------+
|months_between(1997-02-28 10:30:00, 1996-10-30, true)|
+-------------------------------------------------------+
|                                             3.94959677|
+-------------------------------------------------------+
```

```
SELECT months_between('1997-02-28 10:30:00', '1996-10-30', false);
+----------------------------------------------------------+
|months_between(1997-02-28 10:30:00, 1996-10-30, false)|
+----------------------------------------------------------+
|                                  3.9495967741935485|
+----------------------------------------------------------+
-- next_day
SELECT next_day('2015-01-14', 'TU');
+------------------------+
|next_day(2015-01-14, TU)|
+------------------------+
|              2015-01-20|
+------------------------+
-- now
SELECT now();
+--------------------+
|               now()|
+--------------------+
|2024-02-24 16:36:...|
+--------------------+
-- quarter
SELECT quarter('2016-08-31');
+-------------------+
|quarter(2016-08-31)|
+-------------------+
|                  3|
+-------------------+
-- second
SELECT second('2009-07-30 12:58:59');
+---------------------------+
|second(2009-07-30 12:58:59)|
+---------------------------+
|                         59|
+---------------------------+
-- session_window
SELECT a, session_window.start, session_window.end, count(*) as cnt FROM VALUES ('A1',
 '2021-01-01 00:00:00'), ('A1', '2021-01-01 00:04:30'), ('A1', '2021-01-01 00:10:00'),
 ('A2', '2021-01-01 00:01:00') AS tab(a, b) GROUP by a, session_window(b, '5 minutes')
 ORDER BY a, start;
+---+-------------------+-------------------+---+
|  a|              start|                end|cnt|
+---+-------------------+-------------------+---+
| A1|2021-01-01 00:00:00|2021-01-01 00:09:30|  2|
| A1|2021-01-01 00:10:00|2021-01-01 00:15:00|  1|
```

```
| A2|2021-01-01 00:01:00|2021-01-01 00:06:00|  1|
+---+-------------------+-------------------+---+
SELECT a, session_window.start, session_window.end, count(*) as cnt FROM VALUES ('A1',
 '2021-01-01 00:00:00'), ('A1', '2021-01-01 00:04:30'), ('A1', '2021-01-01 00:10:00'),
 ('A2', '2021-01-01 00:01:00'), ('A2', '2021-01-01 00:04:30') AS tab(a, b) GROUP by a,
 session_window(b, CASE WHEN a = 'A1' THEN '5 minutes' WHEN a = 'A2' THEN '1 minute'
 ELSE '10 minutes' END) ORDER BY a, start;
+---+-------------------+-------------------+---+
|  a|              start|                end|cnt|
+---+-------------------+-------------------+---+
| A1|2021-01-01 00:00:00|2021-01-01 00:09:30|  2|
| A1|2021-01-01 00:10:00|2021-01-01 00:15:00|  1|
| A2|2021-01-01 00:01:00|2021-01-01 00:02:00|  1|
| A2|2021-01-01 00:04:30|2021-01-01 00:05:30|  1|
+---+-------------------+-------------------+---+
-- timestamp_micros
SELECT timestamp_micros(1230219000123123);
+----------------------------------+
|timestamp_micros(1230219000123123)|
+----------------------------------+
|              2008-12-26 00:30:...|
+----------------------------------+
-- timestamp_millis
SELECT timestamp_millis(1230219000123);
+-------------------------------+
|timestamp_millis(1230219000123)|
+-------------------------------+
|           2008-12-26 00:30:...|
+-------------------------------+
-- timestamp_seconds
SELECT timestamp_seconds(1230219000);
+-----------------------------+
|timestamp_seconds(1230219000)|
+-----------------------------+
|          2008-12-26 00:30:00|
+-----------------------------+
SELECT timestamp_seconds(1230219000.123);
+---------------------------------+
|timestamp_seconds(1230219000.123)|
+---------------------------------+
|             2008-12-26 00:30:...|
+---------------------------------+
-- to_date
SELECT to_date('2009-07-30 04:17:52');
```

```
+---------------------------+
|to_date(2009-07-30 04:17:52)|
+---------------------------+
|                 2009-07-30|
+---------------------------+
SELECT to_date('2016-12-31', 'yyyy-MM-dd');
+-------------------------------+
|to_date(2016-12-31, yyyy-MM-dd)|
+-------------------------------+
|                     2016-12-31|
+-------------------------------+
-- to_timestamp
SELECT to_timestamp('2016-12-31 00:12:00');
+---------------------------------+
|to_timestamp(2016-12-31 00:12:00)|
+---------------------------------+
|              2016-12-31 00:12:00|
+---------------------------------+
SELECT to_timestamp('2016-12-31', 'yyyy-MM-dd');
+-----------------------------------+
|to_timestamp(2016-12-31, yyyy-MM-dd)|
+-----------------------------------+
|                2016-12-31 00:00:00|
+-----------------------------------+
-- to_timestamp_ltz
SELECT to_timestamp_ltz('2016-12-31 00:12:00');
+-----------------------------------+
|to_timestamp_ltz(2016-12-31 00:12:00)|
+-----------------------------------+
|                 2016-12-31 00:12:00|
+-----------------------------------+
SELECT to_timestamp_ltz('2016-12-31', 'yyyy-MM-dd');
+---------------------------------------+
|to_timestamp_ltz(2016-12-31, yyyy-MM-dd)|
+---------------------------------------+
|                     2016-12-31 00:00:00|
+---------------------------------------+
-- to_timestamp_ntz
SELECT to_timestamp_ntz('2016-12-31 00:12:00');
+-----------------------------------+
|to_timestamp_ntz(2016-12-31 00:12:00)|
+-----------------------------------+
|                 2016-12-31 00:12:00|
+-----------------------------------+
```

```
SELECT to_timestamp_ntz('2016-12-31', 'yyyy-MM-dd');
+----------------------------------------+
|to_timestamp_ntz(2016-12-31, yyyy-MM-dd)|
+----------------------------------------+
|                     2016-12-31 00:00:00|
+----------------------------------------+
-- to_unix_timestamp
SELECT to_unix_timestamp('2016-04-08', 'yyyy-MM-dd');
+----------------------------------------+
|to_unix_timestamp(2016-04-08, yyyy-MM-dd)|
+----------------------------------------+
|                                1460041200|
+----------------------------------------+
-- to_utc_timestamp
SELECT to_utc_timestamp('2016-08-31', 'Asia/Seoul');
+----------------------------------------+
|to_utc_timestamp(2016-08-31, Asia/Seoul)|
+----------------------------------------+
|                     2016-08-30 15:00:00|
+----------------------------------------+
-- trunc
SELECT trunc('2019-08-04', 'week');
+-----------------------+
|trunc(2019-08-04, week)|
+-----------------------+
|             2019-07-29|
+-----------------------+
SELECT trunc('2019-08-04', 'quarter');
+--------------------------+
|trunc(2019-08-04, quarter)|
+--------------------------+
|                2019-07-01|
+--------------------------+
SELECT trunc('2009-02-12', 'MM');
+---------------------+
|trunc(2009-02-12, MM)|
+---------------------+
|           2009-02-01|
+---------------------+
SELECT trunc('2015-10-27', 'YEAR');
+-----------------------+
|trunc(2015-10-27, YEAR)|
+-----------------------+
|             2015-01-01|
```

```
+-----------------------+
-- try_to_timestamp
SELECT try_to_timestamp('2016-12-31 00:12:00');
+-----------------------------------+
|try_to_timestamp(2016-12-31 00:12:00)|
+-----------------------------------+
|                  2016-12-31 00:12:00|
+-----------------------------------+
SELECT try_to_timestamp('2016-12-31', 'yyyy-MM-dd');
+---------------------------------------+
|try_to_timestamp(2016-12-31, yyyy-MM-dd)|
+---------------------------------------+
|                     2016-12-31 00:00:00|
+---------------------------------------+
SELECT try_to_timestamp('foo', 'yyyy-MM-dd');
+---------------------------------+
|try_to_timestamp(foo, yyyy-MM-dd)|
+---------------------------------+
|                             NULL|
+---------------------------------+
-- unix_date
SELECT unix_date(DATE("1970-01-02"));
+---------------------+
|unix_date(1970-01-02)|
+---------------------+
|                    1|
+---------------------+
-- unix_micros
SELECT unix_micros(TIMESTAMP('1970-01-01 00:00:01Z'));
+--------------------------------+
|unix_micros(1970-01-01 00:00:01Z)|
+--------------------------------+
|                         1000000|
+--------------------------------+
-- unix_millis
SELECT unix_millis(TIMESTAMP('1970-01-01 00:00:01Z'));
+--------------------------------+
|unix_millis(1970-01-01 00:00:01Z)|
+--------------------------------+
|                            1000|
+--------------------------------+
-- unix_seconds
SELECT unix_seconds(TIMESTAMP('1970-01-01 00:00:01Z'));
+--------------------------------+
```

```
|unix_seconds(1970-01-01 00:00:01Z)|
+----------------------------------+
|                                 1|
+----------------------------------+
-- unix_timestamp
SELECT unix_timestamp();
+-----------------------------------------------------------+
|unix_timestamp(current_timestamp(), yyyy-MM-dd HH:mm:ss)|
+-----------------------------------------------------------+
|                                                 1708760216|
+-----------------------------------------------------------+
SELECT unix_timestamp('2016-04-08', 'yyyy-MM-dd');
+--------------------------------------+
|unix_timestamp(2016-04-08, yyyy-MM-dd)|
+--------------------------------------+
|                            1460041200|
+--------------------------------------+
-- weekday
SELECT weekday('2009-07-30');
+-------------------+
|weekday(2009-07-30)|
+-------------------+
|                  3|
+-------------------+
-- weekofyear
SELECT weekofyear('2008-02-20');
+----------------------+
|weekofyear(2008-02-20)|
+----------------------+
|                     8|
+----------------------+
-- window
SELECT a, window.start, window.end, count(*) as cnt FROM VALUES ('A1', '2021-01-01
 00:00:00'), ('A1', '2021-01-01 00:04:30'), ('A1', '2021-01-01 00:06:00'), ('A2',
 '2021-01-01 00:01:00') AS tab(a, b) GROUP by a, window(b, '5 minutes') ORDER BY a,
 start;
+---+-------------------+-------------------+---+
|  a|              start|                end|cnt|
+---+-------------------+-------------------+---+
| A1|2021-01-01 00:00:00|2021-01-01 00:05:00|  2|
| A1|2021-01-01 00:05:00|2021-01-01 00:10:00|  1|
| A2|2021-01-01 00:00:00|2021-01-01 00:05:00|  1|
+---+-------------------+-------------------+---+
```

```
SELECT a, window.start, window.end, count(*) as cnt FROM VALUES ('A1', '2021-01-01
 00:00:00'), ('A1', '2021-01-01 00:04:30'), ('A1', '2021-01-01 00:06:00'), ('A2',
 '2021-01-01 00:01:00') AS tab(a, b) GROUP by a, window(b, '10 minutes', '5 minutes')
 ORDER BY a, start;
+---+------------------+------------------+---+
|  a|             start|               end|cnt|
+---+------------------+------------------+---+
| A1|2020-12-31 23:55:00|2021-01-01 00:05:00|  2|
| A1|2021-01-01 00:00:00|2021-01-01 00:10:00|  3|
| A1|2021-01-01 00:05:00|2021-01-01 00:15:00|  1|
| A2|2020-12-31 23:55:00|2021-01-01 00:05:00|  1|
| A2|2021-01-01 00:00:00|2021-01-01 00:10:00|  1|
+---+------------------+------------------+---+
-- window_time
SELECT a, window.start as start, window.end as end, window_time(window), cnt FROM
 (SELECT a, window, count(*) as cnt FROM VALUES ('A1', '2021-01-01 00:00:00'), ('A1',
 '2021-01-01 00:04:30'), ('A1', '2021-01-01 00:06:00'), ('A2', '2021-01-01 00:01:00')
 AS tab(a, b) GROUP by a, window(b, '5 minutes') ORDER BY a, window.start);
+---+------------------+------------------+-------------------+---+
|  a|             start|               end| window_time(window)|cnt|
+---+------------------+------------------+-------------------+---+
| A1|2021-01-01 00:00:00|2021-01-01 00:05:00|2021-01-01 00:04:...|  2|
| A1|2021-01-01 00:05:00|2021-01-01 00:10:00|2021-01-01 00:09:...|  1|
| A2|2021-01-01 00:00:00|2021-01-01 00:05:00|2021-01-01 00:04:...|  1|
+---+------------------+------------------+-------------------+---+
-- year
SELECT year('2016-07-30');
+----------------+
|year(2016-07-30)|
+----------------+
|            2016|
+----------------+
```

Funzioni di aggregazione

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, consulta. the section called “Comandi SQL supportati”

Le funzioni di aggregazione operano sui valori delle righe per eseguire calcoli matematici come somma, media, conteggio, minimum/maximum valori, deviazione standard e stima, oltre ad alcune operazioni non matematiche.

Sintassi

```
aggregate_function(input1 [, input2, ...]) FILTER (WHERE boolean_expression)
```

Parametri

- `boolean_expression`- Specifica qualsiasi espressione che restituisce un risultato di tipo booleano. Due o più espressioni possono essere combinate insieme utilizzando gli operatori logici (AND, OR).

Funzioni aggregate ordinate

Queste funzioni di aggregazione utilizzano una sintassi diversa rispetto alle altre funzioni di aggregazione in modo da specificare un'espressione (in genere un nome di colonna) in base alla quale ordinare i valori.

Sintassi

```
{ PERCENTILE_CONT | PERCENTILE_DISC }(percentile) WITHIN GROUP (ORDER BY
 { order_by_expression [ ASC | DESC ] [ NULLS { FIRST | LAST } ] [ , ... ] }) FILTER
 (WHERE boolean_expression)
```

Parametri

- `percentile`- Il percentile del valore che vuoi trovare. Il percentile deve essere una costante compresa tra 0,0 e 1,0.
- `order_by_expression`- L'espressione (in genere il nome di una colonna) in base alla quale ordinare i valori prima di aggregarli.
- `boolean_expression`- Speciifica qualsiasi espressione che restituisce un risultato di tipo booleano. Due o più espressioni possono essere combinate insieme utilizzando gli operatori logici (AND, OR).

Esempi

```
CREATE OR REPLACE TEMPORARY VIEW basic_pays AS SELECT * FROM VALUES
```

```
('Jane Doe','Accounting',8435),
('Akua Mansa','Accounting',9998),
('John Doe','Accounting',8992),
('Juan Li','Accounting',8870),
('Carlos Salazar','Accounting',11472),
('Arnav Desai','Accounting',6627),
('Saanvi Sarkar','IT',8113),
('Shirley Rodriguez','IT',5186),
('Nikki Wolf','Sales',9181),
('Alejandro Rosalez','Sales',9441),
('Nikhil Jayashankar','Sales',6660),
('Richard Roe','Sales',10563),
('Pat Candella','SCM',10449),
('Gerard Hernandez','SCM',6949),
('Pamela Castillo','SCM',11303),
('Paulo Santos','SCM',11798),
('Jorge Souza','SCM',10586)
AS basic_pays(employee_name, department, salary);
SELECT * FROM basic_pays;
+------------------+----------+------+
|    employee_name  |department|salary|
+------------------+----------+------+
| Arnav Desai       |Accounting|  6627|
| Jorge Souza       |       SCM| 10586|
| Jane Doe          |Accounting|  8435|
| Nikhil Jayashankar|     Sales|  6660|
| Diego Vanauf      |     Sales| 10563|
| Carlos Salazar    |Accounting| 11472|
| Gerard Hernandez  |       SCM|  6949|
| John Doe          |Accounting|  8992|
| Nikki Wolf        |     Sales|  9181|
| Paulo Santos      |       SCM| 11798|
| Saanvi Sarkar     |        IT|  8113|
| Shirley Rodriguez |        IT|  5186|
| Pat Candella      |       SCM| 10449|
| Akua Mansa        |Accounting|  9998|
| Pamela Castillo   |       SCM| 11303|
| Alejandro Rosalez |     Sales|  9441|
| Juan Li           |Accounting|  8870|
+------------------+----------+------+
SELECT
department,
percentile_cont(0.25) WITHIN GROUP (ORDER BY salary) AS pc1,
```

```
percentile_cont(0.25) WITHIN GROUP (ORDER BY salary) FILTER (WHERE employee_name LIKE
 '%Bo%') AS pc2,
percentile_cont(0.25) WITHIN GROUP (ORDER BY salary DESC) AS pc3,
percentile_cont(0.25) WITHIN GROUP (ORDER BY salary DESC) FILTER (WHERE employee_name
 LIKE '%Bo%') AS pc4,
percentile_disc(0.25) WITHIN GROUP (ORDER BY salary) AS pd1,
percentile_disc(0.25) WITHIN GROUP (ORDER BY salary) FILTER (WHERE employee_name LIKE
 '%Bo%') AS pd2,
percentile_disc(0.25) WITHIN GROUP (ORDER BY salary DESC) AS pd3,
percentile_disc(0.25) WITHIN GROUP (ORDER BY salary DESC) FILTER (WHERE employee_name
 LIKE '%Bo%') AS pd4
FROM basic_pays
GROUP BY department
ORDER BY department;
+----------+-------+--------+-------+--------+-----+-----+-----+-----+
|department|    pc1|     pc2|    pc3|     pc4|  pd1|  pd2|  pd3|  pd4|
+----------+-------+--------+-------+--------+-----+-----+-----+-----+
|Accounting|8543.75| 7838.25| 9746.5|10260.75| 8435| 6627| 9998|11472|
|        IT|5917.75|    NULL|7381.25|    NULL| 5186| NULL| 8113| NULL|
|     Sales|8550.75|    NULL| 9721.5|    NULL| 6660| NULL|10563| NULL|
|       SCM|10449.0|10786.25|11303.0|11460.75|10449|10449|11303|11798|
+----------+-------+--------+-------+--------+-----+-----+-----+-----+
```

Funzioni condizionali

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedithe section called “Comandi SQL supportati”.

Funzione	Description
coalesce (espr 1, espr 2,...)	Restituisce il primo argomento non nullo, se esiste. Altrimenti, null.
se (espr 1, espr 2, espr 3)	Se `expr1` restituisce true, restituisce; altriment i restituisce. `expr2` `expr3`
ifnull (espr 1, espr 2)	Restituisce `expr2` se `expr1` è nullo o meno. `expr1`

Funzione	Description
nanvl (espr 1, espr 2)	Restituisce `expr1` se non è NaN o `expr2` altro.
nullif (espr 1, espr 2)	Restituisce null se è uguale a o meno. `expr1` `expr2` `expr1`
nvl (espr 1, espr 2)	Restituisce `expr2` se `expr1` è nullo o meno. `expr1`
nvl2 (espr 1, espr 2, espr 3)	Restituisce se non è nullo o altrimenti. `expr2` `expr1` `expr3`
CASO IN CUI expr1 POI expr2 [QUANDO expr3 THEN expr4] * [ELSE expr5] END	When `expr1` = true, restituisce`expr2`; altrimenti quando = true, restituisce; altrimenti ritorna. `expr3` `expr4` `expr5`

Esempi

```
-- coalesce
SELECT coalesce(NULL, 1, NULL);
+----------------------+
|coalesce(NULL, 1, NULL)|
+----------------------+
|                     1|
+----------------------+
-- if
SELECT if(1 < 2, 'a', 'b');
+------------------+
|(IF((1 < 2), a, b))|
+------------------+
|                 a|
+------------------+
-- ifnull
SELECT ifnull(NULL, array('2'));
+----------------------+
|ifnull(NULL, array(2))|
+----------------------+
|                  [2]|
```

```
+----------------------+
-- nanvl
SELECT nanvl(cast('NaN' as double), 123);
+------------------------------+
|nanvl(CAST(NaN AS DOUBLE), 123)|
+------------------------------+
|                         123.0|
+------------------------------+
-- nullif
SELECT nullif(2, 2);
+------------+
|nullif(2, 2)|
+------------+
|        NULL|
+------------+
-- nvl
SELECT nvl(NULL, array('2'));
+-------------------+
|nvl(NULL, array(2))|
+-------------------+
|                [2]|
+-------------------+
-- nvl2
SELECT nvl2(NULL, 2, 1);
+----------------+
|nvl2(NULL, 2, 1)|
+----------------+
|               1|
+----------------+
-- when
SELECT CASE WHEN 1 > 0 THEN 1 WHEN 2 > 0 THEN 2.0 ELSE 1.2 END;
+-------------------------------------------------------------+
|CASE WHEN (1 > 0) THEN 1 WHEN (2 > 0) THEN 2.0 ELSE 1.2 END|
+-------------------------------------------------------------+
|                                                          1.0|
+-------------------------------------------------------------+
SELECT CASE WHEN 1 < 0 THEN 1 WHEN 2 > 0 THEN 2.0 ELSE 1.2 END;
+-------------------------------------------------------------+
|CASE WHEN (1 < 0) THEN 1 WHEN (2 > 0) THEN 2.0 ELSE 1.2 END|
+-------------------------------------------------------------+
|                                                          2.0|
+-------------------------------------------------------------+
SELECT CASE WHEN 1 < 0 THEN 1 WHEN 2 < 0 THEN 2.0 END;
+-------------------------------------------------+
```

```
|CASE WHEN (1 < 0) THEN 1 WHEN (2 < 0) THEN 2.0 END|
+--------------------------------------------------+
|                                              NULL|
+--------------------------------------------------+
```

Funzioni JSON

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedithe section called “Comandi SQL supportati”.

Funzione	Description
from_json (JsonStr, schema [, opzioni])	Restituisce un valore di struttura con i dati `JsonStr` e `schema`.
get_json_ object (json_txt, percorso)	Estrae un oggetto json da `path`.
json_arra y_length (JsonArray)	Restituisce il numero di elementi nell'array JSON più esterno.
json_obje ct_keys (json_object)	Restituisce tutte le chiavi dell'oggetto JSON più esterno come array.
json_tuple (JsonStr, p1, p2,..., pn)	Restituisce una tupla come la funzione get_json_object, ma accetta più nomi. Tutti i parametri di input e i tipi di colonne di output sono stringhe.

Funzione	Description
schema_of _json (json [, opzioni])	Restituisce lo schema nel formato DDL della stringa JSON.
to_json (expr [, opzioni])	Restituisce una stringa JSON con un determinato valore di struttura

Esempi

```
-- from_json
SELECT from_json('{"a":1, "b":0.8}', 'a INT, b DOUBLE');
+--------------------------+
| from_json({"a":1, "b":0.8}) |
+--------------------------+
| {1, 0.8}                 |
+--------------------------+

SELECT from_json('{"time":"26/08/2015"}', 'time Timestamp', map('timestampFormat', 'dd/
MM/yyyy'));
+--------------------------------+
| from_json({"time":"26/08/2015"}) |
+--------------------------------+
| {2015-08-26 00:00...           |
+--------------------------------+

SELECT from_json('{"teacher": "Alice", "student": [{"name": "Bob", "rank": 1}, {"name":
 "Charlie", "rank": 2}]}', 'STRUCT<teacher: STRING, student: ARRAY<STRUCT<name: STRING,
 rank: INT>>>');
+-----------------------------------------------------------------------------------------------
+
| from_json({"teacher": "Alice", "student": [{"name": "Bob", "rank": 1}, {"name":
 "Charlie", "rank": 2}]}) |
+-----------------------------------------------------------------------------------------------
+
| {Alice, [{Bob, 1}...
                  |
+-----------------------------------------------------------------------------------------------
+
```

```
-- get_json_object
SELECT get_json_object('{"a":"b"}', '$.a');
+-------------------------------+
| get_json_object({"a":"b"}, $.a) |
+-------------------------------+
| b                             |
+-------------------------------+

-- json_array_length
SELECT json_array_length('[1,2,3,4]');
+----------------------------+
| json_array_length([1,2,3,4]) |
+----------------------------+
| 4                          |
+----------------------------+

SELECT json_array_length('[1,2,3,{"f1":1,"f2":[5,6]},4]');
+-------------------------------------------------+
| json_array_length([1,2,3,{"f1":1,"f2":[5,6]},4]) |
+-------------------------------------------------+
| 5                                               |
+-------------------------------------------------+

SELECT json_array_length('[1,2');
+----------------------+
| json_array_length([1,2) |
+----------------------+
| NULL                 |
+----------------------+

-- json_object_keys
SELECT json_object_keys('{}');
+--------------------+
| json_object_keys({}) |
+--------------------+
| []                 |
+--------------------+

SELECT json_object_keys('{"key": "value"}');
+---------------------------------+
| json_object_keys({"key": "value"}) |
+---------------------------------+
| [key]                           |
+---------------------------------+
```

```
SELECT json_object_keys('{"f1":"abc","f2":{"f3":"a", "f4":"b"}}');
+---------------------------------------------------------+
| json_object_keys({"f1":"abc","f2":{"f3":"a", "f4":"b"}}) |
+---------------------------------------------------------+
| [f1, f2]                                                |
+---------------------------------------------------------+

-- json_tuple
SELECT json_tuple('{"a":1, "b":2}', 'a', 'b');
+---+---+
| c0| c1|
+---+---+
|  1|  2|
+---+---+

-- schema_of_json
SELECT schema_of_json('[{"col":0}]');
+---------------------------+
| schema_of_json([{"col":0}]) |
+---------------------------+
| ARRAY<STRUCT<col:...      |
+---------------------------+

SELECT schema_of_json('[{"col":01}]', map('allowNumericLeadingZeros', 'true'));
+----------------------------+
| schema_of_json([{"col":01}]) |
+----------------------------+
| ARRAY<STRUCT<col:...       |
+----------------------------+

-- to_json
SELECT to_json(named_struct('a', 1, 'b', 2));
+---------------------------------+
| to_json(named_struct(a, 1, b, 2)) |
+---------------------------------+
| {"a":1,"b":2}                   |
+---------------------------------+

SELECT to_json(named_struct('time', to_timestamp('2015-08-26', 'yyyy-MM-dd')),
 map('timestampFormat', 'dd/MM/yyyy'));
+-------------------------------------------------------------------+
| to_json(named_struct(time, to_timestamp(2015-08-26, yyyy-MM-dd))) |
+-------------------------------------------------------------------+
```

```
| {"time":"26/08/20...                                                  |
+-----------------------------------------------------------------------+

SELECT to_json(array(named_struct('a', 1, 'b', 2)));
+----------------------------------------+
| to_json(array(named_struct(a, 1, b, 2))) |
+----------------------------------------+
| [{"a":1,"b":2}]                        |
+----------------------------------------+

SELECT to_json(map('a', named_struct('b', 1)));
+-----------------------------------+
| to_json(map(a, named_struct(b, 1))) |
+-----------------------------------+
| {"a":{"b":1}}                     |
+-----------------------------------+

SELECT to_json(map(named_struct('a', 1),named_struct('b', 2)));
+-------------------------------------------------------+
| to_json(map(named_struct(a, 1), named_struct(b, 2))) |
+-------------------------------------------------------+
| {"[1]":{"b":2}}                                       |
+-------------------------------------------------------+

SELECT to_json(map('a', 1));
+------------------+
| to_json(map(a, 1)) |
+------------------+
| {"a":1}          |
+------------------+

SELECT to_json(array(map('a', 1)));
+--------------------------+
| to_json(array(map(a, 1))) |
+--------------------------+
| [{"a":1}]                |
+--------------------------+
```

Funzioni di array

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Funzione	Description
array (expr,...)	Restituisce un array con gli elementi dati.
array_append (array, elemento)	Aggiunge l'elemento alla fine dell'array passato come primo argomento. Il tipo di elemento deve essere simile al tipo degli elementi dell'array. All'array viene inoltre aggiunto un elemento nullo. Ma se l'array è passato, è NULL, l'output è NULL
array_compact (matrice)	Rimuove i valori nulli dall'array.
array_contains (array, valore)	Restituisce true se l'array contiene il valore.
array_distinct (array)	Rimuove i valori duplicati dall'array.
array_except (matrice1, matrice2)	Restituisce una matrice degli elementi in array1 ma non in array2, senza duplicati.
array_insert (x, pos, val)	Inserisce val nell'indice pos dell'array x. Gli indici degli array iniziano da 1. L'indice negativo massimo è -1 per il quale la funzione inserisce un nuovo elemento dopo l'ultimo elemento corrente. L'indice superiore alla dimensione dell'array aggiunge l'array o lo precede se l'indice è negativo, con elementi 'null'.
array_intersect (matrice1, matrice2)	Restituisce una matrice degli elementi nell'intersezione di array1 e array2, senza duplicati.

Funzione	Description
array_join (array, delimiter [, nullReplacement])	Concatena gli elementi dell'array specifica to utilizzando il delimitatore e una stringa opzionale per sostituire i valori null. Se non è impostato alcun valore per NullReplacement, viene filtrato qualsiasi valore nullo.
array_max (matrice)	Restituisce il valore massimo nell'array. NaN è maggiore di qualsiasi elemento non NaN per double/float tipo. Gli elementi NULL vengono ignorati.
array_min (matrice)	Restituisce il valore minimo nell'array. NaN è maggiore di qualsiasi elemento non NaN per double/float tipo. Gli elementi NULL vengono ignorati.
array_position (array, elemento)	Restituisce l'indice (a base 1) del primo elemento corrispondente dell'array se è lungo, oppure 0 se non viene trovata alcuna corrispon denza.
array_prepend (array, elemento)	Aggiunge l'elemento all'inizio dell'array passato come primo argomento. Il tipo di elemento deve essere lo stesso del tipo degli elementi dell'array. All'array viene inoltre aggiunto un elemento nullo. Ma se l'array passato è NULL, l'output è NULL
array_remove (array, elemento)	Rimuove tutti gli elementi uguali all'elemento dall'array.
array_repeat (elemento, conteggio)	Restituisce l'array contenente i tempi di conteggio degli elementi.
array_union (matrice1, matrice2)	Restituisce un array degli elementi nell'unione di array1 e array2, senza duplicati.

Funzione	Description
arrays_overlap (a1, a2)	Restituisce true se a1 contiene almeno un elemento non nullo presente anche in a2. Se gli array non hanno alcun elemento comune e sono entrambi non vuoti e uno di essi contiene un elemento nullo, viene restituito null, false in caso contrario.
arrays_zip (a1, a2,...)	Restituisce un array unito di strutture in cui la struttura N-esima contiene tutti gli N-esimi valori degli array di input.
arrayOfArraysappiattire ()	Trasforma una matrice di matrici in una singola matrice.
get (array, index)	Restituisce un elemento dell'array in un determinato indice (basato su 0). Se l'indice punta al di fuori dei limiti dell'array, questa funzione restituisce NULL.
sequenza (inizio, arresto, passo)	Genera una serie di elementi dall'inizio alla fine (incluso), incrementando passo dopo passo. Il tipo degli elementi restituiti è lo stesso del tipo di espressioni di argomento. I tipi supportati sono: byte, short, integer, long, date, timestamp. Le espressioni di inizio e fine devono avere lo stesso tipo. Se le espressioni di inizio e fine vengono risolte nel tipo «data» o «timestamp», l'espressione step deve risolvere nel tipo «intervallo» o «intervallo anno-mese » o «intervallo giorno-mese», altrimenti nello stesso tipo delle espressioni di inizio e fine.
shuffle (matrice)	Restituisce una permutazione casuale dell'arra y dato.

Funzione	Description
slice (x, inizio, lunghezza)	Sottoinsiemi dell'array x a partire dall'inizio dell'indice (gli indici dell'array iniziano da 1 o iniziano dalla fine se l'inizio è negativo) con la lunghezza specificata.
sort_array (array [, AscendingOrder])	Ordina l'array di input in ordine crescente o decrescente in base all'ordine naturale degli elementi dell'array. NaN è maggiore di qualsiasi elemento non NaN per double/float tipo. Gli elementi nulli verranno posizionati all'inizio dell'array restituito in ordine crescente o alla fine dell'array restituito in ordine decrescente.

Esempi

```
-- array
SELECT array(1, 2, 3);
+--------------+
|array(1, 2, 3)|
+--------------+
|     [1, 2, 3]|
+--------------+
-- array_append
SELECT array_append(array('b', 'd', 'c', 'a'), 'd');
+---------------------------------+
|array_append(array(b, d, c, a), d)|
+---------------------------------+
|                   [b, d, c, a, d]|
+---------------------------------+
SELECT array_append(array(1, 2, 3, null), null);
+---------------------------------------+
|array_append(array(1, 2, 3, NULL), NULL)|
+---------------------------------------+
|                    [1, 2, 3, NULL, N...|
+---------------------------------------+
SELECT array_append(CAST(null as Array<Int>), 2);
+--------------------+
```

```
|array_append(NULL, 2)|
+---------------------+
|                 NULL|
+---------------------+
-- array_compact
SELECT array_compact(array(1, 2, 3, null));
+-----------------------------------+
|array_compact(array(1, 2, 3, NULL))|
+-----------------------------------+
|                          [1, 2, 3]|
+-----------------------------------+
SELECT array_compact(array("a", "b", "c"));
+----------------------------+
|array_compact(array(a, b, c))|
+----------------------------+
|                    [a, b, c]|
+----------------------------+
-- array_contains
SELECT array_contains(array(1, 2, 3), 2);
+---------------------------------+
|array_contains(array(1, 2, 3), 2)|
+---------------------------------+
|                             true|
+---------------------------------+
-- array_distinct
SELECT array_distinct(array(1, 2, 3, null, 3));
+--------------------------------------+
|array_distinct(array(1, 2, 3, NULL, 3))|
+--------------------------------------+
|                       [1, 2, 3, NULL]|
+--------------------------------------+
-- array_except
SELECT array_except(array(1, 2, 3), array(1, 3, 5));
+-------------------------------------------+
|array_except(array(1, 2, 3), array(1, 3, 5))|
+-------------------------------------------+
|                                        [2]|
+-------------------------------------------+
-- array_insert
SELECT array_insert(array(1, 2, 3, 4), 5, 5);
+-----------------------------------+
|array_insert(array(1, 2, 3, 4), 5, 5)|
+-----------------------------------+
|                     [1, 2, 3, 4, 5]|
```

```
+-----------------------------------+
SELECT array_insert(array(5, 4, 3, 2), -1, 1);
+-----------------------------------+
|array_insert(array(5, 4, 3, 2), -1, 1)|
+-----------------------------------+
|                     [5, 4, 3, 2, 1]|
+-----------------------------------+
SELECT array_insert(array(5, 3, 2, 1), -4, 4);
+-----------------------------------+
|array_insert(array(5, 3, 2, 1), -4, 4)|
+-----------------------------------+
|                     [5, 4, 3, 2, 1]|
+-----------------------------------+
-- array_intersect
SELECT array_intersect(array(1, 2, 3), array(1, 3, 5));
+-----------------------------------------------+
|array_intersect(array(1, 2, 3), array(1, 3, 5))|
+-----------------------------------------------+
|                                         [1, 3]|
+-----------------------------------------------+
-- array_join
SELECT array_join(array('hello', 'world'), ' ');
+---------------------------------+
|array_join(array(hello, world),  )|
+---------------------------------+
|                     hello world|
+---------------------------------+
SELECT array_join(array('hello', null ,'world'), ' ');
+---------------------------------------+
|array_join(array(hello, NULL, world),  )|
+---------------------------------------+
|                           hello world|
+---------------------------------------+
SELECT array_join(array('hello', null ,'world'), ' ', ',');
+-----------------------------------------+
|array_join(array(hello, NULL, world),  , ,)|
+-----------------------------------------+
|                            hello , world|
+-----------------------------------------+
-- array_max
SELECT array_max(array(1, 20, null, 3));
+--------------------------------+
|array_max(array(1, 20, NULL, 3))|
+--------------------------------+
```

```
|                              20|
+--------------------------------+
-- array_min
SELECT array_min(array(1, 20, null, 3));
+--------------------------------+
|array_min(array(1, 20, NULL, 3))|
+--------------------------------+
|                               1|
+--------------------------------+
-- array_position
SELECT array_position(array(312, 773, 708, 708), 708);
+-----------------------------------------------+
|array_position(array(312, 773, 708, 708), 708)|
+-----------------------------------------------+
|                                              3|
+-----------------------------------------------+
SELECT array_position(array(312, 773, 708, 708), 414);
+-----------------------------------------------+
|array_position(array(312, 773, 708, 708), 414)|
+-----------------------------------------------+
|                                              0|
+-----------------------------------------------+
-- array_prepend
SELECT array_prepend(array('b', 'd', 'c', 'a'), 'd');
+----------------------------------+
|array_prepend(array(b, d, c, a), d)|
+----------------------------------+
|                    [d, b, d, c, a]|
+----------------------------------+
SELECT array_prepend(array(1, 2, 3, null), null);
+---------------------------------------+
|array_prepend(array(1, 2, 3, NULL), NULL)|
+---------------------------------------+
|                     [NULL, 1, 2, 3, N...|
+---------------------------------------+
SELECT array_prepend(CAST(null as Array<Int>), 2);
+----------------------+
|array_prepend(NULL, 2)|
+----------------------+
|                  NULL|
+----------------------+
-- array_remove
SELECT array_remove(array(1, 2, 3, null, 3), 3);
+---------------------------------------+
```

```
|array_remove(array(1, 2, 3, NULL, 3), 3)|
+----------------------------------------+
|                            [1, 2, NULL]|
+----------------------------------------+
-- array_repeat
SELECT array_repeat('123', 2);
+--------------------+
|array_repeat(123, 2)|
+--------------------+
|          [123, 123]|
+--------------------+
-- array_union
SELECT array_union(array(1, 2, 3), array(1, 3, 5));
+-------------------------------------------+
|array_union(array(1, 2, 3), array(1, 3, 5))|
+-------------------------------------------+
|                               [1, 2, 3, 5]|
+-------------------------------------------+
-- arrays_overlap
SELECT arrays_overlap(array(1, 2, 3), array(3, 4, 5));
+-----------------------------------------------+
|arrays_overlap(array(1, 2, 3), array(3, 4, 5))|
+-----------------------------------------------+
|                                           true|
+-----------------------------------------------+
-- arrays_zip
SELECT arrays_zip(array(1, 2, 3), array(2, 3, 4));
+-----------------------------------------+
|arrays_zip(array(1, 2, 3), array(2, 3, 4))|
+-----------------------------------------+
|                     [{1, 2}, {2, 3}, ...|
+-----------------------------------------+
SELECT arrays_zip(array(1, 2), array(2, 3), array(3, 4));
+-------------------------------------------------+
|arrays_zip(array(1, 2), array(2, 3), array(3, 4))|
+-------------------------------------------------+
|                             [{1, 2, 3}, {2, 3...|
+-------------------------------------------------+
-- flatten
SELECT flatten(array(array(1, 2), array(3, 4)));
+----------------------------------------+
|flatten(array(array(1, 2), array(3, 4)))|
+----------------------------------------+
|                            [1, 2, 3, 4]|
```

```
+--------------------------------------+
-- get
SELECT get(array(1, 2, 3), 0);
+---------------------+
|get(array(1, 2, 3), 0)|
+---------------------+
|                    1|
+---------------------+
SELECT get(array(1, 2, 3), 3);
+---------------------+
|get(array(1, 2, 3), 3)|
+---------------------+
|                 NULL|
+---------------------+
SELECT get(array(1, 2, 3), -1);
+----------------------+
|get(array(1, 2, 3), -1)|
+----------------------+
|                  NULL|
+----------------------+
-- sequence
SELECT sequence(1, 5);
+---------------+
| sequence(1, 5)|
+---------------+
|[1, 2, 3, 4, 5]|
+---------------+
SELECT sequence(5, 1);
+---------------+
| sequence(5, 1)|
+---------------+
|[5, 4, 3, 2, 1]|
+---------------+
SELECT sequence(to_date('2018-01-01'), to_date('2018-03-01'), interval 1 month);
+------------------------------------------------------------------------+
|sequence(to_date(2018-01-01), to_date(2018-03-01), INTERVAL '1' MONTH)|
+------------------------------------------------------------------------+
|                                                    [2018-01-01, 2018...|
+------------------------------------------------------------------------+
SELECT sequence(to_date('2018-01-01'), to_date('2018-03-01'), interval '0-1' year to
 month);
+-----------------------------------------------------------------------------------+
|sequence(to_date(2018-01-01), to_date(2018-03-01), INTERVAL '0-1' YEAR TO MONTH)|
+-----------------------------------------------------------------------------------+
```

```
|                                                                                [2018-01-01, 2018...|
+--------------------------------------------------------------------------------+
-- shuffle
SELECT shuffle(array(1, 20, 3, 5));
+----------------------------+
|shuffle(array(1, 20, 3, 5))|
+----------------------------+
|              [5, 1, 20, 3]|
+----------------------------+
SELECT shuffle(array(1, 20, null, 3));
+-----------------------------+
|shuffle(array(1, 20, NULL, 3))|
+-----------------------------+
|              [1, NULL, 20, 3]|
+-----------------------------+
-- slice
SELECT slice(array(1, 2, 3, 4), 2, 2);
+-----------------------------+
|slice(array(1, 2, 3, 4), 2, 2)|
+-----------------------------+
|                        [2, 3]|
+-----------------------------+
SELECT slice(array(1, 2, 3, 4), -2, 2);
+------------------------------+
|slice(array(1, 2, 3, 4), -2, 2)|
+------------------------------+
|                         [3, 4]|
+------------------------------+
-- sort_array
SELECT sort_array(array('b', 'd', null, 'c', 'a'), true);
+-----------------------------------------+
|sort_array(array(b, d, NULL, c, a), true)|
+-----------------------------------------+
|                       [NULL, a, b, c, d]|
+-----------------------------------------+
```

Funzioni finestra

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Le funzioni della finestra operano su un gruppo di righe, denominato finestra, e calcolano un valore restituito per ogni riga in base al gruppo di righe. Le funzioni della finestra sono utili per elaborare attività come il calcolo di una media mobile, il calcolo di una statistica cumulativa o l'accesso al valore delle righe in base alla posizione relativa della riga corrente.

Sintassi

```
window_function [ nulls_option ] OVER ( [ { PARTITION | DISTRIBUTE } BY
 partition_col_name = partition_col_val ( [ , ... ] ) ] { ORDER | SORT } BY expression
 [ ASC | DESC ] [ NULLS { FIRST | LAST } ] [ , ... ] [ window_frame ] )
```

Parametri

- Funzioni di classificazione

  Sintassi: `RANK | DENSE_RANK | PERCENT_RANK | NTILE | ROW_NUMBER`

  Funzioni analitiche

  Sintassi: `CUME_DIST | LAG | LEAD | NTH_VALUE | FIRST_VALUE | LAST_VALUE`

  Funzioni di aggregazione

  Sintassi: `MAX | MIN | COUNT | SUM | AVG | ...`
- `nulls_option`- Speciifica se ignorare o meno i valori nulli durante la valutazione della funzione finestra. RESPECT NULLS significa non saltare i valori nulli, mentre IGNORE NULLS significa saltare. Se non viene specificato, l'impostazione predefinita è RESPECT NULLS.

  Sintassi: `{ IGNORE | RESPECT } NULLS`

  Nota: `Only LAG | LEAD | NTH_VALUE | FIRST_VALUE | LAST_VALUE` può essere usato con`IGNORE NULLS`.
- `window_frame`- Specificate su quale riga iniziare la finestra e dove terminarla.

  Sintassi: `{ RANGE | ROWS } { frame_start | BETWEEN frame_start AND frame_end }`

  frame_start e frame_end hanno la seguente sintassi:

  Sintassi: `UNBOUNDED PRECEDING | offset PRECEDING | CURRENT ROW | offset FOLLOWING | UNBOUNDED FOLLOWING`

offset: specifica l'offset dalla posizione della riga corrente.

Nota Se frame_end viene omesso, il valore predefinito è CURRENT ROW.

Esempi

```
CREATE TABLE employees (name STRING, dept STRING, salary INT, age INT);
INSERT INTO employees VALUES ("Lisa", "Sales", 10000, 35);
INSERT INTO employees VALUES ("Evan", "Sales", 32000, 38);
INSERT INTO employees VALUES ("Fred", "Engineering", 21000, 28);
INSERT INTO employees VALUES ("Alex", "Sales", 30000, 33);
INSERT INTO employees VALUES ("Tom", "Engineering", 23000, 33);
INSERT INTO employees VALUES ("Jane", "Marketing", 29000, 28);
INSERT INTO employees VALUES ("Jeff", "Marketing", 35000, 38);
INSERT INTO employees VALUES ("Paul", "Engineering", 29000, 23);
INSERT INTO employees VALUES ("Chloe", "Engineering", 23000, 25);
SELECT * FROM employees;
+-----+-----------+------+-----+
| name|       dept|salary|  age|
+-----+-----------+------+-----+
|Chloe|Engineering| 23000|   25|
| Fred|Engineering| 21000|   28|
| Paul|Engineering| 29000|   23|
|Helen|  Marketing| 29000|   40|
|  Tom|Engineering| 23000|   33|
| Jane|  Marketing| 29000|   28|
| Jeff|  Marketing| 35000|   38|
| Evan|      Sales| 32000|   38|
| Lisa|      Sales| 10000|   35|
| Alex|      Sales| 30000|   33|
+-----+-----------+------+-----+
SELECT name, dept, salary, RANK() OVER (PARTITION BY dept ORDER BY salary) AS rank FROM
 employees;
+-----+-----------+------+----+
| name|       dept|salary|rank|
+-----+-----------+------+----+
| Lisa|      Sales| 10000|   1|
| Alex|      Sales| 30000|   2|
| Evan|      Sales| 32000|   3|
| Fred|Engineering| 21000|   1|
|  Tom|Engineering| 23000|   2|
|Chloe|Engineering| 23000|   2|
```

```
| Paul|Engineering| 29000|   4|
|Helen|  Marketing| 29000|   1|
| Jane|  Marketing| 29000|   1|
| Jeff|  Marketing| 35000|   3|
+-----+-----------+------+----+
SELECT name, dept, salary, DENSE_RANK() OVER (PARTITION BY dept ORDER BY salary ROWS
 BETWEEN
UNBOUNDED PRECEDING AND CURRENT ROW) AS dense_rank FROM employees;
+-----+-----------+------+----------+
| name|       dept|salary|dense_rank|
+-----+-----------+------+----------+
| Lisa|      Sales| 10000|         1|
| Alex|      Sales| 30000|         2|
| Evan|      Sales| 32000|         3|
| Fred|Engineering| 21000|         1|
|  Tom|Engineering| 23000|         2|
|Chloe|Engineering| 23000|         2|
| Paul|Engineering| 29000|         3|
|Helen|  Marketing| 29000|         1|
| Jane|  Marketing| 29000|         1|
| Jeff|  Marketing| 35000|         2|
+-----+-----------+------+----------+
SELECT name, dept, age, CUME_DIST() OVER (PARTITION BY dept ORDER BY age
RANGE BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW) AS cume_dist FROM employees;
+-----+-----------+------+------------------+
| name|       dept|age   |         cume_dist|
+-----+-----------+------+------------------+
| Alex|      Sales|    33|0.3333333333333333|
| Lisa|      Sales|    35|0.6666666666666666|
| Evan|      Sales|    38|               1.0|
| Paul|Engineering|    23|              0.25|
|Chloe|Engineering|    25|              0.75|
| Fred|Engineering|    28|              0.25|
|  Tom|Engineering|    33|               1.0|
| Jane|  Marketing|    28|0.3333333333333333|
| Jeff|  Marketing|    38|0.6666666666666666|
|Helen|  Marketing|    40|               1.0|
+-----+-----------+------+------------------+
SELECT name, dept, salary, MIN(salary) OVER (PARTITION BY dept ORDER BY salary) AS min
FROM employees;
+-----+-----------+------+-----+
| name|       dept|salary|  min|
+-----+-----------+------+-----+
| Lisa|      Sales| 10000|10000|
```

```
| Alex|      Sales| 30000|10000|
| Evan|      Sales| 32000|10000|
|Helen|  Marketing| 29000|29000|
| Jane|  Marketing| 29000|29000|
| Jeff|  Marketing| 35000|29000|
| Fred|Engineering| 21000|21000|
|  Tom|Engineering| 23000|21000|
|Chloe|Engineering| 23000|21000|
| Paul|Engineering| 29000|21000|
+-----+-----------+------+-----+
SELECT name, salary,
LAG(salary) OVER (PARTITION BY dept ORDER BY salary) AS lag,
LEAD(salary, 1, 0) OVER (PARTITION BY dept ORDER BY salary) AS lead
FROM employees;
+-----+-----------+------+-----+-----+
| name|       dept|salary|  lag| lead|
+-----+-----------+------+-----+-----+
| Lisa|      Sales| 10000|NULL |30000|
| Alex|      Sales| 30000|10000|32000|
| Evan|      Sales| 32000|30000|    0|
| Fred|Engineering| 21000| NULL|23000|
|Chloe|Engineering| 23000|21000|23000|
|  Tom|Engineering| 23000|23000|29000|
| Paul|Engineering| 29000|23000|    0|
|Helen|  Marketing| 29000| NULL|29000|
| Jane|  Marketing| 29000|29000|35000|
| Jeff|  Marketing| 35000|29000|    0|
+-----+-----------+------+-----+-----+
SELECT id, v,
LEAD(v, 0) IGNORE NULLS OVER w lead,
LAG(v, 0) IGNORE NULLS OVER w lag,
NTH_VALUE(v, 2) IGNORE NULLS OVER w nth_value,
FIRST_VALUE(v) IGNORE NULLS OVER w first_value,
LAST_VALUE(v) IGNORE NULLS OVER w last_value
FROM test_ignore_null
WINDOW w AS (ORDER BY id)
ORDER BY id;
+--+----+----+----+---------+-----------+----------+
|id|   v|lead| lag|nth_value|first_value|last_value|
+--+----+----+----+---------+-----------+----------+
| 0|NULL|NULL|NULL|     NULL|       NULL|      NULL|
| 1|   x|   x|   x|     NULL|          x|         x|
| 2|NULL|NULL|NULL|     NULL|          x|         x|
| 3|NULL|NULL|NULL|     NULL|          x|         x|
```

```
| 4|   y|   y|   y|        y|          x|         y|
| 5|NULL|NULL|NULL|        y|          x|         y|
| 6|   z|   z|   z|        y|          x|         z|
| 7|   v|   v|   v|        y|          x|         v|
| 8|NULL|NULL|NULL|        y|          x|         v|
+--+----+----+----+---------+-----------+----------+
```

Funzioni di conversione

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Funzione	Description
bigint (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `bigint`.
binario (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `binary`.
booleano (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `boolean`.
cast (tipo expr AS)	Trasmette il valore `expr` al tipo di dati di destinazione `type`.
data (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `date`.
decimale (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `decimal`.
doppio (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `double`.
float (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `float`.
int (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `int`.
smallint (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `smallint`.
stringa (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `string`.
timestamp (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `timestamp`.

Funzione	Description
tinyint (expr)	Trasmette il valore `expr` al tipo di dati di destinazione `tinyint`.

Esempi

```
-- cast
SELECT cast(field as int);
+---------------+
|CAST(field AS INT)|
+---------------+
|             10|
+---------------+
```

Funzioni di predicato

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Funzione	Description
! espr	Logico no.
espr 1 < espr 2	Restituisce vero se `expr1` è minore di `expr2`.
espr 1 <= espr 2	Restituisce vero se `expr1` è minore o uguale a `expr2`.
expr1 <=> espr 2	Restituisce lo stesso risultato dell'operatore EQUAL (=) per gli operandi non nulli, ma restituisce true se entrambi sono nulli, false se uno di essi è nullo.
espr 1 = espr 2	Restituisce true se `expr1` è uguale a `expr2`, altrimenti restituisce falso.

Funzione	Description
espr 1 == espr 2	Restituisce true se `expr1` è uguale a `expr2`, altrimenti restituisce false.
expr1 > expr2	Restituisce vero se `expr1` è maggiore di `expr2`.
espr 1 >= espr 2	Restituisce vero se `expr1` è maggiore o uguale a `expr2`.
expr1 ed expr2	AND logico.
pattern simile a una stella [ESCAPE escape]	Restituisce true se str corrisponde a `pattern` con `escape` senza distinzione tra maiuscole e minuscole, null se alcuni argomenti sono nulli, false in caso contrario.
espr1 in (expr2, expr3,...)	Restituisce true se `expr` è uguale a qualsiasi ValN.
isnan (expr)	Restituisce true se `expr` è NaN, altrimenti restituisce false.
isnotnull (expr)	Restituisce true se `expr` non è nullo, altrimenti restituisce false.
isnull (espr)	Restituisce true se `expr` è nullo, altrimenti restituisce falso.
str like pattern [ESCAPE escape]	Restituisce true se str corrisponde a `pattern` con `escape`, null se qualche argomento è nullo, falso altrimenti.
non expr	Logico no.
expr1 o expr2	OR logico.
regexp (str, regexp)	Restituisce true se `str` corrisponde a `regexp`, altrimenti restituisce false.
regexp_like (str, regexp)	Restituisce true se `str` corrisponde a `regexp`, altrimenti restituisce false.
rlike (str, regexp)	Restituisce true se `str` corrisponde a `regexp`, altrimenti restituisce false.

Esempi

```
-- !
SELECT ! true;
+----------+
|(NOT true)|
+----------+
|     false|
+----------+
SELECT ! false;
+-----------+
|(NOT false)|
+-----------+
|       true|
+-----------+
SELECT ! NULL;
+----------+
|(NOT NULL)|
+----------+
|      NULL|
+----------+
-- <
SELECT to_date('2009-07-30 04:17:52') < to_date('2009-07-30 04:17:52');
+-------------------------------------------------------------+
|(to_date(2009-07-30 04:17:52) < to_date(2009-07-30 04:17:52))|
+-------------------------------------------------------------+
|                                                        false|
+-------------------------------------------------------------+
SELECT to_date('2009-07-30 04:17:52') < to_date('2009-08-01 04:17:52');
+-------------------------------------------------------------+
|(to_date(2009-07-30 04:17:52) < to_date(2009-08-01 04:17:52))|
+-------------------------------------------------------------+
|                                                         true|
+-------------------------------------------------------------+
SELECT 1 < NULL;
+----------+
|(1 < NULL)|
+----------+
|      NULL|
+----------+
-- <=
SELECT 2 <= 2;
+--------+
|(2 <= 2)|
+--------+
```

```
|    true|
+--------+
SELECT 1.0 <= '1';
+----------+
|(1.0 <= 1)|
+----------+
|      true|
+----------+
SELECT to_date('2009-07-30 04:17:52') <= to_date('2009-07-30 04:17:52');
+------------------------------------------------------------+
|(to_date(2009-07-30 04:17:52) <= to_date(2009-07-30 04:17:52))|
+------------------------------------------------------------+
|                                                        true|
+------------------------------------------------------------+
SELECT to_date('2009-07-30 04:17:52') <= to_date('2009-08-01 04:17:52');
+------------------------------------------------------------+
|(to_date(2009-07-30 04:17:52) <= to_date(2009-08-01 04:17:52))|
+------------------------------------------------------------+
|                                                        true|
+------------------------------------------------------------+
SELECT 1 <= NULL;
+-----------+
|(1 <= NULL)|
+-----------+
|       NULL|
+-----------+
-- <=>
SELECT 2 <=> 2;
+---------+
|(2 <=> 2)|
+---------+
|     true|
+---------+
SELECT 1 <=> '1';
+---------+
|(1 <=> 1)|
+---------+
|     true|
+---------+
SELECT true <=> NULL;
+---------------+
|(true <=> NULL)|
+---------------+
|          false|
```

```
+---------------+
SELECT NULL <=> NULL;
+---------------+
|(NULL <=> NULL)|
+---------------+
|           true|
+---------------+
-- =
SELECT 2 = 2;
+-------+
|(2 = 2)|
+-------+
|   true|
+-------+
SELECT 1 = '1';
+-------+
|(1 = 1)|
+-------+
|   true|
+-------+
SELECT true = NULL;
+-------------+
|(true = NULL)|
+-------------+
|         NULL|
+-------------+
SELECT NULL = NULL;
+-------------+
|(NULL = NULL)|
+-------------+
|         NULL|
+-------------+
-- ==
SELECT 2 == 2;
+-------+
|(2 = 2)|
+-------+
|   true|
+-------+
SELECT 1 == '1';
+-------+
|(1 = 1)|
+-------+
|   true|
```

```
+-------+
SELECT true == NULL;
+-------------+
|(true = NULL)|
+-------------+
|         NULL|
+-------------+
SELECT NULL == NULL;
+-------------+
|(NULL = NULL)|
+-------------+
|         NULL|
+-------------+
-- >
SELECT 2 > 1;
+-------+
|(2 > 1)|
+-------+
|   true|
+-------+
SELECT 2 > 1.1;
+-------+
|(2 > 1)|
+-------+
|   true|
+-------+
SELECT to_date('2009-07-30 04:17:52') > to_date('2009-07-30 04:17:52');
+------------------------------------------------------------+
|(to_date(2009-07-30 04:17:52) > to_date(2009-07-30 04:17:52))|
+------------------------------------------------------------+
|                                                       false|
+------------------------------------------------------------+
SELECT to_date('2009-07-30 04:17:52') > to_date('2009-08-01 04:17:52');
+------------------------------------------------------------+
|(to_date(2009-07-30 04:17:52) > to_date(2009-08-01 04:17:52))|
+------------------------------------------------------------+
|                                                       false|
+------------------------------------------------------------+
SELECT 1 > NULL;
+----------+
|(1 > NULL)|
+----------+
|      NULL|
+----------+
```

```
-- >=
SELECT 2 >= 1;
+--------+
|(2 >= 1)|
+--------+
|    true|
+--------+
SELECT 2.0 >= '2.1';
+------------+
|(2.0 >= 2.1)|
+------------+
|       false|
+------------+
SELECT to_date('2009-07-30 04:17:52') >= to_date('2009-07-30 04:17:52');
+-------------------------------------------------------------+
|(to_date(2009-07-30 04:17:52) >= to_date(2009-07-30 04:17:52))|
+-------------------------------------------------------------+
|                                                         true|
+-------------------------------------------------------------+
SELECT to_date('2009-07-30 04:17:52') >= to_date('2009-08-01 04:17:52');
+-------------------------------------------------------------+
|(to_date(2009-07-30 04:17:52) >= to_date(2009-08-01 04:17:52))|
+-------------------------------------------------------------+
|                                                        false|
+-------------------------------------------------------------+
SELECT 1 >= NULL;
+-----------+
|(1 >= NULL)|
+-----------+
|       NULL|
+-----------+
-- and
SELECT true and true;
+---------------+
|(true AND true)|
+---------------+
|           true|
+---------------+
SELECT true and false;
+----------------+
|(true AND false)|
+----------------+
|           false|
+----------------+
```

```
SELECT true and NULL;
+---------------+
|(true AND NULL)|
+---------------+
|           NULL|
+---------------+
SELECT false and NULL;
+----------------+
|(false AND NULL)|
+----------------+
|           false|
+----------------+
-- ilike
SELECT ilike('Wagon', '_Agon');
+-------------------+
|ilike(Wagon, _Agon)|
+-------------------+
|               true|
+-------------------+
SELECT '%SystemDrive%\Users\John' ilike '\%SystemDrive\%\\users%';
+-----------------------------------------------------------+
|ilike(%SystemDrive%\Users\John, \%SystemDrive\%\\users%)|
+-----------------------------------------------------------+
|                                                       true|
+-----------------------------------------------------------+
SELECT '%SystemDrive%\\USERS\\John' ilike '\%SystemDrive\%\\\\Users%';
+-----------------------------------------------------------+
|ilike(%SystemDrive%\USERS\John, \%SystemDrive\%\\Users%)|
+-----------------------------------------------------------+
|                                                       true|
+-----------------------------------------------------------+
SELECT '%SystemDrive%/Users/John' ilike '/%SYSTEMDrive/%//Users%' ESCAPE '/';
+-----------------------------------------------------------+
|ilike(%SystemDrive%/Users/John, /%SYSTEMDrive/%//Users%)|
+-----------------------------------------------------------+
|                                                       true|
+-----------------------------------------------------------+
-- in
SELECT 1 in(1, 2, 3);
+----------------+
|(1 IN (1, 2, 3))|
+----------------+
|            true|
+----------------+
```

```
SELECT 1 in(2, 3, 4);
+----------------+
|(1 IN (2, 3, 4))|
+----------------+
|           false|
+----------------+
SELECT named_struct('a', 1, 'b', 2) in(named_struct('a', 1, 'b', 1), named_struct('a',
 1, 'b', 3));
+-----------------------------------------------------------------------------------+
|(named_struct(a, 1, b, 2) IN (named_struct(a, 1, b, 1), named_struct(a, 1, b, 3)))|
+-----------------------------------------------------------------------------------+
|                                                                              false|
+-----------------------------------------------------------------------------------+
SELECT named_struct('a', 1, 'b', 2) in(named_struct('a', 1, 'b', 2), named_struct('a',
 1, 'b', 3));
+-----------------------------------------------------------------------------------+
|(named_struct(a, 1, b, 2) IN (named_struct(a, 1, b, 2), named_struct(a, 1, b, 3)))|
+-----------------------------------------------------------------------------------+
|                                                                               true|
+-----------------------------------------------------------------------------------+
-- isnan
SELECT isnan(cast('NaN' as double));
+--------------------------+
|isnan(CAST(NaN AS DOUBLE))|
+--------------------------+
|                      true|
+--------------------------+
-- isnotnull
SELECT isnotnull(1);
+---------------+
|(1 IS NOT NULL)|
+---------------+
|           true|
+---------------+
-- isnull
SELECT isnull(1);
+-----------+
|(1 IS NULL)|
+-----------+
|      false|
+-----------+
-- like
SELECT like('Wagon', '_Agon');
+----------------+
```

```
|Wagon LIKE _Agon|
+----------------+
|            true|
+----------------+
-- not
SELECT not true;
+----------+
|(NOT true)|
+----------+
|     false|
+----------+
SELECT not false;
+-----------+
|(NOT false)|
+-----------+
|       true|
+-----------+
SELECT not NULL;
+----------+
|(NOT NULL)|
+----------+
|      NULL|
+----------+
-- or
SELECT true or false;
+---------------+
|(true OR false)|
+---------------+
|           true|
+---------------+
SELECT false or false;
+----------------+
|(false OR false)|
+----------------+
|           false|
+----------------+
SELECT true or NULL;
+--------------+
|(true OR NULL)|
+--------------+
|          true|
+--------------+
SELECT false or NULL;
+--------------+
```

```
|(false OR NULL)|
+---------------+
|           NULL|
+---------------+
```

Funzioni della mappa

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedithe section called “Comandi SQL supportati”.

Funzione	Description
element_at (array, indice)	Restituisce un elemento dell'array in un determinato indice (a base 1).
element_at (mappa, chiave)	Restituisce il valore per una determinata chiave. La funzione restituisce NULL se la chiave non è contenuta nella mappa.
mappa (chiave0, valore0, chiave1, valore1,...)	Crea una mappa con le coppie date. key/value
map_concat (mappa,...)	Restituisce l'unione di tutte le mappe fornite
map_contains_key (mappa, chiave)	Restituisce true se la mappa contiene la chiave.
map_entries (mappa)	Restituisce un array non ordinato di tutte le voci nella mappa data.
map_from_arrays (chiavi, valori)	Crea una mappa con un paio di array dati. key/value Tutti gli elementi nelle chiavi non devono essere nulli
map_from_entries () arrayOfEntries	Restituisce una mappa creata dall'array di voci specificato.

Funzione	Description
map_keys (mappa)	Restituisce un array non ordinato contenente le chiavi della mappa.
map_values (mappa)	Restituisce un array non ordinato contenente i valori della mappa.
str_to_map (text [, pairDelim [,]]) keyValueDelim	Crea una mappa dopo aver suddiviso il testo in coppie chiave/valore utilizzando delimitatori. I delimitatori predefiniti sono ',' per `PairDelim` e ':' per ``. keyValueDelim Sia `PairDelim` che `` vengono trattati come espressioni regolari. keyValueDelim
try_element_at (array, indice)	Restituisce un elemento dell'array in un determinato indice (a base 1). Se l'indice è 0, il sistema genererà un errore. Se index < 0, accede agli elementi dall'ultimo al primo. La funzione restituisce sempre NULL se l'indice supera la lunghezza dell'array.
try_element_at (mappa, chiave)	Restituisce il valore per una determinata chiave. La funzione restituisce sempre NULL se la chiave non è contenuta nella mappa.

Esempi

```
-- element_at
SELECT element_at(array(1, 2, 3), 2);
+----------------------------+
|element_at(array(1, 2, 3), 2)|
+----------------------------+
|                           2|
+----------------------------+
SELECT element_at(map(1, 'a', 2, 'b'), 2);
+----------------------------+
|element_at(map(1, a, 2, b), 2)|
+----------------------------+
```

```
|                              b|
+-------------------------------+
-- map
SELECT map(1.0, '2', 3.0, '4');
+--------------------+
| map(1.0, 2, 3.0, 4)|
+--------------------+
|{1.0 -> 2, 3.0 -> 4}|
+--------------------+
-- map_concat
SELECT map_concat(map(1, 'a', 2, 'b'), map(3, 'c'));
+--------------------------------------+
|map_concat(map(1, a, 2, b), map(3, c))|
+--------------------------------------+
|                  {1 -> a, 2 -> b, ...|
+--------------------------------------+
-- map_contains_key
SELECT map_contains_key(map(1, 'a', 2, 'b'), 1);
+-----------------------------------+
|map_contains_key(map(1, a, 2, b), 1)|
+-----------------------------------+
|                               true|
+-----------------------------------+
SELECT map_contains_key(map(1, 'a', 2, 'b'), 3);
+-----------------------------------+
|map_contains_key(map(1, a, 2, b), 3)|
+-----------------------------------+
|                              false|
+-----------------------------------+
-- map_entries
SELECT map_entries(map(1, 'a', 2, 'b'));
+---------------------------+
|map_entries(map(1, a, 2, b))|
+---------------------------+
|            [{1, a}, {2, b}]|
+---------------------------+
-- map_from_arrays
SELECT map_from_arrays(array(1.0, 3.0), array('2', '4'));
+-------------------------------------------+
|map_from_arrays(array(1.0, 3.0), array(2, 4))|
+-------------------------------------------+
|                        {1.0 -> 2, 3.0 -> 4}|
+-------------------------------------------+
-- map_from_entries
```

```
SELECT map_from_entries(array(struct(1, 'a'), struct(2, 'b')));
+----------------------------------------------------+
|map_from_entries(array(struct(1, a), struct(2, b)))|
+----------------------------------------------------+
|                                {1 -> a, 2 -> b}|
+----------------------------------------------------+
-- map_keys
SELECT map_keys(map(1, 'a', 2, 'b'));
+-------------------------+
|map_keys(map(1, a, 2, b))|
+-------------------------+
|                   [1, 2]|
+-------------------------+
-- map_values
SELECT map_values(map(1, 'a', 2, 'b'));
+---------------------------+
|map_values(map(1, a, 2, b))|
+---------------------------+
|                     [a, b]|
+---------------------------+
-- str_to_map
SELECT str_to_map('a:1,b:2,c:3', ',', ':');
+-----------------------------+
|str_to_map(a:1,b:2,c:3, ,, :)|
+-----------------------------+
|         {a -> 1, b -> 2, ...|
+-----------------------------+
SELECT str_to_map('a');
+-------------------+
|str_to_map(a, ,, :)|
+-------------------+
|        {a -> NULL}|
+-------------------+
-- try_element_at
SELECT try_element_at(array(1, 2, 3), 2);
+---------------------------------+
|try_element_at(array(1, 2, 3), 2)|
+---------------------------------+
|                                2|
+---------------------------------+
SELECT try_element_at(map(1, 'a', 2, 'b'), 2);
+---------------------------------+
|try_element_at(map(1, a, 2, b), 2)|
+---------------------------------+
```

```
|                                   b|
+------------------------------------+
```

Funzioni matematiche

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Funzione	Description
espr 1% expr 2	Restituisce il resto dopo `expr1`/`expr2`.
espr 1 * espr 2	Restituisce `expr1`*`expr2`.
espr 1 + espr 2	Restituisce `expr1`+`expr2`.
espr 1 - espr 2	Restituisce `expr1`-`expr2`.
espr 1/espr 2	Restituisce `expr1`/`expr2`. Esegue sempre la divisione in virgola mobile.
abs (expr)	Restituisce il valore assoluto del valore numerico o dell'intervallo.
acos (expr)	Restituisce il coseno inverso (alias arc coseno) di `expr`, come se fosse calcolato da `java.lang.math.ACOS`.
acosh (expr)	Restituisce il coseno iperbolico inverso di `expr`.
asino (expr)	Restituisce il seno inverso (alias arc sine) l'arco sin di `expr`, come se fosse calcolato da `java.lang.Math.ASIN`.
asing (expr)	Restituisce il seno iperbolico inverso di `expr`.

Funzione	Description
satan (expr)	Restituisce la tangente inversa (alias arcotangente) di `expr`, come se fosse calcolata da `java.lang.math.ATAN`
atan2 (ExprY, expRX)	Restituisce l'angolo in radianti tra l'asse x positivo di un piano e il punto dato dalle coordinate (`expRX`, `Expry`), come se fosse calcolato da `java.lang.math.ATAN2`.
atanh (expr)	Restituisce la tangente iperbolica inversa di `expr`.
bin (expr)	Restituisce la rappresentazione in formato stringa del valore lungo `expr` rappresentato in binario.
terreno (expr, d)	Restituisce `expr` arrotondato alle cifre decimali `d` utilizzando la modalità di arrotonda mento HALF_EVEN.
cbrt (espr)	Restituisce la radice cubica di `expr`.
ceil (expr [, scala])	Restituisce il numero più piccolo dopo l'arroton damento per eccesso che non sia inferiore a `expr`. È possibile specificare un parametro `scale` opzionale per controllare il comportam ento di arrotondamento.
soffitto (expr [, scale])	Restituisce il numero più piccolo dopo l'arroton damento per eccesso che non sia inferiore a `expr`. È possibile specificare un parametro `scale` opzionale per controllare il comportam ento di arrotondamento.
conv (num, from_base, to_base)	Converte `num` da `from_base` a `to_base`.

Funzione	Description
cos (expr)	Restituisce il coseno di `expr`, come se fosse calcolato da `java.lang.math.COS`.
cosh (expr)	Restituisce il coseno iperbolico di `expr`, come se fosse calcolato da `java.lang.math.cosh`.
cot (expr)	Restituisce la cotangente di `expr`, come se fosse calcolata da `1/java.lang.math.TAN`.
csc (expr)	Restituisce la cosecante di `expr`, come se fosse calcolata da `1/java.lang.math.sin`.
gradi (expr)	Converte i radianti in gradi.
espr 1 div expr 2	Dividi `expr1` per `expr2`. Restituisce NULL se un operando è NULL o `expr2` è 0. Il risultato viene espresso troppo a lungo.
(e)	Restituisce il numero di Eulero, e.
exp (espr)	Restituisce e alla potenza di `expr`.
expm1 (expr) - Restituisce exp (`expr`)	1
fattoriale (expr)	Restituisce il fattoriale di `expr`. `expr` è [0.. 20]. Altrimenti, null.
pavimento (expr [, scala])	Restituisce il numero più grande dopo l'arroton damento per difetto che non sia maggiore di `expr`. È possibile specificare un parametro `scale` opzionale per controllare il comportam ento di arrotondamento.
massimo (expr,...)	Restituisce il valore massimo di tutti i parametri , ignorando i valori nulli.
esadecimale (expr)	Converte `expr` in esadecimale.

Funzione	Description
hypot (espr 1, espr 2)	Restituisce sqrt (`expr1`**2 + `expr2`**2).
minimo (expr,...)	Restituisce il valore minimo di tutti i parametri, ignorando i valori nulli.
ln (expr)	Restituisce il logaritmo naturale (base e) di `expr`.
log (base, expr)	Restituisce il logaritmo di `expr` con `base`.
log10 (expr)	Restituisce il logaritmo di `expr` con base 10.
log1p (expr)	Restituisce log (1 + `expr`).
log2 (espr)	Restituisce il logaritmo di `expr` con base 2.
espr 1 mod expr 2	Restituisce il resto dopo `expr1`/`expr2`.
negativo (expr)	Restituisce il valore negato di `expr`.
pi ()	Restituisce pi.
pmod (espr 1, espr 2)	Restituisce il valore positivo di `expr1` mod `expr2`.
positivo (expr)	Restituisce il valore di `expr`.
pow (espr 1, espr 2)	Eleva `expr1` alla potenza di `expr2`.
potenza (expr1, expr2)	Eleva `expr1` alla potenza di `expr2`.
radianti (expr)	Converte i gradi in radianti.
rand ([seme])	Restituisce un valore casuale con valori indipendenti e distribuiti in modo identico (i.i.d.) uniformemente distribuiti in [0, 1).

Funzione	Description
randn ([seme])	Restituisce un valore casuale con valori indipendenti e distribuiti in modo identico (i.i.d.) tratti dalla distribuzione normale standard.
casuale ([seme])	Restituisce un valore casuale con valori indipendenti e distribuiti in modo identico (i.i.d.) uniformemente in [0, 1).
stampa (expr)	Restituisce il valore doppio più vicino all'argomento ed è uguale a un numero intero matematico.
rotondo (expr, d)	Restituisce `expr` arrotondato alle cifre decimali `d` utilizzando la modalità di arrotonda mento HALF_UP.
sec (expr)	Restituisce la secante di `expr`, come se fosse calcolata da `1/java.lang.math.COS`.
shiftleft (base, expr)	Spostamento bit per bit a sinistra.
segno (expr)	Restituisce -1,0, 0,0 o 1,0 poiché `expr` è negativo, 0 o positivo.
segno (expr)	Restituisce -1,0, 0,0 o 1,0 poiché `expr` è negativo, 0 o positivo.
sin (espr)	Restituisce il seno di `expr`, come se fosse calcolato da `java.lang.math.sin`.
sinh (expr)	Restituisce il seno iperbolico di `expr`, come se fosse calcolato da `java.lang.math.SINH`.
sqrt (expr)	Restituisce la radice quadrata di `expr`.
tan (expr)	Restituisce la tangente di `expr`, come se fosse calcolata da `java.lang.math.tan`.

Funzione	Description
tanh (expr)	Restituisce la tangente iperbolica di `expr`, come se fosse calcolata da `java.lang.math.TANH`.
try_add (expr1, expr2)	Restituisce la somma di `expr1`e `expr2` e il risultato è nullo in caso di overflow. I tipi di input accettabili sono gli stessi con l'operatore `+`.
try_divide (dividendo, divisore)	Restituisce `dividend`/`divisor`. Esegue sempre la divisione in virgola mobile. Il suo risultato è sempre nullo se `expr2` è 0. `dividend` deve essere un numero o un intervallo. `divisore` deve essere un valore numerico.
try_multiply (espr 1, espr 2)	Restituisce `expr1`*`expr2` e il risultato è nullo in caso di overflow. I tipi di input accettabili sono gli stessi con l'operatore `*`.
try_sottract (expr1, expr2)	Restituisce `expr1`-`expr2` e il risultato è nullo in overflow. I tipi di input accettabili sono gli stessi con l'operatore `-`.
unhex (expr)	Converte `expr` esadecimale in binario.
width_bucket (value, min_value, max_value, num_bucket)	Restituisce il numero del bucket a cui verrà assegnato `value` in un istogramma di equilarghezza con bucket `num_bucket`, nell'intervallo da `min_value` a `max_value`.»

Esempi

```
-- %
SELECT 2 % 1.8;
+---------+
|(2 % 1.8)|
+---------+
```

```
|      0.2|
+---------+
SELECT MOD(2, 1.8);
+-----------+
|mod(2, 1.8)|
+-----------+
|        0.2|
+-----------+
-- *
SELECT 2 * 3;
+-------+
|(2 * 3)|
+-------+
|      6|
+-------+
-- +
SELECT 1 + 2;
+-------+
|(1 + 2)|
+-------+
|      3|
+-------+
-- -
SELECT 2 - 1;
+-------+
|(2 - 1)|
+-------+
|      1|
+-------+
-- /
SELECT 3 / 2;
+-------+
|(3 / 2)|
+-------+
|    1.5|
+-------+
SELECT 2L / 2L;
+-------+
|(2 / 2)|
+-------+
|    1.0|
+-------+
-- abs
SELECT abs(-1);
```

```
+-------+
|abs(-1)|
+-------+
|      1|
+-------+
SELECT abs(INTERVAL -'1-1' YEAR TO MONTH);
+----------------------------------+
|abs(INTERVAL '-1-1' YEAR TO MONTH)|
+----------------------------------+
|              INTERVAL '1-1' YE...|
+----------------------------------+
-- acos
SELECT acos(1);
+-------+
|ACOS(1)|
+-------+
|    0.0|
+-------+
SELECT acos(2);
+-------+
|ACOS(2)|
+-------+
|    NaN|
+-------+
-- acosh
SELECT acosh(1);
+--------+
|ACOSH(1)|
+--------+
|     0.0|
+--------+
SELECT acosh(0);
+--------+
|ACOSH(0)|
+--------+
|     NaN|
+--------+
-- asin
SELECT asin(0);
+-------+
|ASIN(0)|
+-------+
|    0.0|
+-------+
```

```
SELECT asin(2);
+-------+
|ASIN(2)|
+-------+
|    NaN|
+-------+
-- asinh
SELECT asinh(0);
+--------+
|ASINH(0)|
+--------+
|     0.0|
+--------+
-- atan
SELECT atan(0);
+-------+
|ATAN(0)|
+-------+
|    0.0|
+-------+
-- atan2
SELECT atan2(0, 0);
+-----------+
|ATAN2(0, 0)|
+-----------+
|        0.0|
+-----------+
-- atanh
SELECT atanh(0);
+--------+
|ATANH(0)|
+--------+
|     0.0|
+--------+
SELECT atanh(2);
+--------+
|ATANH(2)|
+--------+
|     NaN|
+--------+
-- bin
SELECT bin(13);
+-------+
|bin(13)|
```

```
+-------+
|   1101|
+-------+
SELECT bin(-13);
+--------------------+
|            bin(-13)|
+--------------------+
|1111111111111111...|
+--------------------+
SELECT bin(13.3);
+---------+
|bin(13.3)|
+---------+
|     1101|
+---------+
-- bround
SELECT bround(2.5, 0);
+--------------+
|bround(2.5, 0)|
+--------------+
|             2|
+--------------+
SELECT bround(25, -1);
+--------------+
|bround(25, -1)|
+--------------+
|            20|
+--------------+
-- cbrt
SELECT cbrt(27.0);
+----------+
|CBRT(27.0)|
+----------+
|       3.0|
+----------+
-- ceil
SELECT ceil(-0.1);
+----------+
|CEIL(-0.1)|
+----------+
|         0|
+----------+
SELECT ceil(5);
+-------+
```

```
|CEIL(5)|
+-------+
|      5|
+-------+
SELECT ceil(3.1411, 3);
+---------------+
|ceil(3.1411, 3)|
+---------------+
|          3.142|
+---------------+
SELECT ceil(3.1411, -3);
+----------------+
|ceil(3.1411, -3)|
+----------------+
|            1000|
+----------------+
-- ceiling
SELECT ceiling(-0.1);
+-------------+
|ceiling(-0.1)|
+-------------+
|            0|
+-------------+
SELECT ceiling(5);
+----------+
|ceiling(5)|
+----------+
|         5|
+----------+
SELECT ceiling(3.1411, 3);
+------------------+
|ceiling(3.1411, 3)|
+------------------+
|             3.142|
+------------------+
SELECT ceiling(3.1411, -3);
+-------------------+
|ceiling(3.1411, -3)|
+-------------------+
|               1000|
+-------------------+
-- conv
SELECT conv('100', 2, 10);
+----------------+
```

```
|conv(100, 2, 10)|
+----------------+
|               4|
+----------------+
SELECT conv(-10, 16, -10);
+------------------+
|conv(-10, 16, -10)|
+------------------+
|               -16|
+------------------+
-- cos
SELECT cos(0);
+------+
|COS(0)|
+------+
|   1.0|
+------+
-- cosh
SELECT cosh(0);
+-------+
|COSH(0)|
+-------+
|    1.0|
+-------+
-- cot
SELECT cot(1);
+------------------+
|            COT(1)|
+------------------+
|0.6420926159343306|
+------------------+
-- csc
SELECT csc(1);
+------------------+
|            CSC(1)|
+------------------+
|1.1883951057781212|
+------------------+
-- degrees
SELECT degrees(3.141592653589793);
+--------------------------+
|DEGREES(3.141592653589793)|
+--------------------------+
|                     180.0|
```

```
+--------------------------+
-- div
SELECT 3 div 2;
+---------+
|(3 div 2)|
+---------+
|        1|
+---------+
SELECT INTERVAL '1-1' YEAR TO MONTH div INTERVAL '-1' MONTH;
+-----------------------------------------------------------+
|(INTERVAL '1-1' YEAR TO MONTH div INTERVAL '-1' MONTH)|
+-----------------------------------------------------------+
|                                                        -13|
+-----------------------------------------------------------+
-- e
SELECT e();
+-----------------+
|              E()|
+-----------------+
|2.718281828459045|
+-----------------+
-- exp
SELECT exp(0);
+------+
|EXP(0)|
+------+
|   1.0|
+------+
-- expm1
SELECT expm1(0);
+--------+
|EXPM1(0)|
+--------+
|     0.0|
+--------+
-- factorial
SELECT factorial(5);
+------------+
|factorial(5)|
+------------+
|         120|
+------------+
-- floor
SELECT floor(-0.1);
```

```
+-----------+
|FLOOR(-0.1)|
+-----------+
|         -1|
+-----------+
SELECT floor(5);
+--------+
|FLOOR(5)|
+--------+
|       5|
+--------+
SELECT floor(3.1411, 3);
+----------------+
|floor(3.1411, 3)|
+----------------+
|           3.141|
+----------------+
SELECT floor(3.1411, -3);
+-----------------+
|floor(3.1411, -3)|
+-----------------+
|                0|
+-----------------+
-- greatest
SELECT greatest(10, 9, 2, 4, 3);
+------------------------+
|greatest(10, 9, 2, 4, 3)|
+------------------------+
|                      10|
+------------------------+
-- hex
SELECT hex(17);
+-------+
|hex(17)|
+-------+
|     11|
+-------+
SELECT hex('SQL');
+------------------+
|    hex(SQL)|
+------------------+
|53514C|
+------------------+
-- hypot
```

```
SELECT hypot(3, 4);
+-----------+
|HYPOT(3, 4)|
+-----------+
|        5.0|
+-----------+
-- least
SELECT least(10, 9, 2, 4, 3);
+---------------------+
|least(10, 9, 2, 4, 3)|
+---------------------+
|                    2|
+---------------------+
-- ln
SELECT ln(1);
+-----+
|ln(1)|
+-----+
|  0.0|
+-----+
-- log
SELECT log(10, 100);
+------------+
|LOG(10, 100)|
+------------+
|         2.0|
+------------+
-- log10
SELECT log10(10);
+---------+
|LOG10(10)|
+---------+
|      1.0|
+---------+
-- log1p
SELECT log1p(0);
+--------+
|LOG1P(0)|
+--------+
|     0.0|
+--------+
-- log2
SELECT log2(2);
+-------+
```

```
|LOG2(2)|
+-------+
|    1.0|
+-------+
-- mod
SELECT 2 % 1.8;
+---------+
|(2 % 1.8)|
+---------+
|      0.2|
+---------+
SELECT MOD(2, 1.8);
+-----------+
|mod(2, 1.8)|
+-----------+
|        0.2|
+-----------+
-- negative
SELECT negative(1);
+-----------+
|negative(1)|
+-----------+
|         -1|
+-----------+
-- pi
SELECT pi();
+-----------------+
|             PI()|
+-----------------+
|3.141592653589793|
+-----------------+
-- pmod
SELECT pmod(10, 3);
+-----------+
|pmod(10, 3)|
+-----------+
|          1|
+-----------+
SELECT pmod(-10, 3);
+------------+
|pmod(-10, 3)|
+------------+
|           2|
+------------+
```

```
-- positive
SELECT positive(1);
+-----+
|(+ 1)|
+-----+
|    1|
+-----+
-- pow
SELECT pow(2, 3);
+---------+
|pow(2, 3)|
+---------+
|      8.0|
+---------+
-- power
SELECT power(2, 3);
+-----------+
|POWER(2, 3)|
+-----------+
|        8.0|
+-----------+
-- radians
SELECT radians(180);
+-----------------+
|     RADIANS(180)|
+-----------------+
|3.141592653589793|
+-----------------+
-- rand
SELECT rand();
+------------------+
|            rand()|
+------------------+
|0.7211420708112387|
+------------------+
SELECT rand(0);
+------------------+
|           rand(0)|
+------------------+
|0.7604953758285915|
+------------------+
SELECT rand(null);
+------------------+
|        rand(NULL)|
```

```
+------------------+
|0.7604953758285915|
+------------------+
-- randn
SELECT randn();
+-------------------+
|            randn()|
+-------------------+
|-0.8175603217732732|
+-------------------+
SELECT randn(0);
+------------------+
|          randn(0)|
+------------------+
|1.6034991609278433|
+------------------+
SELECT randn(null);
+------------------+
|       randn(NULL)|
+------------------+
|1.6034991609278433|
+------------------+
-- random
SELECT random();
+-----------------+
|           rand()|
+-----------------+
|0.394205008255365|
+-----------------+
SELECT random(0);
+------------------+
|           rand(0)|
+------------------+
|0.7604953758285915|
+------------------+
SELECT random(null);
+------------------+
|        rand(NULL)|
+------------------+
|0.7604953758285915|
+------------------+
-- rint
SELECT rint(12.3456);
+-------------+
```

```
|rint(12.3456)|
+-------------+
|         12.0|
+-------------+
-- round
SELECT round(2.5, 0);
+-------------+
|round(2.5, 0)|
+-------------+
|            3|
+-------------+
-- sec
SELECT sec(0);
+------+
|SEC(0)|
+------+
|   1.0|
+------+
-- shiftleft
SELECT shiftleft(2, 1);
+---------------+
|shiftleft(2, 1)|
+---------------+
|              4|
+---------------+
-- sign
SELECT sign(40);
+--------+
|sign(40)|
+--------+
|     1.0|
+--------+
SELECT sign(INTERVAL -'100' YEAR);
+--------------------------+
|sign(INTERVAL '-100' YEAR)|
+--------------------------+
|                      -1.0|
+--------------------------+
-- signum
SELECT signum(40);
+----------+
|SIGNUM(40)|
+----------+
|       1.0|
```

```
+----------+
SELECT signum(INTERVAL -'100' YEAR);
+----------------------------+
|SIGNUM(INTERVAL '-100' YEAR)|
+----------------------------+
|                        -1.0|
+----------------------------+
-- sin
SELECT sin(0);
+------+
|SIN(0)|
+------+
|   0.0|
+------+
-- sinh
SELECT sinh(0);
+-------+
|SINH(0)|
+-------+
|    0.0|
+-------+
-- sqrt
SELECT sqrt(4);
+-------+
|SQRT(4)|
+-------+
|    2.0|
+-------+
-- tan
SELECT tan(0);
+------+
|TAN(0)|
+------+
|   0.0|
+------+
-- tanh
SELECT tanh(0);
+-------+
|TANH(0)|
+-------+
|    0.0|
+-------+
-- try_add
SELECT try_add(1, 2);
```

```
+-------------+
|try_add(1, 2)|
+-------------+
|            3|
+-------------+
SELECT try_add(2147483647, 1);
+----------------------+
|try_add(2147483647, 1)|
+----------------------+
|                  NULL|
+----------------------+
SELECT try_add(date'2021-01-01', 1);
+----------------------------+
|try_add(DATE '2021-01-01', 1)|
+----------------------------+
|                  2021-01-02|
+----------------------------+
SELECT try_add(date'2021-01-01', interval 1 year);
+-------------------------------------------+
|try_add(DATE '2021-01-01', INTERVAL '1' YEAR)|
+-------------------------------------------+
|                                 2022-01-01|
+-------------------------------------------+
SELECT try_add(timestamp'2021-01-01 00:00:00', interval 1 day);
+-----------------------------------------------------------+
|try_add(TIMESTAMP '2021-01-01 00:00:00', INTERVAL '1' DAY)|
+-----------------------------------------------------------+
|                                      2021-01-02 00:00:00|
+-----------------------------------------------------------+
SELECT try_add(interval 1 year, interval 2 year);
+-----------------------------------------+
|try_add(INTERVAL '1' YEAR, INTERVAL '2' YEAR)|
+-----------------------------------------+
|                           INTERVAL '3' YEAR|
+-----------------------------------------+
-- try_divide
SELECT try_divide(3, 2);
+----------------+
|try_divide(3, 2)|
+----------------+
|             1.5|
+----------------+
SELECT try_divide(2L, 2L);
+----------------+
```

```
|try_divide(2, 2)|
+----------------+
|             1.0|
+----------------+
SELECT try_divide(1, 0);
+----------------+
|try_divide(1, 0)|
+----------------+
|            NULL|
+----------------+
SELECT try_divide(interval 2 month, 2);
+--------------------------------+
|try_divide(INTERVAL '2' MONTH, 2)|
+--------------------------------+
|             INTERVAL '0-1' YE...|
+--------------------------------+
SELECT try_divide(interval 2 month, 0);
+--------------------------------+
|try_divide(INTERVAL '2' MONTH, 0)|
+--------------------------------+
|                            NULL|
+--------------------------------+
-- try_multiply
SELECT try_multiply(2, 3);
+------------------+
|try_multiply(2, 3)|
+------------------+
|                 6|
+------------------+
SELECT try_multiply(-2147483648, 10);
+-----------------------------+
|try_multiply(-2147483648, 10)|
+-----------------------------+
|                         NULL|
+-----------------------------+
SELECT try_multiply(interval 2 year, 3);
+---------------------------------+
|try_multiply(INTERVAL '2' YEAR, 3)|
+---------------------------------+
|              INTERVAL '6-0' YE...|
+---------------------------------+
-- try_subtract
SELECT try_subtract(2, 1);
+------------------+
```

```
|try_subtract(2, 1)|
+------------------+
|                 1|
+------------------+
SELECT try_subtract(-2147483648, 1);
+----------------------------+
|try_subtract(-2147483648, 1)|
+----------------------------+
|                        NULL|
+----------------------------+
SELECT try_subtract(date'2021-01-02', 1);
+---------------------------------+
|try_subtract(DATE '2021-01-02', 1)|
+---------------------------------+
|                       2021-01-01|
+---------------------------------+
SELECT try_subtract(date'2021-01-01', interval 1 year);
+-------------------------------------------------+
|try_subtract(DATE '2021-01-01', INTERVAL '1' YEAR)|
+-------------------------------------------------+
|                                       2020-01-01|
+-------------------------------------------------+
SELECT try_subtract(timestamp'2021-01-02 00:00:00', interval 1 day);
+-----------------------------------------------------------------+
|try_subtract(TIMESTAMP '2021-01-02 00:00:00', INTERVAL '1' DAY)|
+-----------------------------------------------------------------+
|                                              2021-01-01 00:00:00|
+-----------------------------------------------------------------+
SELECT try_subtract(interval 2 year, interval 1 year);
+-------------------------------------------------+
|try_subtract(INTERVAL '2' YEAR, INTERVAL '1' YEAR)|
+-------------------------------------------------+
|                                 INTERVAL '1' YEAR|
+-------------------------------------------------+
-- unhex
SELECT decode(unhex('53514C'), 'UTF-8');
+---------------------------------------+
|decode(unhex(53514C), UTF-8)|
+---------------------------------------+
|                                    SQL|
+---------------------------------------+
-- width_bucket
SELECT width_bucket(5.3, 0.2, 10.6, 5);
+-------------------------------+
```

```
|width_bucket(5.3, 0.2, 10.6, 5)|
+-------------------------------+
|                              3|
+-------------------------------+
SELECT width_bucket(-2.1, 1.3, 3.4, 3);
+-------------------------------+
|width_bucket(-2.1, 1.3, 3.4, 3)|
+-------------------------------+
|                              0|
+-------------------------------+
SELECT width_bucket(8.1, 0.0, 5.7, 4);
+-------------------------------+
|width_bucket(8.1, 0.0, 5.7, 4)|
+-------------------------------+
|                             5|
+-------------------------------+
SELECT width_bucket(-0.9, 5.2, 0.5, 2);
+-------------------------------+
|width_bucket(-0.9, 5.2, 0.5, 2)|
+-------------------------------+
|                              3|
+-------------------------------+
SELECT width_bucket(INTERVAL '0' YEAR, INTERVAL '0' YEAR, INTERVAL '10' YEAR, 10);
+-----------------------------------------------------------------------------+
|width_bucket(INTERVAL '0' YEAR, INTERVAL '0' YEAR, INTERVAL '10' YEAR, 10)|
+-----------------------------------------------------------------------------+
|                                                                            1|
+-----------------------------------------------------------------------------+
SELECT width_bucket(INTERVAL '1' YEAR, INTERVAL '0' YEAR, INTERVAL '10' YEAR, 10);
+-----------------------------------------------------------------------------+
|width_bucket(INTERVAL '1' YEAR, INTERVAL '0' YEAR, INTERVAL '10' YEAR, 10)|
+-----------------------------------------------------------------------------+
|                                                                            2|
+-----------------------------------------------------------------------------+
SELECT width_bucket(INTERVAL '0' DAY, INTERVAL '0' DAY, INTERVAL '10' DAY, 10);
+-------------------------------------------------------------------------+
|width_bucket(INTERVAL '0' DAY, INTERVAL '0' DAY, INTERVAL '10' DAY, 10)|
+-------------------------------------------------------------------------+
|                                                                        1|
+-------------------------------------------------------------------------+
SELECT width_bucket(INTERVAL '1' DAY, INTERVAL '0' DAY, INTERVAL '10' DAY, 10);
+-------------------------------------------------------------------------+
|width_bucket(INTERVAL '1' DAY, INTERVAL '0' DAY, INTERVAL '10' DAY, 10)|
+-------------------------------------------------------------------------+
```

```
|                                                                    2|
+-------------------------------------------------------------------+
```

Funzioni del generatore

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano queste funzioni SQL, consultathe section called “Comandi SQL supportati”.

Funzione	Description
esplodere (expr)	Separa gli elementi dell'array `expr` in più righe o gli elementi della mappa `expr` in più righe e colonne. Se non diversamente specificato, utilizza il nome di colonna predefinito `col` per gli elementi dell'array o `key` e `value` per gli elementi della mappa.
explode_outer (expr)	Separa gli elementi dell'array `expr` in più righe o gli elementi della mappa `expr` in più righe e colonne. Se non diversamente specificato, utilizza il nome di colonna predefinito `col` per gli elementi dell'array o `key` e `value` per gli elementi della mappa.
in linea (expr)	Esplode una serie di strutture in una tabella. Utilizza i nomi di colonna col1, col2, ecc. per impostazione predefinita, se non diversamente specificato.
inline_outer (expr)	Esplode una matrice di strutture in una tabella. Utilizza i nomi di colonna col1, col2, ecc. per impostazione predefinita, se non diversamente specificato.
posexplode (expr)	Separa gli elementi dell'array `expr` in più righe con posizioni o gli elementi della mappa `expr` in più righe e colonne con posizioni. Se non diversamente specificato, utilizza il nome di colonna `pos` per la posizione, `col` per gli elementi dell'array o `key` e `value` per gli elementi della mappa.
posexplode_outer (expr)	Separa gli elementi dell'array `expr` in più righe con posizioni o gli elementi della mappa `expr` in più righe e colonne con posizioni. Se

Funzione	Description
	non diversamente specificato, utilizza il nome di colonna `pos` per la posizione, `col` per gli elementi dell'array o `key` e `value` per gli elementi della mappa.
pila (n, expr1,..., exprk)	Separa `expr1`,..., `exprk` in `n` righe. Utilizza i nomi di colonna col0, col1, ecc. per impostazione predefinita, se non diversamente specificato.

Esempi

```
-- explode
SELECT explode(array(10, 20));
+---+
|col|
+---+
| 10|
| 20|
+---+

SELECT explode(collection => array(10, 20));
+---+
|col|
+---+
| 10|
| 20|
+---+

SELECT * FROM explode(collection => array(10, 20));
+---+
|col|
+---+
| 10|
| 20|
+---+

-- explode_outer
SELECT explode_outer(array(10, 20));
+---+
|col|
+---+
| 10|
```

```
| 20|
+---+

SELECT explode_outer(collection => array(10, 20));
+---+
|col|
+---+
| 10|
| 20|
+---+

SELECT * FROM explode_outer(collection => array(10, 20));
+---+
|col|
+---+
| 10|
| 20|
+---+

-- inline
SELECT inline(array(struct(1, 'a'), struct(2, 'b')));
+----+----+
|col1|col2|
+----+----+
|   1|   a|
|   2|   b|
+----+----+

-- inline_outer
SELECT inline_outer(array(struct(1, 'a'), struct(2, 'b')));
+----+----+
|col1|col2|
+----+----+
|   1|   a|
|   2|   b|
+----+----+

-- posexplode
SELECT posexplode(array(10,20));
+---+---+
|pos|col|
+---+---+
|  0| 10|
|  1| 20|
```

```
+---+---+

SELECT * FROM posexplode(array(10,20));
+---+---+
|pos|col|
+---+---+
|  0| 10|
|  1| 20|
+---+---+

-- posexplode_outer
SELECT posexplode_outer(array(10,20));
+---+---+
|pos|col|
+---+---+
|  0| 10|
|  1| 20|
+---+---+

SELECT * FROM posexplode_outer(array(10,20));
+---+---+
|pos|col|
+---+---+
|  0| 10|
|  1| 20|
+---+---+

-- stack
SELECT stack(2, 1, 2, 3);
+----+----+
|col0|col1|
+----+----+
|   1|   2|
|   3|NULL|
+----+----+
```

Clausola SELECT

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

OpenSearch SQL supporta un'`SELECT`istruzione utilizzata per recuperare i set di risultati da una o più tabelle. La sezione seguente descrive la sintassi generale delle query e i diversi costrutti di una query.

Sintassi

```
select_statement
[ { UNION | INTERSECT | EXCEPT } [ ALL | DISTINCT ] select_statement, ... ]
[ ORDER BY
    { expression [ ASC | DESC ] [ NULLS { FIRST | LAST } ]
    [ , ... ]
    }
]
[ SORT BY
    { expression [ ASC | DESC ] [ NULLS { FIRST | LAST } ]
    [ , ... ]
    }
]
[ WINDOW { named_window [ , WINDOW named_window, ... ] } ]
[ LIMIT { ALL | expression } ]
```

While `select_statement` è definito come:

```
SELECT [ ALL | DISTINCT ] { [ [ named_expression ] [ , ... ] ] }
FROM { from_item [ , ... ] }
[ PIVOT clause ]
[ UNPIVOT clause ]
[ LATERAL VIEW clause ] [ ... ]
[ WHERE boolean_expression ]
[ GROUP BY expression [ , ... ] ]
[ HAVING boolean_expression ]
```

Parametri

- TUTTO

  Seleziona tutte le righe corrispondenti dalla relazione ed è abilitata per impostazione predefinita.
- DISTINTO

  Seleziona tutte le righe corrispondenti dalla relazione dopo aver rimosso i duplicati nei risultati.
- named_expression

Un'espressione con un nome assegnato. In generale, denota un'espressione di colonna.

Sintassi: `expression [[AS] alias]`

- from_item

  Relazione tra tabelle

  Relazione di unione

  Relazione pivot

  Relazione Unpivot

  Funzione Table-value

  Tabella in linea

  `[ LATERAL ] ( Subquery )`

- PERNO

  La `PIVOT` clausola viene utilizzata per la prospettiva dei dati. È possibile ottenere i valori aggregati in base al valore di colonna specifico.

- UNPIVOT

  La `UNPIVOT` clausola trasforma le colonne in righe. È l'opposto di`PIVOT`, ad eccezione dell'aggregazione di valori.

- VISTA LATERALE

  La `LATERAL VIEW` clausola viene utilizzata insieme a funzioni generatrici come`EXPLODE`, che genereranno una tabella virtuale contenente una o più righe.

  `LATERAL VIEW`applicherà le righe a ciascuna riga di output originale.

- DOVE

  Filtra il risultato della `FROM` clausola in base ai predicati forniti.

- RAGGRUPPA PER

  Speciifica le espressioni utilizzate per raggruppare le righe.

Viene utilizzato insieme alle funzioni di aggregazione (MIN,, MAX COUNT SUMAVG, e così via) per raggruppare le righe in base alle espressioni di raggruppamento e ai valori aggregati di ciascun gruppo.

Quando una FILTER clausola è associata a una funzione aggregata, solo le righe corrispondenti vengono passate a quella funzione.

- AVENDO

  Speciifica i predicati in base ai quali GROUP BY vengono filtrate le righe prodotte da.

  La HAVING clausola viene utilizzata per filtrare le righe dopo l'esecuzione del raggruppamento.

  Se HAVING viene specificato senzaGROUP BY, indica un'espressione GROUP BY senza raggruppamento (aggregato globale).

- ORDINA PER

  Speciifica l'ordinamento delle righe del set completo di risultati della query.

  Le righe di output sono ordinate tra le partizioni.

  Questo parametro si esclude a vicenda con SORT BY e DISTRIBUTE BY e non può essere specificato insieme.

- ORDINA PER

  Speciifica l'ordine in base al quale le righe vengono ordinate all'interno di ciascuna partizione.

  Questo parametro si esclude a vicenda ORDER BY e non può essere specificato insieme.

- LIMITE

  Speciifica il numero massimo di righe che possono essere restituite da un'istruzione o una sottoquery.

  Questa clausola viene utilizzata principalmente in combinazione con per produrre un risultato ORDER BY deterministico.

- espressione booleana

  Specifica qualsiasi espressione che restituisce un tipo di risultato booleano.

  Due o più espressioni possono essere combinate insieme utilizzando gli operatori logici (,). AND OR

- espressione

  Speciifica una combinazione di uno o più valori, operatori e funzioni SQL che restituisce un valore.

- finestra_denominata

  Specificate gli alias per una o più specifiche della finestra di origine.

  È possibile fare riferimento alle specifiche della finestra di origine nelle definizioni della finestra dell'interrogazione.

Clausola WHERE

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

La `WHERE` clausola viene utilizzata per limitare i risultati della `FROM` clausola di una query o di una sottoquery in base alla condizione specificata.

Sintassi

```
WHERE boolean_expression
```

Parametri

- espressione booleana

  Specifica qualsiasi espressione che restituisce un risultato di tipo booleano.

  Due o più espressioni possono essere combinate insieme utilizzando gli operatori logici (,). `AND` `OR`

Esempi

```
CREATE TABLE person (id INT, name STRING, age INT);
INSERT INTO person VALUES
(100, 'John', 30),
(200, 'Mary', NULL),
(300, 'Mike', 80),
```

```
(400, 'Dan',  50);

-- Comparison operator in `WHERE` clause.
SELECT * FROM person WHERE id > 200 ORDER BY id;
+---+----+---+
| id|name|age|
+---+----+---+
|300|Mike| 80|
|400| Dan| 50|
+---+----+---+

-- Comparison and logical operators in `WHERE` clause.
SELECT * FROM person WHERE id = 200 OR id = 300 ORDER BY id;
+---+----+----+
| id|name| age|
+---+----+----+
|200|Mary|null|
|300|Mike|  80|
+---+----+----+

-- IS NULL expression in `WHERE` clause.
SELECT * FROM person WHERE id > 300 OR age IS NULL ORDER BY id;
+---+----+----+
| id|name| age|
+---+----+----+
|200|Mary|null|
|400| Dan|  50|
+---+----+----+

-- Function expression in `WHERE` clause.
SELECT * FROM person WHERE length(name) > 3 ORDER BY id;
+---+----+----+
| id|name| age|
+---+----+----+
|100|John|  30|
|200|Mary|null|
|300|Mike|  80|
+---+----+----+

-- `BETWEEN` expression in `WHERE` clause.
SELECT * FROM person WHERE id BETWEEN 200 AND 300 ORDER BY id;
+---+----+----+
| id|name| age|
+---+----+----+
```

```
|200|Mary|null|
|300|Mike|  80|
+---+----+----+

-- Scalar Subquery in `WHERE` clause.
SELECT * FROM person WHERE age > (SELECT avg(age) FROM person);
+---+----+---+
| id|name|age|
+---+----+---+
|300|Mike| 80|
+---+----+---+

-- Correlated Subquery in `WHERE` clause.
SELECT id FROM person
WHERE exists (SELECT id FROM person where id = 200);
+---+----+----+
|id |name|age |
+---+----+----+
|200|Mary|null|
+---+----+----+
```

Clausola GROUP BY

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedithe section called “Comandi SQL supportati”.

La `GROUP BY` clausola viene utilizzata per raggruppare le righe in base a un insieme di espressioni di raggruppamento specificate e calcolare le aggregazioni sul gruppo di righe in base a una o più funzioni di aggregazione specificate.

Il sistema esegue anche più aggregazioni per lo stesso record di input impostato tramite clausole,. `GROUPING SETS` `CUBE` `ROLLUP` Le espressioni di raggruppamento e le aggregazioni avanzate possono essere mescolate nella clausola e annidate in una `GROUP BY` clausola. `GROUPING SETS` Vedi maggiori dettagli nella sezione. `Mixed/Nested Grouping Analytics`

Quando una `FILTER` clausola è associata a una funzione aggregata, solo le righe corrispondenti vengono passate a quella funzione.

Sintassi

```
GROUP BY group_expression [ , group_expression [ , ... ] ] [ WITH { ROLLUP | CUBE } ]
GROUP BY { group_expression | { ROLLUP | CUBE | GROUPING SETS } (grouping_set
 [ , ...]) } [ , ... ]
```

Mentre le funzioni aggregate sono definite come:

```
aggregate_name ( [ DISTINCT ] expression [ , ... ] ) [ FILTER ( WHERE
 boolean_expression ) ]
```

Parametri

- espressione di gruppo

  Speciifica i criteri in base ai quali le righe vengono raggruppate. Il raggruppamento delle righe viene eseguito in base ai valori dei risultati delle espressioni di raggruppamento.

  Un'espressione di raggruppamento può essere simile al nome di una colonna`GROUP BY a`, alla posizione di una colonna o a un'espressione simile`GROUP BY 0`. `GROUP BY a + b`
- grouping_set

  Un set di raggruppamento è specificato da zero o più espressioni separate da virgole tra parentesi. Quando il set di raggruppamento contiene un solo elemento, le parentesi possono essere omesse.

  Ad esempio, `GROUPING SETS ((a), (b))` è equivalente a `GROUPING SETS (a, b)`.

  Sintassi: `{ ( [ expression [ , ... ] ] ) | expression }`
- SET DI RAGGRUPPAMENTO

  Raggruppa le righe per ogni set di raggruppamento specificato dopo. `GROUPING SETS`

  Ad esempio, `GROUP BY GROUPING SETS ((warehouse), (product))` è semanticamente equivalente all'unione dei risultati di e. `GROUP BY warehouse` `GROUP BY product` Questa clausola è un'abbreviazione di UNION ALL in cui ogni parte dell'`UNION ALL`operatore esegue l'aggregazione di ogni set di raggruppamento specificato nella clausola. `GROUPING SETS`

  Analogamente, `GROUP BY GROUPING SETS ((warehouse, product), (product), ())` è semanticamente equivalente all'unione dei risultati di un aggregato globale. `GROUP BY warehouse, product, GROUP BY product`
- ROLLUP

Speciifica più livelli di aggregazioni in una singola istruzione. Questa clausola viene utilizzata per calcolare aggregazioni basate su più set di raggruppamento. `ROLLUP`è un'abbreviazione di. `GROUPING SETS`

Ad esempio, `GROUP BY warehouse, product WITH ROLLUP or GROUP BY ROLLUP(warehouse, product)` è uguale a `GROUP BY GROUPING SETS((warehouse, product), (warehouse), ())`.

`GROUP BY ROLLUP(warehouse, product, (warehouse, location))` equivale a `GROUP BY GROUPING SETS((warehouse, product, location), (warehouse, product), (warehouse), ())`.

Gli N elementi di una specifica ROLLUP danno come risultato N+1 GROUPING SETS.

- CUBE

  La clausola CUBE viene utilizzata per eseguire aggregazioni basate sulla combinazione di colonne di raggruppamento specificate nella clausola GROUP BY. CUBE è l'abbreviazione di GROUPING SETS.

  Ad esempio, `GROUP BY warehouse, product WITH CUBE or GROUP BY CUBE(warehouse, product)` è uguale a `GROUP BY GROUPING SETS((warehouse, product), (warehouse), (product), ())`.

  `GROUP BY CUBE(warehouse, product, (warehouse, location))` equivale a `GROUP BY GROUPING SETS((warehouse, product, location), (warehouse, product), (warehouse, location), (product, warehouse, location), (warehouse), (product), (warehouse, product), ())`. Gli N elementi di una `CUBE` specifica danno come risultato 2^N. `GROUPING SETS`

- Analisi di raggruppamento misto/annidato

  Una `GROUP BY` clausola può includere più espressioni di gruppo e più espressioni. `CUBE|ROLLUP|GROUPING SETS` `GROUPING SETS`può anche avere `CUBE|ROLLUP|GROUPING SETS` clausole annidate, ad esempio,. `GROUPING SETS(ROLLUP(warehouse, location)` `CUBE(warehouse, location))` `GROUPING SETS(warehouse, GROUPING SETS(location, GROUPING SETS(ROLLUP(warehouse, location), CUBE(warehouse, location))))`

`CUBE|ROLLUP`è solo uno zucchero di sintassi per. `GROUPING SETS` Consulta le sezioni precedenti per sapere come `CUBE|ROLLUP` tradurre in`GROUPING SETS`. `group_expression`può essere trattato come un singolo gruppo `GROUPING SETS` in questo contesto.

Per i multipli `GROUPING SETS` della `GROUP BY` clausola, ne generiamo uno singolo `GROUPING SETS` facendo un prodotto incrociato dell'originale. `GROUPING SETS` Poiché è inserito `GROUPING SETS` nella `GROUPING SETS` clausola, prendiamo semplicemente i suoi set di raggruppamento e li eliminiamo.

Ad esempio, `GROUP BY warehouse, GROUPING SETS((product), ()), GROUPING SETS((location, size), (location), (size), ()) and GROUP BY warehouse, ROLLUP(product), CUBE(location, size)` è uguale a `GROUP BY GROUPING SETS( (warehouse, product, location, size), (warehouse, product, location), (warehouse, product, size), (warehouse, product), (warehouse, location, size), (warehouse, location), (warehouse, size), (warehouse))`.

`GROUP BY GROUPING SETS(GROUPING SETS(warehouse), GROUPING SETS((warehouse, product)))` equivale a `GROUP BY GROUPING SETS((warehouse), (warehouse, product))`.

- nome_aggregato

  Specificate il nome di una funzione aggregata (`MIN`,,, `MAX` `COUNT` `SUMAVG`, e così via).

- DISTINTO

  Rimuove i duplicati nelle righe di input prima che vengano passati alle funzioni aggregate.

- FILTRO

  Filtra le righe di input per le quali la `WHERE` clausola `boolean_expression` in the restituisce true vengono passate alla funzione di aggregazione; le altre righe vengono scartate.

Esempi

```
CREATE TABLE dealer (id INT, city STRING, car_model STRING, quantity INT);
INSERT INTO dealer VALUES
(100, 'Fremont', 'Honda Civic', 10),
(100, 'Fremont', 'Honda Accord', 15),
(100, 'Fremont', 'Honda CRV', 7),
(200, 'Dublin', 'Honda Civic', 20),
```

```
(200, 'Dublin', 'Honda Accord', 10),
(200, 'Dublin', 'Honda CRV', 3),
(300, 'San Jose', 'Honda Civic', 5),
(300, 'San Jose', 'Honda Accord', 8);

-- Sum of quantity per dealership. Group by `id`.
SELECT id, sum(quantity) FROM dealer GROUP BY id ORDER BY id;
+---+-------------+
| id|sum(quantity)|
+---+-------------+
|100|           32|
|200|           33|
|300|           13|
+---+-------------+

-- Use column position in GROUP by clause.
SELECT id, sum(quantity) FROM dealer GROUP BY 1 ORDER BY 1;
+---+-------------+
| id|sum(quantity)|
+---+-------------+
|100|           32|
|200|           33|
|300|           13|
+---+-------------+

-- Multiple aggregations.
-- 1. Sum of quantity per dealership.
-- 2. Max quantity per dealership.
SELECT id, sum(quantity) AS sum, max(quantity) AS max FROM dealer GROUP BY id ORDER BY
 id;
+---+---+---+
| id|sum|max|
+---+---+---+
|100| 32| 15|
|200| 33| 20|
|300| 13|  8|
+---+---+---+

-- Count the number of distinct dealer cities per car_model.
SELECT car_model, count(DISTINCT city) AS count FROM dealer GROUP BY car_model;
+------------+-----+
|   car_model|count|
+------------+-----+
| Honda Civic|    3|
```

```
|   Honda CRV|    2|
|Honda Accord|    3|
+------------+-----+

-- Sum of only 'Honda Civic' and 'Honda CRV' quantities per dealership.
SELECT id, sum(quantity) FILTER (
WHERE car_model IN ('Honda Civic', 'Honda CRV')
) AS `sum(quantity)` FROM dealer
GROUP BY id ORDER BY id;
+---+-------------+
| id|sum(quantity)|
+---+-------------+
|100|           17|
|200|           23|
|300|            5|
+---+-------------+

-- Aggregations using multiple sets of grouping columns in a single statement.
-- Following performs aggregations based on four sets of grouping columns.
-- 1. city, car_model
-- 2. city
-- 3. car_model
-- 4. Empty grouping set. Returns quantities for all city and car models.
SELECT city, car_model, sum(quantity) AS sum FROM dealer
GROUP BY GROUPING SETS ((city, car_model), (city), (car_model), ())
ORDER BY city;
+---------+------------+---+
|     city|   car_model|sum|
+---------+------------+---+
|     null|        null| 78|
|     null| HondaAccord| 33|
|     null|    HondaCRV| 10|
|     null|  HondaCivic| 35|
|   Dublin|        null| 33|
|   Dublin| HondaAccord| 10|
|   Dublin|    HondaCRV|  3|
|   Dublin|  HondaCivic| 20|
|  Fremont|        null| 32|
|  Fremont| HondaAccord| 15|
|  Fremont|    HondaCRV|  7|
|  Fremont|  HondaCivic| 10|
| San Jose|        null| 13|
| San Jose| HondaAccord|  8|
| San Jose|  HondaCivic|  5|
```

```
+---------+------------+---+

-- Group by processing with `ROLLUP` clause.
-- Equivalent GROUP BY GROUPING SETS ((city, car_model), (city), ())
SELECT city, car_model, sum(quantity) AS sum FROM dealer
GROUP BY city, car_model WITH ROLLUP
ORDER BY city, car_model;
+---------+------------+---+
|     city|   car_model|sum|
+---------+------------+---+
|     null|        null| 78|
|   Dublin|        null| 33|
|   Dublin| HondaAccord| 10|
|   Dublin|    HondaCRV|  3|
|   Dublin|  HondaCivic| 20|
|  Fremont|        null| 32|
|  Fremont| HondaAccord| 15|
|  Fremont|    HondaCRV|  7|
|  Fremont|  HondaCivic| 10|
| San Jose|        null| 13|
| San Jose| HondaAccord|  8|
| San Jose|  HondaCivic|  5|
+---------+------------+---+

-- Group by processing with `CUBE` clause.
-- Equivalent GROUP BY GROUPING SETS ((city, car_model), (city), (car_model), ())
SELECT city, car_model, sum(quantity) AS sum FROM dealer
GROUP BY city, car_model WITH CUBE
ORDER BY city, car_model;
+---------+------------+---+
|     city|   car_model|sum|
+---------+------------+---+
|     null|        null| 78|
|     null| HondaAccord| 33|
|     null|    HondaCRV| 10|
|     null|  HondaCivic| 35|
|   Dublin|        null| 33|
|   Dublin| HondaAccord| 10|
|   Dublin|    HondaCRV|  3|
|   Dublin|  HondaCivic| 20|
|  Fremont|        null| 32|
|  Fremont| HondaAccord| 15|
|  Fremont|    HondaCRV|  7|
|  Fremont|  HondaCivic| 10|
```

```
| San Jose|        null| 13|
| San Jose| HondaAccord|  8|
| San Jose|  HondaCivic|  5|
+---------+------------+---+

--Prepare data for ignore nulls example
CREATE TABLE person (id INT, name STRING, age INT);
INSERT INTO person VALUES
(100, 'Mary', NULL),
(200, 'John', 30),
(300, 'Mike', 80),
(400, 'Dan', 50);

--Select the first row in column age
SELECT FIRST(age) FROM person;
+--------------------+
| first(age, false)  |
+--------------------+
| NULL               |
+--------------------+

--Get the first row in column `age` ignore nulls,last row in column `id` and sum of
 column `id`.
SELECT FIRST(age IGNORE NULLS), LAST(id), SUM(id) FROM person;
+-------------------+------------------+----------+
| first(age, true)  | last(id, false)  | sum(id)  |
+-------------------+------------------+----------+
| 30                | 400              | 1000     |
+-------------------+------------------+----------+
```

Clausola HAVING

> **Note**
>
> Per vedere quali integrazioni di origini AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

La `HAVING` clausola viene utilizzata per filtrare i risultati prodotti da `GROUP BY` in base alla condizione specificata. Viene spesso utilizzata insieme a una `GROUP BY` clausola.

Sintassi

```
HAVING boolean_expression
```

Parametri

- espressione_booleana

  Specifica qualsiasi espressione che restituisce un tipo di risultato booleano. Due o più espressioni possono essere combinate insieme utilizzando gli operatori logici (,). `AND` `OR`

  Nota Le espressioni specificate nella `HAVING` clausola possono fare riferimento solo a:

  1. Costanti
  2. Espressioni che compaiono in `GROUP BY`
  3. Funzioni di aggregazione

Esempi

```
CREATE TABLE dealer (id INT, city STRING, car_model STRING, quantity INT);
INSERT INTO dealer VALUES
(100, 'Fremont', 'Honda Civic', 10),
(100, 'Fremont', 'Honda Accord', 15),
(100, 'Fremont', 'Honda CRV', 7),
(200, 'Dublin', 'Honda Civic', 20),
(200, 'Dublin', 'Honda Accord', 10),
(200, 'Dublin', 'Honda CRV', 3),
(300, 'San Jose', 'Honda Civic', 5),
(300, 'San Jose', 'Honda Accord', 8);

-- `HAVING` clause referring to column in `GROUP BY`.
SELECT city, sum(quantity) AS sum FROM dealer GROUP BY city HAVING city = 'Fremont';
+-------+---+
|   city|sum|
+-------+---+
|Fremont| 32|
+-------+---+

-- `HAVING` clause referring to aggregate function.
SELECT city, sum(quantity) AS sum FROM dealer GROUP BY city HAVING sum(quantity) > 15;
+-------+---+
|   city|sum|
+-------+---+
```

```
| Dublin| 33|
|Fremont| 32|
+-------+---+

-- `HAVING` clause referring to aggregate function by its alias.
SELECT city, sum(quantity) AS sum FROM dealer GROUP BY city HAVING sum > 15;
+-------+---+
|   city|sum|
+-------+---+
| Dublin| 33|
|Fremont| 32|
+-------+---+

-- `HAVING` clause referring to a different aggregate function than what is present in
-- `SELECT` list.
SELECT city, sum(quantity) AS sum FROM dealer GROUP BY city HAVING max(quantity) > 15;
+------+---+
|  city|sum|
+------+---+
|Dublin| 33|
+------+---+

-- `HAVING` clause referring to constant expression.
SELECT city, sum(quantity) AS sum FROM dealer GROUP BY city HAVING 1 > 0 ORDER BY city;
+--------+---+
|    city|sum|
+--------+---+
|  Dublin| 33|
| Fremont| 32|
|San Jose| 13|
+--------+---+

-- `HAVING` clause without a `GROUP BY` clause.
SELECT sum(quantity) AS sum FROM dealer HAVING sum(quantity) > 10;
+---+
|sum|
+---+
| 78|
+---+
```

Clausola ORDER BY

> **Note**
>
> Per vedere quali integrazioni di origini AWS dati supportano questo comando SQL, vedithe section called “Comandi SQL supportati”.

La `ORDER BY` clausola viene utilizzata per restituire le righe dei risultati in modo ordinato nell'ordine specificato dall'utente. A differenza della clausola SORT BY, questa clausola garantisce un ordine totale nell'output.

Sintassi

```
ORDER BY { expression [ sort_direction | nulls_sort_order ] [ , ... ] }
```

Parametri

- ORDINA PER

  Specifica un elenco di espressioni separate da virgole insieme a parametri opzionali `sort_direction` e `nulls_sort_order` che vengono utilizzate per ordinare le righe.
- sort_direction

  Facoltativamente specifica se ordinare le righe in ordine crescente o decrescente.

  I valori validi per la direzione di ordinamento sono per l'ordine crescente e `ASC` per il decrescente. `DESC`

  Se la direzione di ordinamento non è specificata in modo esplicito, per impostazione predefinita le righe vengono ordinate in modo crescente.

  Sintassi: `[ ASC | DESC ]`
- nulls_sort_order

  Specifica facoltativamente se vengono restituiti valori non NULL. `NULL` before/after

  Se null_sort_order non è specificato, ordina prima se l'ordinamento è e NULLS `NULLs` ordina per ultimo se l'ordinamento è`ASC`. `DESC`

1. Se `NULLS FIRST` è specificato, i valori NULL vengono restituiti per primi indipendentemente dal tipo di ordinamento.

2. Se `NULLS LAST` viene specificato, i valori NULL vengono restituiti per ultimi indipendentemente dal tipo di ordinamento.

Sintassi: `[ NULLS { FIRST | LAST } ]`

Esempi

```
CREATE TABLE person (id INT, name STRING, age INT);
INSERT INTO person VALUES
(100, 'John', 30),
(200, 'Mary', NULL),
(300, 'Mike', 80),
(400, 'Jerry', NULL),
(500, 'Dan',  50);

-- Sort rows by age. By default rows are sorted in ascending manner with NULL FIRST.
SELECT name, age FROM person ORDER BY age;
+-----+----+
| name| age|
+-----+----+
|Jerry|null|
| Mary|null|
| John|  30|
|  Dan|  50|
| Mike|  80|
+-----+----+

-- Sort rows in ascending manner keeping null values to be last.
SELECT name, age FROM person ORDER BY age NULLS LAST;
+-----+----+
| name| age|
+-----+----+
| John|  30|
|  Dan|  50|
| Mike|  80|
| Mary|null|
|Jerry|null|
+-----+----+
```

```
-- Sort rows by age in descending manner, which defaults to NULL LAST.
SELECT name, age FROM person ORDER BY age DESC;
+-----+----+
| name| age|
+-----+----+
| Mike|  80|
|  Dan|  50|
| John|  30|
|Jerry|null|
| Mary|null|
+-----+----+

-- Sort rows in ascending manner keeping null values to be first.
SELECT name, age FROM person ORDER BY age DESC NULLS FIRST;
+-----+----+
| name| age|
+-----+----+
|Jerry|null|
| Mary|null|
| Mike|  80|
|  Dan|  50|
| John|  30|
+-----+----+

-- Sort rows based on more than one column with each column having different
-- sort direction.
SELECT * FROM person ORDER BY name ASC, age DESC;
+---+-----+----+
| id| name| age|
+---+-----+----+
|500|  Dan|  50|
|400|Jerry|null|
|100| John|  30|
|200| Mary|null|
|300| Mike|  80|
+---+-----+----+
```

Clausola JOIN

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Un join SQL viene utilizzato per combinare righe di due relazioni in base a criteri di unione. La sezione seguente descrive la sintassi generale del join e i diversi tipi di join insieme a esempi.

Sintassi

```
relation INNER JOIN relation [ join_criteria ]
```

Parametri

- relazione

  Specifica la relazione da unire.
- join_type

  Specifica il tipo di join.

  Sintassi: `INNER | CROSS | LEFT OUTER`
- join_criteria

  Speciifica come le righe di una relazione verranno combinate con le righe di un'altra relazione.

  Sintassi: `ON boolean_expression | USING ( column_name [ , ... ] )`
- espressione_booleana

  Specifica un'espressione con un tipo restituito di tipo booleano.

Tipi di join

- Inner Join

  L'inner join deve essere specificato in modo esplicito. Seleziona le righe con valori corrispondenti in entrambe le relazioni.

Sintassi: `relation INNER JOIN relation [ join_criteria ]`

- Left Join

  Un left join restituisce tutti i valori della relazione sinistra e i valori corrispondenti della relazione destra oppure aggiunge NULL se non c'è corrispondenza. Viene anche chiamato left outer join.

  Sintassi: `relation LEFT OUTER JOIN relation [ join_criteria ]`

- Cross Join

  Un cross join restituisce il prodotto cartesiano di due relazioni.

  Sintassi: `relation CROSS JOIN relation [ join_criteria ]`

Esempi

```
-- Use employee and department tables to demonstrate different type of joins.
SELECT * FROM employee;
+---+-----+------+
| id| name|deptno|
+---+-----+------+
|105|Chloe|     5|
|103| Paul|     3|
|101| John|     1|
|102| Lisa|     2|
|104| Evan|     4|
|106|  Amy|     6|
+---+-----+------+
SELECT * FROM department;
+------+-----------+
|deptno|   deptname|
+------+-----------+
|     3|Engineering|
|     2|      Sales|
|     1|  Marketing|
+------+-----------+

-- Use employee and department tables to demonstrate inner join.
SELECT id, name, employee.deptno, deptname
FROM employee INNER JOIN department ON employee.deptno = department.deptno;
+---+-----+------+-----------|
| id| name|deptno|   deptname|
```

```
+---+-----+------+-----------|
|103| Paul|     3|Engineering|
|101| John|     1|  Marketing|
|102| Lisa|     2|      Sales|
+---+-----+------+-----------|

-- Use employee and department tables to demonstrate left join.
SELECT id, name, employee.deptno, deptname
FROM employee LEFT JOIN department ON employee.deptno = department.deptno;
+---+-----+------+-----------|
| id| name|deptno|   deptname|
+---+-----+------+-----------|
|105|Chloe|     5|       NULL|
|103| Paul|     3|Engineering|
|101| John|     1|  Marketing|
|102| Lisa|     2|      Sales|
|104| Evan|     4|       NULL|
|106|  Amy|     6|       NULL|
+---+-----+------+-----------|

-- Use employee and department tables to demonstrate cross join.
SELECT id, name, employee.deptno, deptname FROM employee CROSS JOIN department;
+---+-----+------+-----------|
| id| name|deptno|   deptname|
+---+-----+------+-----------|
|105|Chloe|     5|Engineering|
|105|Chloe|     5|  Marketing|
|105|Chloe|     5|      Sales|
|103| Paul|     3|Engineering|
|103| Paul|     3|  Marketing|
|103| Paul|     3|      Sales|
|101| John|     1|Engineering|
|101| John|     1|  Marketing|
|101| John|     1|      Sales|
|102| Lisa|     2|Engineering|
|102| Lisa|     2|  Marketing|
|102| Lisa|     2|      Sales|
|104| Evan|     4|Engineering|
|104| Evan|     4|  Marketing|
|104| Evan|     4|      Sales|
|106|  Amy|     4|Engineering|
|106|  Amy|     4|  Marketing|
|106|  Amy|     4|      Sales|
```

```
+---+-----+------+-----------|
```

Clausola LIMIT

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

La `LIMIT` clausola viene utilizzata per limitare il numero di righe restituite dall'istruzione. `SELECT` In generale, questa clausola viene utilizzata insieme a per `ORDER BY` garantire che i risultati siano deterministici.

Sintassi

```
LIMIT { ALL | integer_expression }
```

Parametri

- TUTTO

  Se specificato, la query restituisce tutte le righe. In altre parole, non viene applicato alcun limite se viene specificata questa opzione.
- espressione intera

  Specifica un'espressione pieghevole che restituisce un numero intero.

Esempi

```
CREATE TABLE person (name STRING, age INT);
INSERT INTO person VALUES
('Jane Doe', 25),
('Pat C', 18),
('Nikki W', 16),
('John D', 25),
('Juan L', 18),
('Jorge S', 16);
```

```
-- Select the first two rows.
SELECT name, age FROM person ORDER BY name LIMIT 2;
+-------+---+
|   name|age|
+-------+---+
|  Pat C| 18|
|Jorge S| 16|
+------+---+

-- Specifying ALL option on LIMIT returns all the rows.
SELECT name, age FROM person ORDER BY name LIMIT ALL;
+--------+---+
|    name|age|
+--------+---+
|   Pat C| 18|
| Jorge S| 16|
|  Juan L| 18|
|  John D| 25|
| Nikki W| 16|
|Jane Doe| 25|
+--------+---+

-- A function expression as an input to LIMIT.
SELECT name, age FROM person ORDER BY name LIMIT length('OPENSEARCH');
+-------+---+
|   name|age|
+-------+---+
|  Pat C| 18|
|Jorge S| 16|
| Juan L| 18|
| John D| 25|
|Nikki W| 16|
+-------+---+
```

Clausola CASE

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

La CASE clausola utilizza una regola per restituire un risultato specifico in base alla condizione specificata, in modo simile alle istruzioni if/else in altri linguaggi di programmazione.

Sintassi

```
CASE [ expression ] { WHEN boolean_expression THEN then_expression } [ ... ]
[ ELSE else_expression ]
END
```

Parametri

- espressione_booleana

  Specifica qualsiasi espressione che restituisce un tipo di risultato booleano.

  Due o più espressioni possono essere combinate insieme utilizzando gli operatori logici (,). AND OR
- then_expression

  Specifica l'espressione then in base alla condizione boolean_expression.

  then_expressione else_expression dovrebbero essere tutti dello stesso tipo o compatibili con un tipo comune.
- else_expression

  Specifica l'espressione predefinita.

  then_expressione else_expression devono essere tutti dello stesso tipo o compatibili con un tipo comune.

Esempi

```
CREATE TABLE person (id INT, name STRING, age INT);
INSERT INTO person VALUES
(100, 'John', 30),
(200, 'Mary', NULL),
(300, 'Mike', 80),
(400, 'Dan', 50);
SELECT id, CASE WHEN id > 200 THEN 'bigger' ELSE 'small' END FROM person;
+------+-------------------------------------------------+
|  id  | CASE WHEN (id > 200) THEN bigger ELSE small END  |
+------+-------------------------------------------------+
```

```
| 100  | small                                                |
| 200  | small                                                |
| 300  | bigger                                               |
| 400  | bigger                                               |
+------+------------------------------------------------------+
SELECT id, CASE id WHEN 100 then 'bigger' WHEN  id > 300 THEN '300' ELSE 'small' END
 FROM person;
+------
+-------------------------------------------------------------------------------------------------------
+
|  id  | CASE WHEN (id = 100) THEN bigger WHEN (id = CAST((id > 300) AS INT)) THEN 300
 ELSE small END  |
+------
+-------------------------------------------------------------------------------------------------------
+
| 100  | bigger
                |
| 200  | small
                |
| 300  | small
                |
| 400  | small
                |
+------
+-------------------------------------------------------------------------------------------------------
+
```

Espressione di tabella comune

> **Note**
>
> Per vedere quali integrazioni di origini AWS dati supportano questo comando SQL, vedithe section called “Comandi SQL supportati”.

Un'espressione di tabella comune (CTE) definisce un set di risultati temporaneo a cui un utente può fare riferimento, possibilmente più volte, nell'ambito di un'istruzione SQL. Un CTE viene utilizzato principalmente in una `SELECT` dichiarazione.

Sintassi

```
WITH common_table_expression [ , ... ]
```

While `common_table_expression` è definito come:

```
Syntexpression_name [ ( column_name [ , ... ] ) ] [ AS ] ( query )
```

Parametri

- nome_espressione

  Specificate un nome per l'espressione della tabella comune.
- query

  Una `SELECT` dichiarazione.

Esempi

```
-- CTE with multiple column aliases
WITH t(x, y) AS (SELECT 1, 2)
SELECT * FROM t WHERE x = 1 AND y = 2;
+---+---+
|  x|  y|
+---+---+
|  1|  2|
+---+---+

-- CTE in CTE definition
WITH t AS (
WITH t2 AS (SELECT 1)
SELECT * FROM t2
)
SELECT * FROM t;
+---+
|  1|
+---+
|  1|
+---+

-- CTE in subquery
SELECT max(c) FROM (
WITH t(c) AS (SELECT 1)
SELECT * FROM t
);
+------+
```

```
|max(c)|
+------+
|     1|
+------+

-- CTE in subquery expression
SELECT (
WITH t AS (SELECT 1)
SELECT * FROM t
);
+----------------+
|scalarsubquery()|
+----------------+
|               1|
+----------------+

-- CTE in CREATE VIEW statement
CREATE VIEW v AS
WITH t(a, b, c, d) AS (SELECT 1, 2, 3, 4)
SELECT * FROM t;
SELECT * FROM v;
+---+---+---+---+
|  a|  b|  c|  d|
+---+---+---+---+
|  1|  2|  3|  4|
+---+---+---+---+
```

EXPLAIN

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedithe section called “Comandi SQL supportati”.

L'`EXPLAIN`istruzione viene utilizzata per fornire piani logici/fisici per un'istruzione di input. Per impostazione predefinita, questa clausola fornisce informazioni solo su un piano fisico.

Sintassi

```
EXPLAIN [ EXTENDED | CODEGEN | COST | FORMATTED ] statement
```

Parametri

- ESTESO

  Genera piano logico analizzato, piano logico analizzato, piano logico ottimizzato e piano fisico.

  Il piano logico analizzato è un piano irrisolto estratto dalla query.

  I piani logici analizzati si trasformano e si traducono in oggetti completamente un`resolvedAttribute` tipizzatiun`resolvedRelation`.

  Il piano logico ottimizzato si trasforma attraverso una serie di regole di ottimizzazione, dando origine al piano fisico.
- CODEGEN

  Genera il codice per l'eventuale dichiarazione e un piano fisico.
- COSTO

  Se le statistiche del nodo del piano sono disponibili, genera un piano logico e le statistiche.
- FORMATTATO

  Genera due sezioni: uno schema fisico del piano e i dettagli del nodo.
- Istruzione

  Specifica un'istruzione SQL da spiegare.

Esempi

```
-- Default Output
EXPLAIN select k, sum(v) from values (1, 2), (1, 3) t(k, v) group by k;
+----------------------------------------------------+
|                                                plan|
+----------------------------------------------------+
| == Physical Plan ==
*(2) HashAggregate(keys=[k#33], functions=[sum(cast(v#34 as bigint))])
+- Exchange hashpartitioning(k#33, 200), true, [id=#59]
+- *(1) HashAggregate(keys=[k#33], functions=[partial_sum(cast(v#34 as bigint))])
+- *(1) LocalTableScan [k#33, v#34]
|
+----------------------------------------------------
```

```
-- Using Extended
EXPLAIN EXTENDED select k, sum(v) from values (1, 2), (1, 3) t(k, v) group by k;
+----------------------------------------------------+
|                                                plan|
+----------------------------------------------------+
| == Parsed Logical Plan ==
'Aggregate ['k], ['k, unresolvedalias('sum('v), None)]
 +- 'SubqueryAlias `t`
+- 'UnresolvedInlineTable [k, v], [List(1, 2), List(1, 3)]

 == Analyzed Logical Plan ==
 k: int, sum(v): bigint
 Aggregate [k#47], [k#47, sum(cast(v#48 as bigint)) AS sum(v)#50L]
 +- SubqueryAlias `t`
    +- LocalRelation [k#47, v#48]

 == Optimized Logical Plan ==
 Aggregate [k#47], [k#47, sum(cast(v#48 as bigint)) AS sum(v)#50L]
 +- LocalRelation [k#47, v#48]

 == Physical Plan ==
 *(2) HashAggregate(keys=[k#47], functions=[sum(cast(v#48 as bigint))], output=[k#47,
 sum(v)#50L])
+- Exchange hashpartitioning(k#47, 200), true, [id=#79]
   +- *(1) HashAggregate(keys=[k#47], functions=[partial_sum(cast(v#48 as bigint))],
 output=[k#47, sum#52L])
    +- *(1) LocalTableScan [k#47, v#48]
|
+----------------------------------------------------+

-- Using Formatted
EXPLAIN FORMATTED select k, sum(v) from values (1, 2), (1, 3) t(k, v) group by k;
+----------------------------------------------------+
|                                                plan|
+----------------------------------------------------+
| == Physical Plan ==
 * HashAggregate (4)
 +- Exchange (3)
    +- * HashAggregate (2)
       +- * LocalTableScan (1)


 (1) LocalTableScan [codegen id : 1]
 Output: [k#19, v#20]
```

```
 (2) HashAggregate [codegen id : 1]
 Input: [k#19, v#20]

 (3) Exchange
 Input: [k#19, sum#24L]

 (4) HashAggregate [codegen id : 2]
 Input: [k#19, sum#24L]
|
+-----------------------------------------------------+
```

Clausola LATERAL SUBQUERY

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

`LATERAL SUBQUERY`è una sottoquery preceduta dalla parola chiave. `LATERAL` Fornisce un modo per fare riferimento alle colonne della clausola precedente. `FROM` Senza la `LATERAL` parola chiave, le sottoquery possono fare riferimento solo alle colonne della query esterna, ma non alla clausola. `FROM` `LATERAL SUBQUERY`rende le interrogazioni complicate più semplici ed efficienti.

Sintassi

```
[ LATERAL ] primary_relation [ join_relation ]
```

Parametri

- relazione_primaria

  Specifica la relazione principale. Può essere uno dei seguenti:

  1. Relazione tra tabelle
  2. Interrogazione con alias

     Sintassi: `( query ) [ [ AS ] alias ]`
  3. Relazione con alias

```
Syntax: ( relation ) [ [ AS ] alias ]
```

Esempi

```
CREATE TABLE t1 (c1 INT, c2 INT);
INSERT INTO t1 VALUES (0, 1), (1, 2);
CREATE TABLE t2 (c1 INT, c2 INT);
INSERT INTO t2 VALUES (0, 2), (0, 3);
SELECT * FROM t1,
LATERAL (SELECT * FROM t2 WHERE t1.c1 = t2.c1);
+--------+-------+--------+-------+
|  t1.c1 | t1.c2 |  t2.c1 | t2.c2 |
+-------+--------+--------+-------+
|    0   |   1   |    0   |   3   |
|    0   |   1   |    0   |   2   |
+-------+--------+--------+-------+
SELECT a, b, c FROM t1,
LATERAL (SELECT c1 + c2 AS a),
LATERAL (SELECT c1 - c2 AS b),
LATERAL (SELECT a * b AS c);
+--------+-------+--------+
|    a   |   b   |    c   |
+-------+--------+--------+
|    3   |  -1   |   -3   |
|    1   |  -1   |   -1   |
+-------+--------+--------+
```

Clausola LATERAL VIEW

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

La `LATERAL VIEW` clausola viene utilizzata insieme a funzioni generatrici come`EXPLODE`, che genereranno una tabella virtuale contenente una o più righe. `LATERAL VIEW`applicherà le righe a ciascuna riga di output originale.

Sintassi

```
LATERAL VIEW [ OUTER ] generator_function ( expression [ , ... ] ) [ table_alias ] AS
 column_alias [ , ... ]
```

Parametri

- ESTERNO

  Se OUTER specificato, restituisce null se un input array/map è vuoto o nullo.

- generator_function

  Specifica una funzione generatrice (EXPLODEINLINE, e così via).

- table_alias

  L'alias per, che è facoltativogenerator_function.

- column_alias

  Elenca gli alias di colonna digenerator_function, che possono essere utilizzati nelle righe di output.

  È possibile avere più alias se sono presenti più generator_function colonne di output.

Esempi

```
CREATE TABLE person (id INT, name STRING, age INT, class INT, address STRING);
INSERT INTO person VALUES
(100, 'John', 30, 1, 'Street 1'),
(200, 'Mary', NULL, 1, 'Street 2'),
(300, 'Mike', 80, 3, 'Street 3'),
(400, 'Dan', 50, 4, 'Street 4');
SELECT * FROM person
LATERAL VIEW EXPLODE(ARRAY(30, 60)) tableName AS c_age
LATERAL VIEW EXPLODE(ARRAY(40, 80)) AS d_age;
+------+-------+-------+--------+-----------+--------+--------+
|  id  | name  |  age  | class  |  address  | c_age  | d_age  |
+------+-------+-------+--------+-----------+--------+--------+
| 100  | John  | 30    | 1      | Street 1  | 30     | 40     |
| 100  | John  | 30    | 1      | Street 1  | 30     | 80     |
| 100  | John  | 30    | 1      | Street 1  | 60     | 40     |
| 100  | John  | 30    | 1      | Street 1  | 60     | 80     |
| 200  | Mary  | NULL  | 1      | Street 2  | 30     | 40     |
```

```
| 200  | Mary  | NULL  | 1      | Street 2  | 30     | 80     |
| 200  | Mary  | NULL  | 1      | Street 2  | 60     | 40     |
| 200  | Mary  | NULL  | 1      | Street 2  | 60     | 80     |
| 300  | Mike  | 80    | 3      | Street 3  | 30     | 40     |
| 300  | Mike  | 80    | 3      | Street 3  | 30     | 80     |
| 300  | Mike  | 80    | 3      | Street 3  | 60     | 40     |
| 300  | Mike  | 80    | 3      | Street 3  | 60     | 80     |
| 400  | Dan   | 50    | 4      | Street 4  | 30     | 40     |
| 400  | Dan   | 50    | 4      | Street 4  | 30     | 80     |
| 400  | Dan   | 50    | 4      | Street 4  | 60     | 40     |
| 400  | Dan   | 50    | 4      | Street 4  | 60     | 80     |
+------+-------+-------+--------+-----------+--------+--------+
SELECT c_age, COUNT(1) FROM person
LATERAL VIEW EXPLODE(ARRAY(30, 60)) AS c_age
LATERAL VIEW EXPLODE(ARRAY(40, 80)) AS d_age
GROUP BY c_age;
+--------+-----------+
| c_age  | count(1)  |
+--------+-----------+
| 60     | 8         |
| 30     | 8         |
+--------+-----------+
SELECT * FROM person
LATERAL VIEW EXPLODE(ARRAY()) tableName AS c_age;
+-----+-------+------+--------+----------+--------+
| id  | name  | age  | class  | address  | c_age  |
+-----+-------+------+--------+----------+--------+
+-----+-------+------+--------+----------+--------+
SELECT * FROM person
LATERAL VIEW OUTER EXPLODE(ARRAY()) tableName AS c_age;
+------+-------+-------+--------+-----------+--------+
|  id  | name  |  age  | class  |  address  | c_age  |
+------+-------+-------+--------+-----------+--------+
| 100  | John  | 30    | 1      | Street 1  | NULL   |
| 200  | Mary  | NULL  | 1      | Street 2  | NULL   |
| 300  | Mike  | 80    | 3      | Street 3  | NULL   |
| 400  | Dan   | 50    | 4      | Street 4  | NULL   |
+------+-------+-------+--------+-----------+--------+
```

Predicato LIKE

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Un `LIKE` predicato viene utilizzato per cercare uno schema specifico. Questo predicato supporta anche più modelli con i quantificatori che includonoANY, e. `SOME` `ALL`

Sintassi

```
[ NOT ] { LIKE search_pattern [ ESCAPE esc_char ] | [ RLIKE | REGEXP ] regex_pattern }
[ NOT ] { LIKE quantifiers ( search_pattern [ , ... ]) }
```

Parametri

- modello_di ricerca

  Specifica uno schema di stringhe in cui eseguire la ricerca mediante la clausola LIKE. Può contenere caratteri speciali per la corrispondenza dei modelli:

  - %corrisponde a zero o più caratteri.
  - _corrisponde esattamente a un carattere.
- esc_char

  Specifica il carattere di escape. Il carattere di escape predefinito è. \
- regex_pattern

  Specifica un modello di ricerca per espressioni regolari in cui eseguire la ricerca mediante la clausola or. `RLIKE` `REGEXP`
- quantificatori

  Speciifica che i quantificatori dei predicati includono, e. `ANY` `SOME` `ALL`

  `ANY`o `SOME` significa che se uno dei modelli corrisponde all'input, restituisce true.

  `ALL`significa se tutti i pattern corrispondono all'input, allora restituisce true.

Esempi

```
CREATE TABLE person (id INT, name STRING, age INT);
INSERT INTO person VALUES
(100, 'John', 30),
(200, 'Mary', NULL),
(300, 'Mike', 80),
(400, 'Dan',  50),
(500, 'Evan_w', 16);
SELECT * FROM person WHERE name LIKE 'M%';
+---+----+----+
| id|name| age|
+---+----+----+
|300|Mike|  80|
|200|Mary|null|
+---+----+----+
SELECT * FROM person WHERE name LIKE 'M_ry';
+---+----+----+
| id|name| age|
+---+----+----+
|200|Mary|null|
+---+----+----+
SELECT * FROM person WHERE name NOT LIKE 'M_ry';
+---+------+---+
| id|  name|age|
+---+------+---+
|500|Evan_W| 16|
|300|  Mike| 80|
|100|  John| 30|
|400|   Dan| 50|
+---+------+---+
SELECT * FROM person WHERE name RLIKE 'M+';
+---+----+----+
| id|name| age|
+---+----+----+
|300|Mike|  80|
|200|Mary|null|
+---+----+----+
SELECT * FROM person WHERE name REGEXP 'M+';
+---+----+----+
| id|name| age|
+---+----+----+
|300|Mike|  80|
|200|Mary|null|
```

```
+---+----+----+
SELECT * FROM person WHERE name LIKE '%\_%';
+---+------+---+
| id|  name|age|
+---+------+---+
|500|Evan_W| 16|
+---+------+---+
SELECT * FROM person WHERE name LIKE '%$_%' ESCAPE '$';
+---+------+---+
| id|  name|age|
+---+------+---+
|500|Evan_W| 16|
+---+------+---+
SELECT * FROM person WHERE name LIKE ALL ('%an%', '%an');
+---+----+----+
| id|name| age|
+---+----+----+
|400| Dan|  50|
+---+----+----+
SELECT * FROM person WHERE name LIKE ANY ('%an%', '%an');
+---+------+---+
| id|  name|age|
+---+------+---+
|400|   Dan| 50|
|500|Evan_W| 16|
+---+------+---+
SELECT * FROM person WHERE name LIKE SOME ('%an%', '%an');
+---+------+---+
| id|  name|age|
+---+------+---+
|400|   Dan| 50|
|500|Evan_W| 16|
+---+------+---+
SELECT * FROM person WHERE name NOT LIKE ALL ('%an%', '%an');
+---+----+----+
| id|name| age|
+---+----+----+
|100|John|  30|
|200|Mary|null|
|300|Mike|  80|
+---+----+----+
SELECT * FROM person WHERE name NOT LIKE ANY ('%an%', '%an');
+---+------+----+
| id|  name| age|
```

```
+---+------+----+
|100|  John|  30|
|200|  Mary|null|
|300|  Mike|  80|
|500|Evan_W|  16|
+---+------+----+
SELECT * FROM person WHERE name NOT LIKE SOME ('%an%', '%an');
+---+------+----+
| id|  name| age|
+---+------+----+
|100|  John|  30|
|200|  Mary|null|
|300|  Mike|  80|
|500|Evan_W|  16|
+---+------+----+
```

OFFSET

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedithe section called “Comandi SQL supportati”.

La `OFFSET` clausola viene utilizzata per specificare il numero di righe da saltare prima di iniziare a restituire le righe restituite dall'istruzione. `SELECT` In generale, questa clausola viene utilizzata insieme a per `ORDER BY` garantire che i risultati siano deterministici.

Sintassi

```
OFFSET integer_expression
```

Parametri

- espressione intera

  Specifica un'espressione pieghevole che restituisce un numero intero.

Esempi

```
CREATE TABLE person (name STRING, age INT);
```

```
INSERT INTO person VALUES
('Jane Doe', 25),
('Pat C', 18),
('Nikki W', 16),
('Juan L', 25),
('John D', 18),
('Jorge S', 16);

-- Skip the first two rows.
SELECT name, age FROM person ORDER BY name OFFSET 2;
+-------+---+
|   name|age|
+-------+---+
| John D| 18|
| Juan L| 25|
|Nikki W| 16|
|Jane Doe| 25|
+-------+---+

-- Skip the first two rows and returns the next three rows.
SELECT name, age FROM person ORDER BY name LIMIT 3 OFFSET 2;
+-------+---+
|   name|age|
+-------+---+
| John D| 18|
| Juan L| 25|
|Nikki W| 16|
+-------+---+

-- A function expression as an input to OFFSET.
SELECT name, age FROM person ORDER BY name OFFSET length('WAGON');
+-------+---+
|   name|age|
+-------+---+
|Jane Doe| 25|
+-------+---+
```

Clausola PIVOT

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

La `PIVOT` clausola viene utilizzata per la prospettiva dei dati. Possiamo ottenere i valori aggregati in base a valori di colonna specifici, che verranno trasformati in più colonne utilizzate nella `SELECT` clausola. La `PIVOT` clausola può essere specificata dopo il nome della tabella o della sottoquery.

Sintassi

```
PIVOT ( { aggregate_expression [ AS aggregate_expression_alias ] } [ , ... ] FOR
 column_list IN ( expression_list ) )
```

Parametri

- aggregate_expression

  Specifica un'espressione (`SUM(a)` aggregata e così via.). `COUNT(DISTINCT b)`
- aggregate_expression_alias

  Specifica un alias per l'espressione aggregata.
- column_list

  Contiene colonne nella `FROM` clausola, che specifica le colonne che si desidera sostituire con nuove colonne. È possibile utilizzare parentesi per racchiudere le colonne, ad esempio. `(c1, c2)`
- expression_list

  Speciifica nuove colonne, che vengono utilizzate per abbinare i valori in `column_list` come condizione di aggregazione. È inoltre possibile aggiungere alias per esse.

Esempi

```
CREATE TABLE person (id INT, name STRING, age INT, class INT, address STRING);
INSERT INTO person VALUES
(100, 'John', 30, 1, 'Street 1'),
```

```
(200, 'Mary', NULL, 1, 'Street 2'),
(300, 'Mike', 80, 3, 'Street 3'),
(400, 'Dan', 50, 4, 'Street 4');
SELECT * FROM person
PIVOT (
SUM(age) AS a, AVG(class) AS c
FOR name IN ('John' AS john, 'Mike' AS mike)
);
+------+-----------+---------+---------+---------+---------+
|  id  |  address  | john_a  | john_c  | mike_a  | mike_c  |
+------+-----------+---------+---------+---------+---------+
| 200  | Street 2  | NULL    | NULL    | NULL    | NULL    |
| 100  | Street 1  | 30      | 1.0     | NULL    | NULL    |
| 300  | Street 3  | NULL    | NULL    | 80      | 3.0     |
| 400  | Street 4  | NULL    | NULL    | NULL    | NULL    |
+------+-----------+---------+---------+---------+---------+
SELECT * FROM person
PIVOT (
SUM(age) AS a, AVG(class) AS c
FOR (name, age) IN (('John', 30) AS c1, ('Mike', 40) AS c2)
);
+------+-----------+-------+-------+-------+-------+
|  id  |  address  | c1_a  | c1_c  | c2_a  | c2_c  |
+------+-----------+-------+-------+-------+-------+
| 200  | Street 2  | NULL  | NULL  | NULL  | NULL  |
| 100  | Street 1  | 30    | 1.0   | NULL  | NULL  |
| 300  | Street 3  | NULL  | NULL  | NULL  | NULL  |
| 400  | Street 4  | NULL  | NULL  | NULL  | NULL  |
+------+-----------+-------+-------+-------+-------+
```

Operatori su set

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

Gli operatori di set vengono utilizzati per combinare due relazioni di input in una sola. OpenSearch SQL supporta tre tipi di operatori di set:

- `EXCEPT` o `MINUS`

- INTERSECT
- UNION

Le relazioni di input devono avere lo stesso numero di colonne e tipi di dati compatibili per le rispettive colonne.

TRANNE

EXCEPTe EXCEPT ALL restituisce le righe che si trovano in una relazione ma non nell'altra. EXCEPT(in alternativa,EXCEPT DISTINCT) accetta solo righe distinte senza rimuovere i duplicati dalle righe dei risultati. EXCEPT ALL Nota che MINUS è un alias per. EXCEPT

Sintassi

```
 [ ( ] relation [ ) ] EXCEPT | MINUS [ ALL | DISTINCT ] [ ( ] relation [ ) ]
```

Esempi

```
-- Use table1 and table2 tables to demonstrate set operators in this page.
SELECT * FROM table1;
+---+
|  c|
+---+
|  3|
|  1|
|  2|
|  2|
|  3|
|  4|
+---+
SELECT * FROM table2;
+---+
|  c|
+---+
|  5|
|  1|
|  2|
|  2|
+---+
SELECT c FROM table1 EXCEPT SELECT c FROM table2;
+---+
|  c|
```

```
+---+
|  3|
|  4|
+---+
SELECT c FROM table1 MINUS SELECT c FROM table2;
+---+
|  c|
+---+
|  3|
|  4|
+---+
SELECT c FROM table1 EXCEPT ALL (SELECT c FROM table2);
+---+
|  c|
+---+
|  3|
|  3|
|  4|
+---+
SELECT c FROM table1 MINUS ALL (SELECT c FROM table2);
+---+
|  c|
+---+
|  3|
|  3|
|  4|
+---+
```

INTERSECARE

`INTERSECT`e `INTERSECT ALL` restituisce le righe che si trovano in entrambe le relazioni. `INTERSECT`(in alternativa,`INTERSECT DISTINCT`) accetta solo righe distinte senza rimuovere i duplicati dalle righe dei risultati. `INTERSECT ALL`

Sintassi

```
 [ ( ] relation [ ) ] INTERSECT [ ALL | DISTINCT ] [ ( ] relation [ ) ]
```

Esempi

```
(SELECT c FROM table1) INTERSECT (SELECT c FROM table2);
+---+
|  c|
```

```
+---+
|  1|
|  2|
+---+
(SELECT c FROM table1) INTERSECT DISTINCT (SELECT c FROM table2);
+---+
|  c|
+---+
|  1|
|  2|
+---+
(SELECT c FROM table1) INTERSECT ALL (SELECT c FROM table2);
+---+
|  c|
+---+
|  1|
|  2|
|  2|
+---+
```

UNIONE

`UNION`e `UNION ALL` restituisce le righe che si trovano in entrambe le relazioni. `UNION`(in alternativa,`UNION DISTINCT`) accetta solo righe distinte senza rimuovere i duplicati dalle righe dei risultati. `UNION ALL`

Sintassi

```
 [ ( ] relation [ ) ] UNION [ ALL | DISTINCT ] [ ( ] relation [ ) ]
```

Esempi

```
(SELECT c FROM table1) UNION (SELECT c FROM table2);
+---+
|  c|
+---+
|  1|
|  3|
|  5|
|  4|
|  2|
+---+
(SELECT c FROM table1) UNION DISTINCT (SELECT c FROM table2);
```

```
+---+
|  c|
+---+
|  1|
|  3|
|  5|
|  4|
|  2|
+---+
SELECT c FROM table1 UNION ALL (SELECT c FROM table2);
+---+
|  c|
+---+
|  3|
|  1|
|  2|
|  2|
|  3|
|  4|
|  5|
|  1|
|  2|
|  2|
+---+
```

Clausola SORT BY

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

La `SORT BY` clausola viene utilizzata per restituire le righe dei risultati ordinate all'interno di ciascuna partizione nell'ordine specificato dall'utente. Quando c'è più di una partizione `SORT BY` può restituire un risultato parzialmente ordinato. Questo è diverso dalla `ORDER BY` clausola che garantisce un ordine totale dell'output.

Sintassi

```
SORT BY { expression [ sort_direction | nulls_sort_order ] [ , ... ] }
```

Parametri

- ORDINA PER

  Specifica un elenco di espressioni separate da virgole insieme ai parametri opzionali sort_direction e nulls_sort_order che vengono utilizzati per ordinare le righe all'interno di ciascuna partizione.

- sort_direction

  Facoltativamente specifica se ordinare le righe in ordine crescente o decrescente.

  I valori validi per la direzione di ordinamento sono per l'ordine crescente e `ASC` per il decrescente. `DESC`

  Se la direzione di ordinamento non è specificata in modo esplicito, per impostazione predefinita le righe vengono ordinate in modo crescente.

  Sintassi: `[ ASC | DESC ]`

- nulls_sort_order

  Specifica facoltativamente se i valori NULL vengono restituiti prima/dopo valori non NULL.

  Se non `null_sort_order` è specificato, NULLs ordina prima se il criterio di ordinamento è `ASC` e NULLS ordina per ultimo se il criterio di ordinamento è. `DESC`

  1. Se `NULLS FIRST` viene specificato, i valori NULL vengono restituiti per primi indipendentemente dal tipo di ordinamento.

  2. Se `NULLS LAST` viene specificato, i valori NULL vengono restituiti per ultimi indipendentemente dal tipo di ordinamento.

  Sintassi: `[ NULLS { FIRST | LAST } ]`

Esempi

```
CREATE TABLE person (zip_code INT, name STRING, age INT);
INSERT INTO person VALUES
(94588, 'Shirley Rodriguez', 50),
(94588, 'Juan Li', 18),
(94588, 'Anil K', 27),
(94588, 'John D', NULL),
(94511, 'David K', 42),
```

```
(94511, 'Aryan B.', 18),
(94511, 'Lalit B.', NULL);
-- Sort rows by `name` within each partition in ascending manner
SELECT name, age, zip_code FROM person SORT BY name;
+------------------+----+--------+
|              name| age|zip_code|
+------------------+----+--------+
|            Anil K|  27|   94588|
|           Juan Li|  18|   94588|
|            John D|null|   94588|
| Shirley Rodriguez|  50|   94588|
|          Aryan B.|  18|   94511|
|           David K|  42|   94511|
|          Lalit B.|null|   94511|
+------------------+----+--------+
-- Sort rows within each partition using column position.
SELECT name, age, zip_code FROM person SORT BY 1;
+----------------+----+----------+
|              name| age|zip_code|
+------------------+----+--------+
|            Anil K|  27|   94588|
|           Juan Li|  18|   94588|
|            John D|null|   94588|
| Shirley Rodriguez|  50|   94588|
|          Aryan B.|  18|   94511|
|           David K|  42|   94511|
|          Lalit B.|null|   94511|
+------------------+----+--------+

-- Sort rows within partition in ascending manner keeping null values to be last.
SELECT age, name, zip_code FROM person SORT BY age NULLS LAST;
+----+-----------------+--------+
| age|             name|zip_code|
+----+-----------------+--------+
|  18|           Juan Li|   94588|
|  27|            Anil K|   94588|
|  50| Shirley Rodriguez|   94588|
|null|            John D|   94588|
|  18|          Aryan B.|   94511|
|  42|           David K|   94511|
|null|          Lalit B.|   94511|
+--------------+--------+--------+
```

```
-- Sort rows by age within each partition in descending manner, which defaults to NULL
 LAST.
SELECT age, name, zip_code FROM person SORT BY age DESC;
+----+-----------------+--------+
| age|             name|zip_code|
+----+-----------------+--------+
|  50|          Shirley Rodriguez|   94588|
|  27|            Anil K|   94588|
|  18|           Juan Li|   94588|
|null|            John D|   94588|
|  42|           David K|   94511|
|  18|          Aryan B.|   94511|
|null|          Lalit B.|   94511|
+----+-----------------+--------+

-- Sort rows by age within each partition in descending manner keeping null values to
 be first.
SELECT age, name, zip_code FROM person SORT BY age DESC NULLS FIRST;
+----+-----------------+--------+
| age|             name|zip_code|
+----+-----------------+--------+
|null|            John D|   94588|
|  50| Shirley Rodriguez|   94588|
|  27|            Anil K|   94588|
|  18|           Juan Li|   94588|
|null|          Lalit B.|   94511|
|  42|           David K|   94511|
|  18|          Aryan B.|   94511|
+----+-----------------+--------+

-- Sort rows within each partition based on more than one column with each column
 having
-- different sort direction.
SELECT name, age, zip_code FROM person
SORT BY name ASC, age DESC;
+-----------------+----+--------+
|             name| age|zip_code|
+-----------------+----+--------+
|           Anil K|  27|   94588|
|          Juan Li|  18|   94588|
|           John D|null|   94588|
| Shirley Rodriguez|  50|   94588|
|         Aryan B.|  18|   94511|
|          David K|  42|   94511|
```

```
|          Lalit B.|null|   94511|
+------------------+----+--------+
```

UNPIVOT

> **Note**
>
> Per vedere quali integrazioni di origini AWS dati supportano questo comando SQL, vedi. the section called “Comandi SQL supportati”

La `UNPIVOT` clausola trasforma più colonne in più righe utilizzate nella clausola. `SELECT` La `UNPIVOT` clausola può essere specificata dopo il nome della tabella o della sottoquery.

Sintassi

```
UNPIVOT [ { INCLUDE | EXCLUDE } NULLS ] (
    { single_value_column_unpivot | multi_value_column_unpivot }
) [[AS] alias]

single_value_column_unpivot:
    values_column
    FOR name_column
    IN (unpivot_column [[AS] alias] [, ...])

multi_value_column_unpivot:
    (values_column [, ...])
    FOR name_column
    IN ((unpivot_column [, ...]) [[AS] alias] [, ...])
```

Parametri

- unpivot_column

  Contiene le colonne nella `FROM` clausola, che specifica le colonne su cui vogliamo eseguire l'unpivot.
- nome_colonna

  Il nome della colonna che contiene i nomi delle colonne non ruotate.
- valori_colonna

Il nome della colonna che contiene i valori delle colonne non pivotate.

Esempi

```
CREATE TABLE sales_quarterly (year INT, q1 INT, q2 INT, q3 INT, q4 INT);
INSERT INTO sales_quarterly VALUES
(2020, null, 1000, 2000, 2500),
(2021, 2250, 3200, 4200, 5900),
(2022, 4200, 3100, null, null);
-- column names are used as unpivot columns
SELECT * FROM sales_quarterly
UNPIVOT (
sales FOR quarter IN (q1, q2, q3, q4)
);
+------+---------+-------+
| year | quarter | sales |
+------+---------+-------+
| 2020 | q2      | 1000  |
| 2020 | q3      | 2000  |
| 2020 | q4      | 2500  |
| 2021 | q1      | 2250  |
| 2021 | q2      | 3200  |
| 2021 | q3      | 4200  |
| 2021 | q4      | 5900  |
| 2022 | q1      | 4200  |
| 2022 | q2      | 3100  |
+------+---------+-------+
-- NULL values are excluded by default, they can be included
-- unpivot columns can be alias
-- unpivot result can be referenced via its alias
SELECT up.* FROM sales_quarterly
UNPIVOT INCLUDE NULLS (
sales FOR quarter IN (q1 AS Q1, q2 AS Q2, q3 AS Q3, q4 AS Q4)
) AS up;
+------+---------+-------+
| year | quarter | sales |
+------+---------+-------+
| 2020 | Q1      | NULL  |
| 2020 | Q2      | 1000  |
| 2020 | Q3      | 2000  |
| 2020 | Q4      | 2500  |
| 2021 | Q1      | 2250  |
```

```
| 2021 | Q2      | 3200  |
| 2021 | Q3      | 4200  |
| 2021 | Q4      | 5900  |
| 2022 | Q1      | 4200  |
| 2022 | Q2      | 3100  |
| 2022 | Q3      | NULL  |
| 2022 | Q4      | NULL  |
+------+---------+-------+
-- multiple value columns can be unpivoted per row
SELECT * FROM sales_quarterly
UNPIVOT EXCLUDE NULLS (
(first_quarter, second_quarter)
FOR half_of_the_year IN (
(q1, q2) AS H1,
(q3, q4) AS H2
)
);
+------+-----------------+---------------+----------------+
|  id  | half_of_the_year | first_quarter | second_quarter |
+------+-----------------+---------------+----------------+
| 2020 | H1              | NULL          | 1000           |
| 2020 | H2              | 2000          | 2500           |
| 2021 | H1              | 2250          | 3200           |
| 2021 | H2              | 4200          | 5900           |
| 2022 | H1              | 4200          | 3100           |
+------+-----------------+---------------+----------------+
```

Comandi PPL supportati

Le tabelle seguenti mostrano quali comandi PPL sono supportati da OpenSearch Dashboards per interrogare CloudWatch Logs, Amazon S3 o Security Lake e quali comandi supporta Logs Insights. CloudWatch CloudWatch Logs Insights utilizza la stessa sintassi PPL di OpenSearch Dashboards per interrogare CloudWatch Logs e le tabelle si riferiscono a entrambi come Logs. CloudWatch

> **Note**
>
> Quando si analizzano i dati al di fuori del OpenSearch servizio, i comandi potrebbero essere eseguiti in modo diverso rispetto agli indici. OpenSearch

Argomenti

- Comandi
- Funzioni
- Informazioni aggiuntive per gli utenti di CloudWatch Logs Insights che utilizzano PPL OpenSearch

Comandi

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “campi”	Visualizza una serie di campi che necessitano di proiezione.	Supportato	Supportato	Supportato	`fields field1, field2`
the section called “dove”	Filtra i dati in base alle condizioni specificate.	Supportato	Supportato	Supportato	`where field1="success" | where field2 != "i-023fe0a90929d8822" | fields field3, col4, col5, col6 | head 1000`
the section called “stats”	Esegue aggregazioni e calcoli.	Supportato	Supportato	Supportato	`stats count(), count(`field1`),`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
					`min(`field1`), max(`field1`), avg(`field1`) by field2 | head 1000`
the section called “parse”	Estrae un pattern di espressione regolare (regex) da una stringa e visualizza il pattern estratto. Il pattern estratto può essere ulteriormente utilizzato per creare nuovi campi o filtrare i dati.	Supportato	Supportato	Supportato	`parse `field1` ".*/(?<field2>[^/]+$)" | where field2 = "requestId" | fields field2, `field2` | head 1000`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “modelli”	Estrae i modelli di registro da un campo di testo e aggiunge i risultati al risultato della ricerca. Il raggruppamento dei log in base ai relativi schemi semplifica l'aggregazione delle statistiche provenienti da grandi volumi di dati di registro per l'analisi e la risoluzione dei problemi.	Non supportato	Supportato	Supportato	`patterns new_field='no_numbers' pattern='[0-9]' message | fields message, no_numbers`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “sort”	Ordina i risultati visualizzati in base al nome di un campo. Usa sort - FieldName per ordinare in ordine decrescente.	Supportato	Supportato	Supportato	`stats count(), count(`field1`), min(`field1`) as field1Alias, max(`field1`), avg(`field1`) by field2 | sort - field1Alias | head 1000`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “eval”	Modifica o elabora il valore di un campo e lo memorizza in un campo diverso. Ciò è utile per modificare matematic amente una colonna, applicare funzioni di stringa a una colonna o applicare funzioni di data a una colonna.	Supportato	Supportato	Supportato	`eval field2 = `field1` * 2 | fields field1, field2 | head 20`
the section called “rinominare”	Rinomina uno o più campi nei risultati della ricerca.	Supportato	Supportato	Supportato	`rename field2 as field1 | fields field1`
the section called “head”	Limita i risultati dell'inte rrogazione visualizzati alle prime N righe.	Supportato	Supportato	Supportato	`fields `@message` | head 20`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “grok”	Analizza un campo di testo con un pattern grok basato su un'espressione regolare e aggiunge i risultati al risultato della ricerca.	Supportato	Supportato	Supportato	`grok email '.+@%{HOSTNAME:host}' | fields email`
the section called “top”	Trova i valori più frequenti per un campo.	Supportato	Supportato	Supportato	`top 2 Field1 by Field2`
the section called “dedup”	Rimuove le voci duplicate in base ai campi specificati.	Supportato	Supportato	Supportato	`dedup field1 | fields field1, field2, field3`
the section called “join”	Unisce due set di dati.	Supportato	Supportato	Supportato	`source=customer | join ON c_custkey = o_custkey orders | head 10`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “cercare”	Arricchisce i dati di ricerca aggiungendo o sostituendo i dati da un indice di ricerca (tabella delle dimensioni). È possibile estendere i campi di un indice con i valori di una tabella delle dimensioni, aggiungere o sostituire valori quando viene soddisfatta la condizione di ricerca	Non supportato	Supportato	Supportato	`where orderType = 'Cancelled' | lookup account_list, mkt_id AS mkt_code replace amount, account_name as name | stats count(mkt_code), avg(amount) by name`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “subquery”	Esegue interrogazioni complesse e annidate all'interno delle istruzioni PPL (Piped Processing Language).	Supportato	Supportato	Supportato	`where id in [ subquery source=users | where user in [ subquery source=actions | where action="login" | fields user ] | fields uid ]`
the section called “rari”	Trova i valori meno frequenti di tutti i campi nell'elenco dei campi.	Supportato	Supportato	Supportato	`rare Field1 by Field2`
the section called “linea di tendenza”	Calcola le medie mobili dei campi.	Supportato	Supportato	Supportato	`trendline sma(2, field1) as field1Alias`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “eventstats”	Arricchisce i dati degli eventi con statistiche riassuntive calcolate. Analizza i campi specifici all'interno degli eventi, calcola varie misure statistiche e quindi aggiunge questi risultati a ciascun evento originale come nuovi campi.	Supportato (eccetto) `count()`	Supportato	Supportato	`eventstats sum(field1) by field2`
the section called “flatten”	Appiattisce un campo. Il campo deve essere di questo tipo: `struct<?,?> or array<struct<?,?>>`	Supportato	Supportato	Supportato	`source=table | flatten field1`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “riepilogo del campo”	Calcola le statistiche di base per ogni campo (conteggio, conteggio distinto, min, max, avg, stddev e mean).	Supportato (un campo per query)	Supportato	Supportato	`where field1 != 200 | fieldsummary includefields=field1 nulls=true`
the section called “fillnull”	Riempie i campi nulli con il valore fornito. Può essere utilizzato in uno o più campi.	Supportato	Supportato	Supportato	`fields field1 | eval field2=field1 | fillnull value=0 field1`
the section called “Espansione”	Suddivide un campo contenente più valori in righe separate, creando una nuova riga per ogni valore nel campo specificato.	Supportato	Supportato	Supportato	`expand employee | stats max(salary) as max by state, company`

Comando PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “describe”	Ottiene informazioni dettagliate sulla struttura e i metadati di tabelle, schemi e cataloghi	Non supportato	Supportato	Supportato	`describe schema.table`

Funzioni

Funzione PPL	Description	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “Stringa” (`CONCAT`, `CONCAT_WS` , `LENGTH`, `LOWER`, `LTRIM`, `POSITION`, `REVERSE`, `RIGHT`, `RTRIM`, `SUBSTRING` , `TRIM`, `UPPER`)	Funzioni integrate in PPL in grado di manipolare e trasformare stringhe e dati di testo all'interno di query PPL. Ad	Supportato	Supportato	Supportato	`eval col1Len = LENGTH(col1) | fields col1Len`

Funzione PPL	Descripti on	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
	esempio, convertire maiuscole e minuscole , combinare stringhe, estrarre parti e pulire il testo.				

Funzione PPL	Descripti on	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “Data e ora” (DAY, DAYOFMONTH , DAY_OF_MONTH ,DAYOFWEEK , DAY_OF_WEEK , DAYOFYEAR , DAY_OF_YEAR , DAYNAME, FROM_UNIXTIME , HOUR, HOUR_OF_DAY , LAST_DAY, LOCALTIMESTAMP , LOCALTIME , MAKE_DATE , MINUTE, MINUTE_OF_HOUR , MONTH, MONTHNAME , MONTH_OF_YEAR , NOW, QUARTER, SECOND, SECOND_OF_MINUTE , SUBDATE, SYSDATE, TIMESTAMP , UNIX_TIMESTAMP , WEEK,	Funzioni integrate per la gestione e la trasforma zione dei dati di data e ora nelle query PPL. Ad esempio, date_add, date_form at, datediff e current_d ate.	Supportato	Supportat o	Supportat o	`eval newDate = ADDDATE(DATE('2020-08-26'), 1) | fields newDate`

Funzione PPL	Descripti on	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
WEEKDAY, WEEK_OF_Y EAR , DATE_ADD, DATE_SUB, TIMESTAMP ADD , TIMESTAMP DIFF , UTC_TIMES TAMP , CURRENT_T IMEZONE)					
the section called “Condizione” (EXISTS, IF, IFNULL, ISNOTNULL , ISNULL, NULLIF)	Funzioni integrate che eseguono calcoli su più righe per produrre un unico valore riepiloga to. Ad esempio, sum, count, avg, max e min.	Supportato	Supportat o	Supportat o	`eval field2 = isnull(co l1) | fields field2, col1, field3`

Funzione PPL	Descripti on	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “Math” (ABS, ACOS, ASIN, ATAN, ATAN2, CEIL, CEILING, CONV, COS, COT, CRC32, DEGREES, E, EXP, FLOOR, LN, LOG, LOG2, LOG10, MOD, PI. POW, POWER, RADIANS, RAND, ROUND, SIGN, SIN, SQRT, CBRT)	Funzioni integrate per eseguire calcoli e trasforma zioni matematic he nelle query PPL. Ad esempio: abs (valore assoluto) , round (arrotond a i numeri), sqrt (radice quadrata) , pow (calcolo della potenza) e ceil (arrotond a al numero	Supportato	Supportat o	Supportat o	`eval field2 = ACOS(col1) | fields col1`

Funzione PPL	Descripti on	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
	intero più vicino).				
the section called “Espressioni” (Operatori aritmetici (+,,*)-, Operatori predicati (,) `>. < IN`)	Le funzioni integrate per le espressioni, in particolare le espressioni di valore, restituiscono un valore scalare. Le espressioni hanno tipi e forme diversi.	Supportato	Supportato	Supportato	`where age > (25 + 5) | fields age`

Funzione PPL	Descripti on	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “IP address (Indirizzo IP)” (CIDRMATCH)	Funzioni integrate per la gestione di indirizzi IP come CIDR.	Supportato	Supportato	Supportato	`where cidrmatch(ip, '********/24') | fields ip`
the section called “JSON” (ARRAY_LENGTH , ARRAY_LENGTH , JSON, JSON_ARRAY , JSON_EXTRACT , JSON_KEYS , JSON_OBJECT , JSON_VALID , TO_JSON_STRING)	Funzioni integrate per la gestione di JSON, tra cui array, estrazione e convalida.	Supportato	Supportato	Supportato	`eval` `json_extract('{"a":"b"}', '$.a')` = `json_extract('{"a":"b"}', '$a')`
the section called “Lambda” (EXISTS, FILTER, REDUCE, TRANSFORM)	Funzioni integrate per la gestione di JSON, tra cui array, estrazione e convalida.	Non supportata	Supportato	Supportato	`eval array = json_array(1, -1, 2), result = filter(array, x -> x > 0) | fields result`

Funzione PPL	Descripti on	CloudWatch Registri	Simple Storage Service (Amazon S3)	Lago di sicurezza	Esempio di comando
the section called “Crittografico” (MD5, SHA1, SHA2)	Funzioni integrate che consentono di generare impronte digitali uniche di dati, che possono essere utilizzate per la verifica, il confronto o come parte di protocolli di sicurezza più complessi.	Supportato	Supportato	Supportato	`eval `MD5('hello')` = MD5('hello') | fields `MD5('hello')``

Informazioni aggiuntive per gli utenti di CloudWatch Logs Insights che utilizzano PPL OpenSearch

Sebbene CloudWatch Logs Insights supporti la maggior parte dei comandi e delle funzioni OpenSearch PPL, alcuni comandi e funzioni non sono attualmente supportati. Ad esempio,

attualmente non supporta i comandi Lookup in PPL. A partire dal 2 giugno 2025, CloudWatch Logs Insights ora supporta le funzioni JOIN, subqueries, Flatten, Fillnull, Expand, Cidrmatch e JSON in PPL. Per un elenco completo dei comandi e delle funzioni di query supportati, consulta le colonne di Amazon CloudWatch Logs nelle tabelle precedenti.

Esempi di domande e quote

Quanto segue si applica sia agli utenti di CloudWatch Logs Insights che agli utenti che effettuano query sui dati. OpenSearch CloudWatch

Per informazioni sui limiti che si applicano all'interrogazione di CloudWatch Logs from OpenSearch Service, consulta CloudWatch Logs quotas nella Amazon CloudWatch Logs User Guide. I limiti riguardano il numero di gruppi di CloudWatch log su cui è possibile eseguire query, il numero massimo di query simultanee che è possibile eseguire, il tempo massimo di esecuzione delle query e il numero massimo di righe restituite nei risultati. I limiti sono gli stessi indipendentemente dal linguaggio utilizzato per interrogare i log (ovvero OpenSearch PPL, SQL e CloudWatch Logs Insights QL).

Comandi PPL

Argomenti

- comment
- comando di correlazione
- comando dedup
- comando describe
- comando eval
- comando eventstats
- comando expand
- comando explain
- comando fillnull
- comando fields
- comando flatten
- comando grok
- comando head
- comando join

- comando lookup
- comando parse
- comando patterns
- comando raro
- comando di ridenominazione
- comando di ricerca
- comando sort
- comando stats
- comando subquery
- comando top
- comando trendline
- comando where
- riepilogo del campo
- comando expand
- Funzioni PPL

comment

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

PPL supporta sia i commenti di riga che i commenti a blocchi. Il sistema non valuta il testo dei commenti.

Commenti di riga

I commenti alle righe iniziano con due barre//e terminano con una nuova riga.

Esempio:

```
os> source=accounts | top gender // finds most common gender of all the accounts
fetched rows / total rows = 2/2
+----------+
```

```
| gender   |
|----------|
| M        |
| F        |
+----------+
```

Blocca commenti

I commenti di blocco iniziano con una barra seguita da un asterisco\ * e terminano con un asterisco seguito da una barra */.

Esempio:

```
os> source=accounts | dedup 2 gender /* dedup the document with gender field keep 2
 duplication */ | fields account_number, gender
fetched rows / total rows = 3/3
+------------------+----------+
| account_number   | gender   |
|------------------+----------|
| 1                | M        |
| 6                | M        |
| 13               | F        |
+------------------+----------+
```

comando di correlazione

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

È possibile correlare diverse fonti di dati in base a dimensioni e intervalli di tempo comuni.

Questa correlazione è fondamentale quando si ha a che fare con grandi quantità di dati provenienti da vari settori verticali che condividono gli stessi periodi di tempo ma non sono formalmente sincronizzati.

Correlando queste diverse fonti di dati in base a intervalli di tempo e dimensioni simili, puoi arricchire i dati e scoprire informazioni preziose.

Esempio

Il dominio dell'osservabilità ha tre fonti di dati distinte:

- Log
- Metriche
- Tracce

Queste fonti di dati potrebbero condividere dimensioni comuni. Per passare da una fonte di dati all'altra, è necessario correlarle correttamente. Utilizzando le convenzioni di denominazione semantica, puoi identificare gli elementi condivisi tra log, tracce e metriche.

Esempio:

```
{
  "@timestamp": "2018-07-02T22:23:00.186Z",
  "aws": {
    "elb": {
      "backend": {
        "http": {
          "response": {
            "status_code": 500
          }
        },
        "ip": "********",
        "port": "80"
      },
      ...
     "target_port": [
        "10.0.0.1:80"
      ],
      "target_status_code": [
        "500"
      ],
      "traceId": "Root=1-58337262-36d228ad5d99923122bbe354",
      "type": "http"
    }
  },
  "cloud": {
    "provider": "aws"
  },
  "http": {
    "request": {
    ...
```

```
  },
  "communication": {
    "source": {
      "address": "**************",
      "ip": "**************",
      "port": 2817
    }
  },
  "traceId": "Root=1-58337262-36d228ad5d99923122bbe354"
}
```

Questo esempio mostra un log AWS ELB proveniente da un servizio che risiede su. AWS Mostra una risposta HTTP di backend con un codice di stato 500, che indica un errore. Ciò potrebbe attivare un avviso o far parte del normale processo di monitoraggio. Il passo successivo è raccogliere dati pertinenti su questo evento per un'indagine approfondita.

Anche se potreste essere tentati di interrogare tutti i dati relativi al periodo di tempo, questo approccio può rivelarsi faticoso. Potresti finire con troppe informazioni e dedicare più tempo a filtrare i dati irrilevanti che a identificare la causa principale.

Puoi invece utilizzare un approccio più mirato correlando i dati provenienti da fonti diverse. Puoi utilizzare queste dimensioni per la correlazione:

- IP - `"ip": "10.0.0.1" | "ip": "**************"`
- Porta - `"port": 2817 | "target_port": "10.0.0.1:80"`

Supponendo che tu abbia accesso a tracce e indici metrici aggiuntivi e che tu abbia familiarità con la struttura dello schema, puoi creare una query di correlazione più precisa.

Ecco un esempio di documento con indice di traccia contenente informazioni HTTP che potresti voler correlare:

```
{
  "traceId": "c1d985bd02e1dbb85b444011f19a1ecc",
  "spanId": "55a698828fe06a42",
  "traceState": [],
  "parentSpanId": "",
  "name": "mysql",
  "kind": "CLIENT",
  "@timestamp": "2021-11-13T20:20:39+00:00",
  "events": [
```

```
    {
      "@timestamp": "2021-03-25T17:21:03+00:00",
       ...
    }
  ],
  "links": [
    {
      "traceId": "c1d985bd02e1dbb85b444011f19a1ecc",
      "spanId": "55a698828fe06a42w2",
      },
      "droppedAttributesCount": 0
    }
  ],
  "resource": {
    "service@name": "database",
    "telemetry@sdk@name": "opentelemetry",
    "host@hostname": "ip-172-31-10-8.us-west-2.compute.internal"
  },
  "status": {
    ...
  },
  "attributes": {
    "http": {
      "user_agent": {
        "original": "Mozilla/5.0"
      },
      "network": {
         ...
        }
      },
      "request": {
         ...
        }
      },
      "response": {
        "status_code": "200",
        "body": {
          "size": 500
        }
      },
      "client": {
        "server": {
          "socket": {
            "address": "***********",
```

```
            "domain": "example.com",
            "port": 80
          },
          "address": "***********",
          "port": 80
        },
        "resend_count": 0,
        "url": {
          "full": "http://example.com"
        }
      },
      "server": {
        "route": "/index",
        "address": "***********",
        "port": 8080,
        "socket": {
         ...
        },
        "client": {
         ...
         }
        },
        "url": {
         ...
        }
      }
    }
  }
}
```

In questo approccio puoi vedere gli `traceId` e gli http client/server `ip` che possono essere correlati ai log elb per comprendere meglio il comportamento e le condizioni del sistema.

Nuovo comando di interrogazione di correlazione

Ecco il nuovo comando che consentirebbe questo tipo di indagine:

```
source alb_logs, traces | where alb_logs.ip="10.0.0.1" AND
 alb_logs.cloud.provider="aws"|
correlate exact fields(traceId, ip) scope(@timestamp, 1D) mapping(alb_logs.ip =
 traces.attributes.http.server.address, alb_logs.traceId = traces.traceId )
```

Ecco cosa fa ogni parte del comando:

1. `source alb_logs, traces`- Seleziona le fonti di dati che desideri correlare.
2. `where ip="10.0.0.1" AND cloud.provider="aws"`- Questo restringe l'ambito della ricerca.
3. `correlate exact fields(traceId, ip)`- Ciò indica al sistema di correlare i dati in base alle corrispondenze esatte dei seguenti campi:
   - Il `ip` campo ha una condizione di filtro esplicita, quindi verrà utilizzato nella correlazione per tutte le fonti di dati.
   - Il `traceId` campo non ha un filtro esplicito, quindi corrisponderà agli stessi traceID in tutte le fonti di dati.

I nomi dei campi indicano il significato logico della funzione all'interno del comando di correlazione. La condizione di unione effettiva si basa sull'istruzione di mappatura fornita.

Il termine `exact` indica che le istruzioni di correlazione richiederanno che tutti i campi corrispondano per soddisfare l'istruzione di query.

Il termine `approximate` cercherà di corrispondere nel migliore dei casi e non rifiuterà le righe con corrispondenze parziali.

Risolvere diversi tipi di mappatura dei campi

Nei casi in cui lo stesso campo logico (ad esempio`ip`) ha nomi diversi tra le fonti di dati, è necessario fornire la mappatura esplicita dei campi del percorso. Per risolvere questo problema, puoi estendere le condizioni di correlazione in modo che corrispondano nomi di campo diversi con significati logici simili. Ecco come puoi farlo:

```
alb_logs.ip = traces.attributes.http.server.address, alb_logs.traceId = traces.traceId
```

Per ogni campo che partecipa al join di correlazione, è necessario fornire una dichiarazione di mappatura pertinente che includa tutte le tabelle che devono essere unite da questo comando di correlazione.

Esempio

In questo esempio, ci sono 2 fonti: `alb_logs, traces`

Ci sono 2 campi: `traceId, ip`

Esistono 2 istruzioni di mappatura: `alb_logs.ip = traces.attributes.http.server.address, alb_logs.traceId = traces.traceId`

Definizione dei tempi di correlazione

Per semplificare il lavoro svolto dal motore di esecuzione (driver), puoi aggiungere l'istruzione scope. Ciò indica in modo esplicito alla join query l'ora in cui deve spaziare per questa ricerca.

`scope(@timestamp, 1D)i`

In questo esempio, l'ambito di ricerca si concentra su base giornaliera, quindi le correlazioni che appaiono nello stesso giorno vengono raggruppate. Questo meccanismo di definizione dell'ambito semplifica e consente un migliore controllo sui risultati, abilitando una risoluzione incrementale della ricerca in base alle esigenze dell'utente.

Driver di supporto

Il nuovo comando di correlazione è in realtà un comando join «nascosto». Pertanto, solo i seguenti driver PPL supportano questo comando. In questi driver, il comando di correlazione verrà tradotto direttamente nel piano logico Catalyst Join appropriato.

Esempio

```
source alb_logs, traces, metrics | where ip="10.0.0.1" AND
cloud.provider="aws"| correlate exact on (ip, port) scope(@timestamp,
2018-07-02T22:23:00, 1 D)
```

Piano logico:

```
'Project [*]
+- 'Join Inner, ('ip && 'port)
   :- 'Filter (('ip === "10.0.0.1" & 'cloud.provider === "aws") &
 inTimeScope('@timestamp, "2018-07-02T22:23:00", "1 D"))
      +- 'UnresolvedRelation [alb_logs]
   +- 'Join Inner, ('ip & 'port)
      :- 'Filter (('ip === "10.0.0.1" & 'cloud.provider === "aws") &
 inTimeScope('@timestamp, "2018-07-02T22:23:00", "1 D"))
         +- 'UnresolvedRelation [traces]
      +- 'Filter (('ip === "10.0.0.1" & 'cloud.provider === "aws") &
 inTimeScope('@timestamp, "2018-07-02T22:23:00", "1 D"))
         +- 'UnresolvedRelation [metrics]
```

Il motore catalyst ottimizza questa query in base all'ordinamento di giunzione più efficiente.

comando dedup

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `dedup` comando per rimuovere documenti identici dai risultati della ricerca in base ai campi specificati.

Sintassi

Utilizzare la seguente sintassi:

```
dedup [int] <field-list> [keepempty=<bool>] [consecutive=<bool>]
```

`int`

- Opzionale.
- <int>Il `dedup` comando conserva più eventi per ogni combinazione quando viene specificata. Il numero per <int>deve essere maggiore di 0. Se non si specifica un numero, viene mantenuto solo il primo evento che si verifica. Tutti gli altri duplicati vengono rimossi dai risultati.
- Impostazione predefinita: 1

`keepempty`

- Opzionale.
- Se vero, mantiene i documenti in cui qualsiasi campo dell'elenco dei campi ha un valore NULL o è MANCANTE.
- Impostazione predefinita: false

`consecutive`

- Opzionale.
- Se vero, rimuove solo gli eventi con combinazioni di valori duplicate consecutive.

- Impostazione predefinita: false

`field-list`

- Obbligatorio.
- Un elenco di campi delimitato da virgole. È richiesto almeno un campo.

Esempio 1: deduplicazione per un campo

Questo esempio mostra come deduplicare i documenti utilizzando il campo gender.

Interrogazione PPL:

```
os> source=accounts | dedup gender | fields account_number, gender;
fetched rows / total rows = 2/2
+------------------+----------+
| account_number   | gender   |
|------------------+----------|
| 1                | M        |
| 13               | F        |
+------------------+----------+
```

Esempio 2: Conserva 2 documenti duplicati

L'esempio mostra come deduplicare i documenti con il campo relativo al genere, conservando due duplicati.

Interrogazione PPL:

```
os> source=accounts | dedup 2 gender | fields account_number, gender;
fetched rows / total rows = 3/3
+------------------+----------+
| account_number   | gender   |
|------------------+----------|
| 1                | M        |
| 6                | M        |
| 13               | F        |
+------------------+----------+
```

Esempio 3: Mantieni o ignora il campo vuoto per impostazione predefinita

L'esempio mostra come deduplicare il documento mantenendo il campo con valore nullo.

Interrogazione PPL:

```
os> source=accounts | dedup email keepempty=true | fields account_number, email;
fetched rows / total rows = 4/4
+------------------+-----------------------+
| account_number   | email                 |
+------------------+-----------------------+
| 1                | john_doe@example.com  |
| 6                | jane_doe@example.com  |
| 13               | null                  |
| 18               | juan_li@example.com   |
+------------------+-----------------------+
```

L'esempio mostra come deduplicare il documento ignorando il campo del valore vuoto.

Interrogazione PPL:

```
os> source=accounts | dedup email | fields account_number, email;
fetched rows / total rows = 3/3
+------------------+-----------------------+
| account_number   | email                 |
+------------------+-----------------------+
| 1                | john_doe@example.com  |
| 6                | jane_doe@example.com  |
| 18               | juan_li@example.com   |
+------------------+-----------------------+
```

Esempio 4: deduplicazione in documenti consecutivi

L'esempio mostra come eseguire la deduplicazione in documenti consecutivi.

Interrogazione PPL:

```
os> source=accounts | dedup gender consecutive=true | fields account_number, gender;
fetched rows / total rows = 3/3
+------------------+----------+
| account_number   | gender   |
+------------------+----------+
| 1                | M        |
| 13               | F        |
| 18               | M        |
```

```
+------------------+----------+
```

Esempi aggiuntivi

- `source = table | dedup a | fields a,b,c`
- `source = table | dedup a,b | fields a,b,c`
- `source = table | dedup a keepempty=true | fields a,b,c`
- `source = table | dedup a,b keepempty=true | fields a,b,c`
- `source = table | dedup 1 a | fields a,b,c`
- `source = table | dedup 1 a,b | fields a,b,c`
- `source = table | dedup 1 a keepempty=true | fields a,b,c`
- `source = table | dedup 1 a,b keepempty=true | fields a,b,c`
- `source = table | dedup 2 a | fields a,b,c`
- `source = table | dedup 2 a,b | fields a,b,c`
- `source = table | dedup 2 a keepempty=true | fields a,b,c`
- `source = table | dedup 2 a,b keepempty=true | fields a,b,c`
- `source = table | dedup 1 a consecutive=true| fields a,b,c`(la deduplicazione consecutiva non è supportata)

Limitazione

- Per `| dedup 2 a, b keepempty=false`

```
DataFrameDropColumns('_row_number_)
+- Filter ('_row_number_ <= 2) // allowed duplication = 2
   +- Window [row_number() windowspecdefinition('a, 'b, 'a ASC NULLS FIRST, 'b ASC
 NULLS FIRST, specifiedwindowframe(RowFrame, unboundedpreceding$(), currentrow$()))
 AS _row_number_], ['a, 'b], ['a ASC NULLS FIRST, 'b ASC NULLS FIRST]
       +- Filter (isnotnull('a) AND isnotnull('b)) // keepempty=false
          +- Project
             +- UnresolvedRelation
```

- Per `| dedup 2 a, b keepempty=true`

```
Union
:- DataFrameDropColumns('_row_number_)
:  +- Filter ('_row_number_ <= 2)
```

```
:     +- Window [row_number() windowspecdefinition('a, 'b, 'a ASC NULLS FIRST, 'b ASC
 NULLS FIRST, specifiedwindowframe(RowFrame, unboundedpreceding$(), currentrow$()))
 AS _row_number_], ['a, 'b], ['a ASC NULLS FIRST, 'b ASC NULLS FIRST]
:        +- Filter (isnotnull('a) AND isnotnull('b))
:           +- Project
:              +- UnresolvedRelation
+- Filter (isnull('a) OR isnull('b))
   +- Project
      +- UnresolvedRelation
```

comando describe

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `describe` comando per ottenere informazioni dettagliate sulla struttura e i metadati di tabelle, schemi e cataloghi. Di seguito sono riportati vari esempi e casi d'uso del comando. `describe`

Describe

- `describe table`Questo comando è uguale al comando `DESCRIBE EXTENDED table` SQL
- `describe schema.table`
- `describe schema.`table``
- `describe catalog.schema.table`
- `describe catalog.schema.`table``
- `describe `catalog`.`schema`.`table``

comando eval

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Il `eval` comando valuta l'espressione e aggiunge il risultato al risultato della ricerca.

Sintassi

Utilizzare la seguente sintassi:

```
eval <field>=<expression> ["," <field>=<expression> ]...
```

- `field`: Obbligatorio. Se il nome del campo non esiste, viene aggiunto un nuovo campo. Se il nome del campo esiste già, verrà sovrascritto.
- `expression`: Obbligatorio. Qualsiasi espressione supportata dal sistema.

Esempio 1: creare il nuovo campo

Questo esempio mostra come creare un nuovo `doubleAge` campo per ogni documento. Il nuovo `doubleAge` è il risultato della valutazione dell'età moltiplicato per 2.

Interrogazione PPL:

```
os> source=accounts | eval doubleAge = age * 2 | fields age, doubleAge ;
fetched rows / total rows = 4/4
+-------+-------------+
| age   | doubleAge   |
|-------+-------------|
| 32    | 64          |
| 36    | 72          |
| 28    | 56          |
| 33    | 66          |
+-------+-------------+
```

Esempio 2: sovrascrivi il campo esistente

Questo esempio mostra come sovrascrivere il campo età esistente con age plus 1.

Interrogazione PPL:

```
os> source=accounts | eval age = age + 1 | fields age ;
fetched rows / total rows = 4/4
+-------+
| age   |
|-------|
```

```
| 33    |
| 37    |
| 29    |
| 34    |
+-------+
```

Esempio 3: crea il nuovo campo con il campo definito in eval

Questo esempio mostra come creare un nuovo `ddAge` campo con un campo definito nel comando eval. Il nuovo campo `ddAge` è il risultato della valutazione di `doubleAge` moltiplicato per 2, dove `doubleAge` è definito nel comando eval.

Interrogazione PPL:

```
os> source=accounts | eval doubleAge = age * 2, ddAge = doubleAge * 2 | fields age,
 doubleAge, ddAge ;
fetched rows / total rows = 4/4
+-------+-------------+---------+
| age   | doubleAge   | ddAge   |
|-------+-------------+---------|
| 32    | 64          | 128     |
| 36    | 72          | 144     |
| 28    | 56          | 112     |
| 33    | 66          | 132     |
+-------+-------------+---------+
```

Ipotesi:`a`,`b`, `c` sono campi esistenti in `table`

Esempi aggiuntivi

- `source = table | eval f = 1 | fields a,b,c,f`
- `source = table | eval f = 1`(campi di output a, b, c, f)
- `source = table | eval n = now() | eval t = unix_timestamp(a) | fields n,t`
- `source = table | eval f = a | where f > 1 | sort f | fields a,b,c | head 5`
- `source = table | eval f = a * 2 | eval h = f * 2 | fields a,f,h`
- `source = table | eval f = a * 2, h = f * 2 | fields a,f,h`
- `source = table | eval f = a * 2, h = b | stats avg(f) by h`
- `source = table | eval f = ispresent(a)`

- `source = table | eval r = coalesce(a, b, c) | fields r`
- `source = table | eval e = isempty(a) | fields e`
- `source = table | eval e = isblank(a) | fields e`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one', a = 2, 'two', a = 3, 'three', a = 4, 'four', a = 5, 'five', a = 6, 'six', a = 7, 'se7en', a = 8, 'eight', a = 9, 'nine')`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one' else 'unknown')`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one' else concat(a, ' is an incorrect binary digit'))`
- `source = table | eval f = a in ('foo', 'bar') | fields f`
- `source = table | eval f = a not in ('foo', 'bar') | fields f`

Valuta con un esempio di caso:

```
source = table | eval e = eval status_category =
case(a >= 200 AND a < 300, 'Success',
a >= 300 AND a < 400, 'Redirection',
a >= 400 AND a < 500, 'Client Error',
a >= 500, 'Server Error'
else 'Unknown')
```

Valuta con un altro esempio di caso:

Ipotesi:a,b, c sono campi esistenti in `table`

Esempi aggiuntivi

- `source = table | eval f = 1 | fields a,b,c,f`
- `source = table | eval f = 1`(campi di output a, b, c, f)
- `source = table | eval n = now() | eval t = unix_timestamp(a) | fields n,t`
- `source = table | eval f = a | where f > 1 | sort f | fields a,b,c | head 5`
- `source = table | eval f = a * 2 | eval h = f * 2 | fields a,f,h`

- `source = table | eval f = a * 2, h = f * 2 | fields a,f,h`
- `source = table | eval f = a * 2, h = b | stats avg(f) by h`
- `source = table | eval f = ispresent(a)`
- `source = table | eval r = coalesce(a, b, c) | fields r`
- `source = table | eval e = isempty(a) | fields e`
- `source = table | eval e = isblank(a) | fields e`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one', a = 2, 'two', a = 3, 'three', a = 4, 'four', a = 5, 'five', a = 6, 'six', a = 7, 'se7en', a = 8, 'eight', a = 9, 'nine')`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one' else 'unknown')`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one' else concat(a, ' is an incorrect binary digit'))`
- `source = table | eval f = a in ('foo', 'bar') | fields f`
- `source = table | eval f = a not in ('foo', 'bar') | fields f`

Valuta con un esempio di caso:

```
source = table | eval e = eval status_category =
case(a >= 200 AND a < 300, 'Success',
a >= 300 AND a < 400, 'Redirection',
a >= 400 AND a < 500, 'Client Error',
a >= 500, 'Server Error'
else 'Unknown')
```

Valuta con un altro esempio di caso:

```
source = table |  where ispresent(a) |
eval status_category =
 case(a >= 200 AND a < 300, 'Success',
  a >= 300 AND a < 400, 'Redirection',
  a >= 400 AND a < 500, 'Client Error',
  a >= 500, 'Server Error'
  else 'Incorrect HTTP status code'
 )
```

```
| stats count() by status_category
```

Limitazioni

- La sovrascrittura dei campi esistenti non è supportata. Le query che tentano di farlo genereranno delle eccezioni con il messaggio «Il riferimento 'a' è ambiguo».

```
- `source = table | eval a = 10 | fields a,b,c`
- `source = table | eval a = a * 2 | stats avg(a)`
- `source = table | eval a = abs(a) | where a > 0`
- `source = table | eval a = signum(a) | where a < 0`
```

comando eventstats

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `eventstats` comando per arricchire i dati degli eventi con statistiche di riepilogo calcolate. Funziona analizzando campi specifici all'interno degli eventi, calcolando varie misure statistiche e quindi aggiungendo questi risultati come nuovi campi a ciascun evento originale.

Aspetti chiave di eventstats

1. Esegue calcoli sull'intero set di risultati o all'interno di gruppi definiti.
2. Gli eventi originali rimangono intatti, con nuovi campi aggiunti per contenere i risultati statistici.
3. Il comando è particolarmente utile per l'analisi comparativa, l'identificazione di valori anomali o l'aggiunta di un contesto aggiuntivo ai singoli eventi.

Differenza tra stats e eventstats

`eventstats`I comandi `stats` and vengono entrambi utilizzati per il calcolo delle statistiche, ma presentano alcune differenze fondamentali nel modo in cui funzionano e nei risultati che producono.

Formato di output

- `stats`: produce una tabella di riepilogo con solo le statistiche calcolate.

- `eventstats`: aggiunge le statistiche calcolate come nuovi campi agli eventi esistenti, preservando i dati originali.

Conservazione degli eventi

- `stats`: Riduce il set di risultati al solo riepilogo statistico, eliminando i singoli eventi.
- `eventstats`: conserva tutti gli eventi originali e aggiunge nuovi campi con le statistiche calcolate.

Casi d'uso

- `stats`: Ideale per creare report o dashboard di riepilogo. Spesso utilizzato come comando finale per riepilogare i risultati.
- `eventstats`: Utile quando è necessario arricchire gli eventi con un contesto statistico per ulteriori analisi o filtri. Può essere utilizzato durante la ricerca per aggiungere statistiche da utilizzare nei comandi successivi.

Sintassi

Utilizzare la seguente sintassi:

```
eventstats <aggregation>... [by-clause]
```

Aggregazione

- Obbligatorio.
- Una funzione di aggregazione.
- L'argomento dell'aggregazione deve essere un campo.

clausola secondaria

- Opzionale.
- Sintassi: `by [span-expression,] [field,]...`
- La clausola by può includere campi ed espressioni come funzioni scalari e funzioni di aggregazione. È inoltre possibile utilizzare la clausola span per suddividere un campo specifico in bucket di intervalli uguali. Il comando eventstats esegue quindi l'aggregazione in base a questi span bucket.

- Impostazione predefinita: se non si specifica una clausola by, il comando eventstats si aggrega sull'intero set di risultati.

espressione span-expression

- Facoltativo, al massimo uno.
- Sintassi: `span(field_expr, interval_expr)`
- L'unità dell'espressione dell'intervallo è l'unità naturale di default. Tuttavia, per i campi relativi ai tipi di data e ora, è necessario specificare l'unità nell'espressione dell'intervallo quando si utilizzano unità di data/ora.

  Ad esempio, per dividere il campo `age` in bucket per 10 anni, usa. `span(age, 10)` Per i campi basati sul tempo, puoi dividere un `timestamp` campo in intervalli orari utilizzando. `span(timestamp, 1h)`

Unità di tempo disponibili

Unità di intervallo di ampiezza
millisecondo (ms)
secondo/i
minuto (m, distinzione tra maiuscole e minuscole)
ora (h)
giorno (d)
settimana (e)
mese (M, distinzione tra maiuscole e minuscole)
trimestre (q)
anno (y)

Funzioni di aggregazione

COUNT

COUNTrestituisce un conteggio del numero di espr nelle righe recuperate da un'istruzione SELECT.

Per CloudWatch Logs utilizza le query, COUNT non è supportata.

Esempio:

```
os> source=accounts | eventstats count();
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------+--------------------
+------------+--------------------------+--------+-------+---------+
| account_number | balance  | firstname | lastname | age | gender | address
 | employer   | email                    | city   | state | count() |
+----------------+----------+-----------+----------+-----+--------+--------------------
+------------+--------------------------+--------+-------+---------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | *** Any Lane
 | AnyCorp    | janedoe@anycorp.com      | Brogan | IL    | 4       |
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
 | AnyCompany | marymajor@anycompany.com | Dante  | TN    | 4       |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | 789 Any Street
 | AnyOrg     |                          | Nogal  | VA    | 4       |
| 18             | 4180     | Juan      | Li       | 33  | M      | *** Example Court
 |            | juanli@exampleorg.com    | Orick  | MD    | 4       |
+----------------+----------+-----------+----------+-----+--------+--------------------
+------------+--------------------------+--------+-------+---------+
```

SUM

SUM(expr)restituisce la somma di expr.

Esempio:

```
os> source=accounts | eventstats sum(age) by gender;
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+-----------------------+------------+--------------------------+--------+-------
+--------------------+
| account_number | balance  | firstname | lastname | age | gender | address
   | employer   | email                    | city   | state | sum(age) by gender |
+----------------+----------+-----------+----------+-----+--------
+-----------------------+------------+--------------------------+--------+-------
+--------------------+
```

```
| 1              | 39225    | Jane      | Doe      | 32  | M      | 880 Any Lane
   | AnyCorp    | janedoe@anycorp.com      | Brogan | IL    | 101               |
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
   | AnyCompany | marymajor@anycompany.com | Dante  | TN    | 101               |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | 789 Any Street
   | AnyOrg     |                          | Nogal  | VA    | 28                |
| 18             | 4180     | Juan      | Li       | 33  | M      | 467 Example Court
   |            | juanli@exampleorg.com    | Orick  | MD    | 101               |
+----------------+----------+-----------+----------+-----+--------
+-----------------------+------------+--------------------------+--------+-------
+--------------------+
```

AVG

`AVG(expr)`restituisce il valore medio di expr.

Esempio:

```
os> source=accounts | eventstats avg(age) by gender;
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+-----------------------+------------+--------------------------+--------+-------
+--------------------+
| account_number | balance  | firstname | lastname | age | gender | address
   | employer    | email                    | city   | state | avg(age) by gender |
+----------------+----------+-----------+----------+-----+--------
+-----------------------+------------+--------------------------+--------+-------
+--------------------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | 880 Any Lane
   | AnyCorp     | janedoe@anycorp.com      | Brogan | IL    | 33.67             |
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
   | Any Company | marymajor@anycompany.com | Dante  | TN    | 33.67             |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | 789 Any Street
   | AnyOrg      |                          | Nogal  | VA    | 28.00             |
| 18             | 4180     | Juan      | Li       | 33  | M      | 467 Example Court
   |             | juanli@exampleorg.com    | Orick  | MD    | 33.67             |
+----------------+----------+-----------+----------+-----+--------
+-----------------------+-------------+--------------------------+--------+-------
+--------------------+
```

MAX

`MAX(expr)`Restituisce il valore massimo di expr.

Esempio

```
os> source=accounts | eventstats max(age);
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
| account_number | balance  | firstname | lastname | age | gender | address
   | employer    | email                     | city   | state | max(age)  |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | 880 Any Lane
   | AnyCorp     | janedoe@anycorp.com       | Brogan | IL    | 36        |
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
   | Any Company | marymajor@anycompany.com | Dante  | TN    | 36        |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | 789 Any Street
   | AnyOrg      |                           | Nogal  | VA    | 36        |
| 18             | 4180     | Juan      | Li       | 33  | M      | *** Example Court
   |             | juanli@exampleorg.com     | Orick  | MD    | 36        |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
```

MIN

`MIN(expr)`Restituisce il valore minimo di expr.

Esempio

```
os> source=accounts | eventstats min(age);
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
| account_number | balance  | firstname | lastname | age | gender | address
   | employer    | email                     | city   | state | min(age)  |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | 880 Any Lane
   | AnyCorp     | janedoe@anycorp.com       | Brogan | IL    | 28        |
```

```
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
   | Any Company | marymajor@anycompany.com | Dante  | TN    | 28        |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | *** Any Street
   | AnyOrg      |                          | Nogal  | VA    | 28        |
| 18             | 4180     | Juan      | Li       | 33  | M      | *** Example Court
   |             | juanli@exampleorg.com    | Orick  | MD    | 28        |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
```

STDDEV_SAMP

`STDDEV_SAMP(expr)`Restituisce la deviazione standard del campione di expr.

Esempio

```
os> source=accounts | eventstats stddev_samp(age);
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+------------------------+
| account_number | balance  | firstname | lastname | age | gender | address
   | employer    | email                    | city   | state | stddev_samp(age)       |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+------------------------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | *** Any Lane
   | AnyCorp     | janedoe@anycorp.com      | Brogan | IL    | 3.304037933599835      |
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
   | Any Company | marymajor@anycompany.com | Dante  | TN    | 3.304037933599835      |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | 789 Any Street
   | AnyOrg      |                          | Nogal  | VA    | 3.304037933599835      |
| 18             | 4180     | Juan      | Li       | 33  | M      | 467 Example Court
   |             | juanli@exampleorg.com    | Orick  | MD    | 3.304037933599835      |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+------------------------+
```

STDDEV_POP

`STDDEV_POP(expr)`Restituisce la deviazione standard della popolazione di expr.

Esempio

```
os> source=accounts | eventstats stddev_pop(age);
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+------------------------+
| account_number | balance  | firstname | lastname | age | gender | address
   | employer    | email                    | city   | state | stddev_pop(age)        |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+------------------------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | 880 Any Lane
   | AnyCorp     | janedoe@anycorp.com      | Brogan | IL    | 2.****************     |
| 6              | 5686     | Mary      | Major    | 36  | M      | *** Example Street
   | Any Company | marymajor@anycompany.com | Dante  | TN    | 2.****************     |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | *** Any Street
   | AnyOrg      |                          | Nogal  | VA    | 2.****************     |
| 18             | 4180     | Juan      | Li       | 33  | M      | *** Example Court
   |             | juanli@exampleorg.com    | Orick  | MD    | 2.****************     |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+------------------------+
```

PERCENTILE o PERCENTILE_APPROX

`PERCENTILE(expr, percent)`o `PERCENTILE_APPROX(expr, percent)` Restituisce il valore percentile approssimativo di expr alla percentuale specificata.

percentuale

- Il numero deve essere una costante compresa tra 0 e 100.

Esempio

```
os> source=accounts | eventstats percentile(age, 90) by gender;
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+--------------------------------+
| account_number | balance  | firstname | lastname | age | gender | address
    | employer    | email                    | city   | state | percentile(age, 90) by
 gender  |
```

```
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+--------------------------------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | *** Any Lane
   | AnyCorp     | janedoe@anycorp.com      | Brogan | IL    | 36
       |
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
   | Any Company | marymajor@anycompany.com | Dante  | TN    | 36
       |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | 789 Any Street
   | AnyOrg      |                          | Nogal  | VA    | 28
       |
| 18             | 4180     | Juan      | Li       | 33  | M      | *** Example Court
   |             | juanli@exampleorg.com    | Orick  | MD    | 36
       |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+--------------------------------+
```

Esempio 1: calcola la media, la somma e il conteggio di un campo per gruppo

L'esempio mostra come calcolare l'età media, la somma dell'età e il conteggio degli eventi di tutti gli account raggruppati per genere.

```
os> source=accounts | eventstats avg(age) as avg_age, sum(age) as sum_age, count() as
 count by gender;
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+-----------+-------+
| account_number | balance  | firstname | lastname | age | gender | address
    | employer    | email                    | city   | state | avg_age   | sum_age   |
 count |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+-----------+-------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | *** Any Lane
    | AnyCorp     | janedoe@anycorp.com      | Brogan | IL    | 33.666667 | 101       |
 3     |
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
    | Any Company | marymajor@anycompany.com | Dante  | TN    | 33.666667 | 101       |
 3     |
```

```
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | 789 Any Street
    | AnyOrg      |                          | Nogal  | VA    | 28.000000 | 28        |
 1     |
| 18             | 4180     | Juan      | Li       | 33  | M      | *** Example Court
    |             | juanli@exampleorg.com    | Orick  | MD    | 33.666667 | 101       |
 3     |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+-----------+--------+
```

Esempio 2: calcola il conteggio per un intervallo

L'esempio ottiene il conteggio dell'età per un intervallo di 10 anni.

```
os> source=accounts | eventstats count(age) by span(age, 10) as age_span
fetched rows / total rows = 4/4
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
| account_number | balance  | firstname | lastname | age | gender | address
   | employer    | email                    | city   | state | age_span |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
| 1              | 39225    | Jane      | Doe      | 32  | M      | *** Any Lane
   | AnyCorp     | janedoe@anycorp.com      | Brogan | IL    | 3         |
| 6              | 5686     | Mary      | Major    | 36  | M      | 671 Example Street
   | Any Company | marymajor@anycompany.com | Dante  | TN    | 3         |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | 789 Any Street
   | AnyOrg      |                          | Nogal  | VA    | 1         |
| 18             | 4180     | Juan      | Li       | 33  | M      | *** Example Court
   |             | juanli@exampleorg.com    | Orick  | MD    | 3         |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+
```

Esempio 3: calcola il conteggio in base al sesso e all'intervallo

L'esempio ottiene il conteggio dell'età per intervallo di 5 anni e il gruppo per sesso.

```
os> source=accounts | eventstats count() as cnt by span(age, 5) as age_span, gender
fetched rows / total rows = 4/4
```

```
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+---------------------------+--------+-------
+-----+
| account_number | balance  | firstname | lastname | age | gender | address
   | employer    | email                     | city   | state | cnt |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+---------------------------+--------+-------
+-----+
| 1              | 39225    | Jane      | Doe      | 32  | M      | *** Any Lane
   | AnyCorp     | janedoe@anycorp.com       | Brogan | IL    | 2   |
| 6              | 5686     | Mary      | Majo     | 36  | M      | 671 Example Street
   | Any Company | hattiebond@anycompany.com | Dante  | TN    | 1   |
| 13             | 32838    | Nikki     | Wolf     | 28  | F      | *** Any Street
   | AnyOrg      |                           | Nogal  | VA    | 1   |
| 18             | 4180     | Juan      | Li       | 33  | M      | *** Example Court
   |             | juanli@exampleorg.com     | Orick  | MD    | 2   |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+---------------------------+--------+-------
+-----+
```

Utilizzo

- `source = table | eventstats avg(a)`
- `source = table | where a < 50 | eventstats avg(c)`
- `source = table | eventstats max(c) by b`
- `source = table | eventstats count(c) by b | head 5`
- `source = table | eventstats distinct_count(c)`
- `source = table | eventstats stddev_samp(c)`
- `source = table | eventstats stddev_pop(c)`
- `source = table | eventstats percentile(c, 90)`
- `source = table | eventstats percentile_approx(c, 99)`

Aggregazioni con span

- `source = table | eventstats count(a) by span(a, 10) as a_span`
- `source = table | eventstats sum(age) by span(age, 5) as age_span | head 2`

- `source = table | eventstats avg(age) by span(age, 20) as age_span, country | sort - age_span | head 2`

Aggregazioni con intervallo di tempo (funzione tumble windowing)

- `source = table | eventstats sum(productsAmount) by span(transactionDate, 1d) as age_date | sort age_date`
- `source = table | eventstats sum(productsAmount) by span(transactionDate, 1w) as age_date, productId`

Le aggregazioni sono raggruppate per più livelli

- `source = table | eventstats avg(age) as avg_state_age by country, state | eventstats avg(avg_state_age) as avg_country_age by country`
- `source = table | eventstats avg(age) as avg_city_age by country, state, city | eval new_avg_city_age = avg_city_age - 1 | eventstats avg(new_avg_city_age) as avg_state_age by country, state | where avg_state_age > 18 | eventstats avg(avg_state_age) as avg_adult_country_age by country`

comando expand

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `expand` comando per appiattire un campo di tipo:

- `Array<Any>`
- `Map<Any>`

Sintassi

Utilizzare la seguente sintassi:

```
expand <field> [As alias]
```

field

- Il campo da espandere (esplodere). Deve essere di un tipo supportato.

alias

- Opzionale. Il nome da utilizzare al posto del nome del campo originale.

Utilizzo

Il `expand` comando produce una riga per ogni elemento nell'array o nel campo della mappa specificato, dove:

- Gli elementi dell'array diventano righe singole.
- Le coppie chiave-valore della mappa sono suddivise in righe separate, con ogni valore chiave rappresentato come una riga.

- Quando viene fornito un alias, i valori esplosi vengono rappresentati sotto l'alias anziché nel nome del campo originale.
- Questo può essere usato in combinazione con altri comandi, come `statseval`, e `parse` per manipolare o estrarre i dati dopo l'espansione.

Esempi

- `source = table | expand employee | stats max(salary) as max by state, company`
- `source = table | expand employee as worker | stats max(salary) as max by state, company`
- `source = table | expand employee as worker | eval bonus = salary * 3 | fields worker, bonus`
- `source = table | expand employee | parse description '(?<email>.+@.+)' | fields employee, email`

- `source = table | eval array=json_array(1, 2, 3) | expand array as uid | fields name, occupation, uid`
- `source = table | expand multi_valueA as multiA | expand multi_valueB as multiB`

comando explain

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Il `explain` comando consente di comprendere i piani di esecuzione delle query, consentendovi di analizzare e ottimizzare le query per migliorare le prestazioni. Questa introduzione fornisce una panoramica concisa dello scopo del comando explain e della sua importanza nell'ottimizzazione delle query.

Comment

- `source=accounts | top gender // finds most common gender of all the accounts`(commento di riga)
- `source=accounts | dedup 2 gender /* dedup the document with gender field keep 2 duplication */ | fields account_number, gender`(commento del blocco)

Describe

- `describe table`Questo comando è uguale al comando `DESCRIBE EXTENDED table` SQL
- `describe schema.table`
- `describe schema.\`table\``
- `describe catalog.schema.table`
- `describe catalog.schema.\`table\``
- `describe \`catalog\`.\`schema\`.\`table\``

Spiegazione

- `explain simple | source = table | where a = 1 | fields a,b,c`
- `explain extended | source = table`
- `explain codegen | source = table | dedup a | fields a,b,c`
- `explain cost | source = table | sort a | fields a,b,c`
- `explain formatted | source = table | fields - a`
- `explain simple | describe table`

Campi

- `source = table`
- `source = table | fields a,b,c`
- `source = table | fields + a,b,c`
- `source = table | fields - b,c`
- `source = table | eval b1 = b | fields - b1,c`

Riepilogo del campo

- `source = t | fieldsummary includefields=status_code nulls=false`
- `source = t | fieldsummary includefields= id, status_code, request_path nulls=true`
- `source = t | where status_code != 200 | fieldsummary includefields= status_code nulls=true`

Campo annidato

- `source = catalog.schema.table1, catalog.schema.table2 | fields A.nested1, B.nested1`
- `source = catalog.table | where struct_col2.field1.subfield > 'valueA' | sort int_col | fields int_col, struct_col.field1.subfield, struct_col2.field1.subfield`

- `source = catalog.schema.table | where struct_col2.field1.subfield > 'valueA' | sort int_col | fields int_col, struct_col.field1.subfield, struct_col2.field1.subfield`

Filtri

- `source = table | where a = 1 | fields a,b,c`
- `source = table | where a >= 1 | fields a,b,c`
- `source = table | where a < 1 | fields a,b,c`
- `source = table | where b != 'test' | fields a,b,c`
- `source = table | where c = 'test' | fields a,b,c | head 3`
- `source = table | where ispresent(b)`
- `source = table | where isnull(coalesce(a, b)) | fields a,b,c | head 3`
- `source = table | where isempty(a)`
- `source = table | where isblank(a)`
- `source = table | where case(length(a) > 6, 'True' else 'False') = 'True'`
- `source = table | where a not in (1, 2, 3) | fields a,b,c`
- `source = table | where a between 1 and 4`- Nota: questo restituisce un >= 1 e un <= 4, cioè [1, 4]
- `source = table | where b not between '2024-09-10' and '2025-09-10'`- Nota: restituisce b >= '**********' e b <= '2025-09-10'
- `source = table | where cidrmatch(ip, '***********/24')`
- `source = table | where cidrmatch(ipv6, '2003:db8::/32')`
- `source = table | trendline sma(2, temperature) as temp_trend`

Interrogazioni relative all'IP

- `source = table | where cidrmatch(ip, '***************')`
- `source = table | where isV6 = false and isValid = true and cidrmatch(ipAddress, '***************')`
- `source = table | where isV6 = true | eval inRange = case(cidrmatch(ipAddress, '2003:***::/32'), 'in' else 'out') | fields ip, inRange`

Filtri complessi

```
source = table | eval status_category =
case(a >= 200 AND a < 300, 'Success',
    a >= 300 AND a < 400, 'Redirection',
    a >= 400 AND a < 500, 'Client Error',
    a >= 500, 'Server Error'
else 'Incorrect HTTP status code')
| where case(a >= 200 AND a < 300, 'Success',
    a >= 300 AND a < 400, 'Redirection',
    a >= 400 AND a < 500, 'Client Error',
    a >= 500, 'Server Error'
else 'Incorrect HTTP status code'
) = 'Incorrect HTTP status code'
```

```
source = table
| eval factor = case(a > 15, a - 14, isnull(b), a - 7, a < 3, a + 1 else 1)
| where case(factor = 2, 'even', factor = 4, 'even', factor = 6, 'even', factor = 8,
 'even' else 'odd') = 'even'
| stats count() by factor
```

Filtri con condizioni logiche

- `source = table | where c = 'test' AND a = 1 | fields a,b,c`
- `source = table | where c != 'test' OR a > 1 | fields a,b,c | head 1`
- `source = table | where c = 'test' NOT a > 1 | fields a,b,c`

Valutazione

Ipotesi:a,b, c sono campi esistenti in `table`

- `source = table | eval f = 1 | fields a,b,c,f`
- `source = table | eval f = 1`(campi di output a, b, c, f)
- `source = table | eval n = now() | eval t = unix_timestamp(a) | fields n,t`
- `source = table | eval f = a | where f > 1 | sort f | fields a,b,c | head 5`
- `source = table | eval f = a * 2 | eval h = f * 2 | fields a,f,h`

- `source = table | eval f = a * 2, h = f * 2 | fields a,f,h`
- `source = table | eval f = a * 2, h = b | stats avg(f) by h`
- `source = table | eval f = ispresent(a)`
- `source = table | eval r = coalesce(a, b, c) | fields r`
- `source = table | eval e = isempty(a) | fields e`
- `source = table | eval e = isblank(a) | fields e`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one', a = 2, 'two', a = 3, 'three', a = 4, 'four', a = 5, 'five', a = 6, 'six', a = 7, 'se7en', a = 8, 'eight', a = 9, 'nine')`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one' else 'unknown')`
- `source = table | eval f = case(a = 0, 'zero', a = 1, 'one' else concat(a, ' is an incorrect binary digit'))`
- `source = table | eval digest = md5(fieldName) | fields digest`
- `source = table | eval digest = sha1(fieldName) | fields digest`
- `source = table | eval digest = sha2(fieldName,256) | fields digest`
- `source = table | eval digest = sha2(fieldName,512) | fields digest`

comando fillnull

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Description

Utilizzate il `fillnull` comando per sostituire i valori nulli con un valore specificato in uno o più campi dei risultati della ricerca.

Sintassi

Utilizzare la seguente sintassi:

```
fillnull [with <null-replacement> in <nullable-field>["," <nullable-field>]] | [using
 <source-field> = <null-replacement> [","<source-field> = <null-replacement>]]
```

- null-replacement: obbligatorio. Il valore usato per sostituire i valori nulli.
- nullable-field: obbligatorio. Riferimento al campo. I valori nulli in questo campo verranno sostituiti con il valore specificato in null-replacement.

Esempio 1: Riempi un campo

L'esempio mostra come utilizzare fillnull su un singolo campo:

```
os> source=logs | fields status_code | eval input=status_code | fillnull with 0 in
 status_code;
| input | status_code |
|-------|-------------|
| 403   | 403         |
| 403   | 403         |
| NULL  | 0           |
| NULL  | 0           |
| 200   | 200         |
| 404   | 404         |
| 500   | 500         |
| NULL  | 0           |
| 500   | 500         |
| 404   | 404         |
| 200   | 200         |
| 500   | 500         |
| NULL  | 0           |
| NULL  | 0           |
| 404   | 404         |
```

Esempio 2: Fillnull applicato a più campi

L'esempio mostra fillnull applicato a più campi.

```
os> source=logs | fields request_path, timestamp | eval
 input_request_path=request_path, input_timestamp = timestamp | fillnull with '???' in
 request_path, timestamp;
| input_request_path | input_timestamp       | request_path | timestamp             |
|--------------------------------------------------------------------------------|
| /contact           | NULL                  | /contact     | ???                   |
```

```
| /home              | NULL                  | /home       | ???                   |
| /about             | 2023-10-01 10:30:00   | /about      | 2023-10-01 10:30:00   |
| /home              | 2023-10-01 10:15:00   | /home       | 2023-10-01 10:15:00   |
| NULL               | 2023-10-01 10:20:00   | ???         | 2023-10-01 10:20:00   |
| NULL               | 2023-10-01 11:05:00   | ???         | 2023-10-01 11:05:00   |
| /about             | NULL                  | /about      | ???                   |
| /home              | 2023-10-01 10:00:00   | /home       | 2023-10-01 10:00:00   |
| /contact           | NULL                  | /contact    | ???                   |
| NULL               | 2023-10-01 10:05:00   | ???         | 2023-10-01 10:05:00   |
| NULL               | 2023-10-01 10:50:00   | ???         | 2023-10-01 10:50:00   |
| /services          | NULL                  | /services   | ???                   |
| /home              | 2023-10-01 10:45:00   | /home       | 2023-10-01 10:45:00   |
| /services          | 2023-10-01 11:00:00   | /services   | 2023-10-01 11:00:00   |
| NULL               | 2023-10-01 10:35:00   | ???         | 2023-10-01 10:35:00   |
```

Esempio 3: Fillnull applicato a più campi con vari valori sostitutivi nulli.

L'esempio mostra fillnull con vari valori usati per sostituire i valori null.

- `/error`nel `request_path` campo
- `1970-01-01 00:00:00`in `timestamp` campo

```
os> source=logs | fields request_path, timestamp | eval
 input_request_path=request_path, input_timestamp = timestamp | fillnull using
 request_path = '/error', timestamp='1970-01-01 00:00:00';

| input_request_path | input_timestamp       | request_path | timestamp             |
|--------------------------------------------------------------------------------|
| /contact           | NULL                  | /contact    | 1970-01-01 00:00:00   |
| /home              | NULL                  | /home       | 1970-01-01 00:00:00   |
| /about             | 2023-10-01 10:30:00   | /about      | 2023-10-01 10:30:00   |
| /home              | 2023-10-01 10:15:00   | /home       | 2023-10-01 10:15:00   |
| NULL               | 2023-10-01 10:20:00   | /error      | 2023-10-01 10:20:00   |
| NULL               | 2023-10-01 11:05:00   | /error      | 2023-10-01 11:05:00   |
| /about             | NULL                  | /about      | 1970-01-01 00:00:00   |
| /home              | 2023-10-01 10:00:00   | /home       | 2023-10-01 10:00:00   |
| /contact           | NULL                  | /contact    | 1970-01-01 00:00:00   |
| NULL               | 2023-10-01 10:05:00   | /error      | 2023-10-01 10:05:00   |
| NULL               | 2023-10-01 10:50:00   | /error      | 2023-10-01 10:50:00   |
| /services          | NULL                  | /services   | 1970-01-01 00:00:00   |
| /home              | 2023-10-01 10:45:00   | /home       | 2023-10-01 10:45:00   |
| /services          | 2023-10-01 11:00:00   | /services   | 2023-10-01 11:00:00   |
```

```
| NULL                | 2023-10-01 10:35:00   | /error       | 2023-10-01 10:35:00    |
```

comando fields

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `fields` comando per mantenere o rimuovere campi dal risultato della ricerca.

Sintassi

Utilizzare la seguente sintassi:

```
field [+|-] <field-list>
```

- `index`: facoltativo.

  Se si utilizza il segno più (+), verranno conservati solo i campi specificati nell'elenco dei campi.

  Se si utilizza il segno meno (-), tutti i campi specificati nell'elenco dei campi verranno rimossi.

  Impostazione predefinita: +
- `field list`: Obbligatorio. Un elenco di campi delimitati da virgole da conservare o rimuovere.

Esempio 1: Seleziona i campi specificati dal risultato

Questo esempio mostra come recuperare `account_number` `lastname` i campi e dai risultati di ricerca. `firstname`

Interrogazione PPL:

```
os> source=accounts | fields account_number, firstname, lastname;
fetched rows / total rows = 4/4
+------------------+-------------+------------+
| account_number   | firstname   | lastname   |
|------------------+-------------+------------|
```

```
| 1                | Jane        | Doe         |
| 6                | John        | Doe         |
| 13               | Jorge       | Souza       |
| 18               | Juan        | Li          |
+------------------+-------------+-------------+
```

Esempio 2: rimuovere i campi specificati dal risultato

Questo esempio mostra come rimuovere il `account_number` campo dai risultati di ricerca.

Interrogazione PPL:

```
os> source=accounts | fields account_number, firstname, lastname | fields -
 account_number ;
fetched rows / total rows = 4/4
+-------------+------------+
| firstname   | lastname   |
|-------------+------------|
| Jane        | Doe        |
| John        | Doe        |
| Jorge       | Souza      |
| Juan        | Li         |
+-------------+------------+
```

Esempi aggiuntivi

- `source = table`
- `source = table | fields a,b,c`
- `source = table | fields + a,b,c`
- `source = table | fields - b,c`
- `source = table | eval b1 = b | fields - b1,c`

Esempio di campi annidati:

```
`source = catalog.schema.table1, catalog.schema.table2 | fields A.nested1, B.nested1`
`source = catalog.table | where struct_col2.field1.subfield > 'valueA' | sort int_col |
 fields  int_col, struct_col.field1.subfield, struct_col2.field1.subfield`
`source = catalog.schema.table | where struct_col2.field1.subfield > 'valueA' | sort
 int_col | fields  int_col, struct_col.field1.subfield, struct_col2.field1.subfield`
```

comando flatten

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il comando flatten per espandere i campi dei seguenti tipi:

- `struct<?,?>`
- `array<struct<?,?>>`

Sintassi

Utilizzare la seguente sintassi:

```
flatten <field>
```

- field: Il campo da appiattire. Il campo deve essere del tipo supportato.

Schema

nome_col	data_type
_ora	stringa
ponti	<length:bigint, name:string>matrice <struct>
città	stringa
colore	struttura<alt:bigint, lat:double, long:doub le>
paese	stringa

Dati

_tempo	ponti	città	colore	paese
2024-09-13H 12:00:00	[{801, Tower Bridge}, {928, Ponte di Londra}]	Londra	{35, 51,5074, -0,1278}	Inghilterra
2024-09-13H 12:00:00	[{232, Pont Neuf}, {160, Ponte Alessandro III}]	Parigi	{35, 48,8566, 2,3522}	Francia
2024-09-13H 12:00:00	[{48, Ponte di Rialto}, {11, Ponte dei Sospiri}]	Venezia	{2, 45.408, 12.315}	Italia
2024-09-13H 12:00:00	[{***, Ponte Carlo}, {343, Ponte della Legione}]	Praga	{200, 50.0755, 14.4378}	Repubblica Ceca
2024-09-13H 12:00:00	[{375, Ponte delle Catene}, {333, Ponte della Libertà}]	Budapest	{96, 47,4979, 19.0402}	Ungheria
1990-09-13H 12:00:00	NULL	Varsavia	NULL	Polonia

Esempio 1: appiattire la struttura

Questo esempio mostra come appiattire un campo struct.

Interrogazione PPL:

```
source=table | flatten coor
```

_ora	ponti	città	paese	alt	Lat	Long
2024-09-13H 12:00:00	[{801, Tower Bridge}, {928, Ponte di Londra}]	Londra	Inghilterra	35	51,5074	-0,1278
2024-09-13H 12:00:00	[{232, Pont Neuf}, {160, Ponte Alessandro III}]	Parigi	Francia	35	48,8566	2,3522
2024-09-13H 12:00:00	[{48, Ponte di Rialto}, {11, Ponte dei Sospiri}]	Venezia	Italia	2	45.408	12,3155
2024-09-13H 12:00:00	[{516, Ponte Carlo}, {343, Ponte	Praga	Repubblica Ceca	200	50.0755	14,4378

_ora	ponti	città	paese	alt	Lat	Long
	della Legione}]					
2024-09-13H 12:00:00	[{375, Ponte delle Catene}, {333, Ponte della Libertà}]	Budapest	Ungheria	96	47,4979	19,0402
1990-09-13H 12:00:00	NULL	Varsavia	Polonia	NULL	NULL	NULL

Esempio 2: flatten array

L'esempio mostra come appiattire una matrice di campi struct.

Interrogazione PPL:

```
source=table | flatten bridges
```

_ora	città	colore	paese	length	nome
2024-09-13H 12:00:00	Londra	{35, 51,5074, -0,1278}	Inghilterra	801	Tower Bridge
2024-09-13H 12:00:00	Londra	{35, 51,5074, -0,1278}	Inghilterra	928	Ponte di Londra
2024-09-13H 12:00:00	Parigi	{35, 48,8566, 2,3522}	Francia	232	Pont-Neuf

_ora	città	colore	paese	length	nome
2024-09-13H 12:00:00	Parigi	{35, 48,8566, 2,3522}	Francia	160	Ponte Alessandro III
2024-09-13H 12:00:00	Venezia	{2, 45.408, 12.315}	Italia	48	Ponte di Rialto
2024-09-13H 12:00:00	Venezia	{2, 45.408, 12.315}	Italia	11	Ponte dei Sospiri
2024-09-13H 12:00:00	Praga	{200, 50.0755, 14.4378}	Repubblica Ceca	516	Ponte Carlo
2024-09-13H 12:00:00	Praga	{200, 50.0755, 14.4378}	Repubblica Ceca	343	Ponte della Legione
2024-09-13H 12:00:00	Budapest	{96, 47,4979, 19.0402}	Ungheria	375	Ponte delle Catene
2024-09-13H 12:00:00	Budapest	{96, 47,4979, 19.0402}	Ungheria	333	Ponte della Libertà
1990-09-13H 12:00:00	Varsavia	NULL	Polonia	NULL	NULL

Esempio 3: appiattire l'array e la struttura

Questo esempio mostra come appiattire più campi.

Interrogazione PPL:

```
source=table | flatten bridges | flatten coor
```

_ora	città	paese	length	nome	alt	Lat	Long
2024-09-13H 12:00:00	Londra	Inghilterra	801	Tower Bridge	35	51,5074	-0,1278
2024-09-13H 12:00:00	Londra	Inghilterra	928	Ponte di Londra	35	51,5074	-0,1278
2024-09-13H 12:00:00	Parigi	Francia	232	Pont-Neuf	35	48,8566	2,3522
2024-09-13H 12:00:00	Parigi	Francia	160	Ponte Alessandro III	35	48,8566	2,3522
2024-09-13H 12:00:00	Venezia	Italia	48	Ponte di Rialto	2	45,408	12,3155
2024-09-13H 12:00:00	Venezia	Italia	11	Ponte dei Sospiri	2	45.408	12,3155
2024-09-13H 12:00:00	Praga	Repubblica Ceca	516	Ponte Carlo	200	50,0755	14,4378
2024-09-13H 12:00:00	Praga	Repubblica Ceca	343	Ponte della Legione	200	50.0755	14,4378
2024-09-13H 12:00:00	Budape	Ungheria	375	Ponte delle Catene	96	47,4979	19,0402

_ora	città	paese	length	nome	alt	Lat	Long
2024-09-13H 12:00:00	Budape	Ungheria	333	Ponte della Libertà	96	47,4979	19,0402
1990-09-13H 12:00:00	Varsavi	Polonia	NULL	NULL	NULL	NULL	NULL

comando grok

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Il `grok` comando analizza un campo di testo con uno schema grok e aggiunge i risultati al risultato della ricerca.

Sintassi

Utilizzare la seguente sintassi:

```
grok <field> <pattern>
```

field

- Obbligatorio.
- Il campo deve essere un campo di testo.

pattern

- Obbligatorio.
- Il modello grok usato per estrarre nuovi campi dal campo di testo specificato.
- Se esiste già un nuovo nome di campo, sostituirà il campo originale.

Pattern grok

Il pattern grok viene utilizzato per abbinare il campo di testo di ogni documento per estrarre nuovi campi.

Esempio 1: creare il nuovo campo

Questo esempio mostra come creare un nuovo campo `host` per ogni documento. `host`sarà il nome host dopo @ nel `email` campo. L'analisi di un campo nullo restituirà una stringa vuota.

```
os> source=accounts | grok email '.+@%{HOSTNAME:host}' | fields email, host ;
fetched rows / total rows = 4/4
+-------------------------+-------------+
| email                   | host        |
|-------------------------+-------------|
| jane_doe@example.com    | example.com |
| arnav_desai@example.net | example.net |
| null                    |             |
| juan_li@example.org     | example.org |
+-------------------------+-------------+
```

Esempio 2: sovrascrivi il campo esistente

Questo esempio mostra come sostituire il `address` campo esistente con il numero civico rimosso.

```
os> source=accounts | grok address '%{NUMBER} %{GREEDYDATA:address}' | fields address ;
fetched rows / total rows = 4/4
+------------------+
| address          |
|------------------|
| Example Lane     |
| Any Street       |
| Main Street      |
| Example Court    |
+------------------+
```

Esempio 3: Usare grok per analizzare i log

Questo esempio mostra come usare grok per analizzare i log non elaborati.

```
os> source=apache | grok message '%{COMMONAPACHELOG}' | fields COMMONAPACHELOG,
 timestamp, response, bytes ;
fetched rows / total rows = 4/4
```

```
+-----------------------------------------------------------------------------------------------------------
+----------------------------+-----------+---------+
| COMMONAPACHELOG
                                          | timestamp                  | response   |
 bytes   |
|-----------------------------------------------------------------------------------------------------------
+----------------------------+-----------+---------|
| 177.95.8.74 - upton5450 [28/Sep/2022:10:15:57 -0700] "HEAD /e-business/mindshare
 HTTP/1.0" 404 19927                       | 28/Sep/2022:10:15:57 -0700 | 404        |
 19927   |
| 127.45.152.6 - pouros8756 [28/Sep/2022:10:15:57 -0700] "GET /architectures/
convergence/niches/mindshare HTTP/1.0" 100 28722 | 28/Sep/2022:10:15:57 -0700 | 100
    | 28722   |
| *************** - - [28/Sep/2022:10:15:57 -0700] "PATCH /strategize/out-of-the-box
 HTTP/1.0" 401 27439                      | 28/Sep/2022:10:15:57 -0700 | 401        |
 27439   |
| ************** - - [28/Sep/2022:10:15:57 -0700] "POST /users HTTP/1.1" 301 9481
                                        | 28/Sep/2022:10:15:57 -0700 | 301        | 9481
    |
+-----------------------------------------------------------------------------------------------------------
+----------------------------+-----------+---------+
```

Limitazioni

Il comando grok presenta le stesse limitazioni del comando parse.

comando head

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `head` comando per restituire il primo numero N di risultati specificati dopo un offset opzionale nell'ordine di ricerca.

Sintassi

Utilizzare la seguente sintassi:

```
head [<size>] [from <offset>]
```

<size>

- Numero intero opzionale.
- Il numero di risultati da restituire.
- Impostazione predefinita: 10

<offset>

- Numero intero dopo facoltativo`from`.
- Il numero di risultati da saltare.
- Impostazione predefinita: 0

Esempio 1: ottieni i primi 10 risultati

Questo esempio mostra come recuperare un massimo di 10 risultati dall'indice dei conti.

Interrogazione PPL:

```
os> source=accounts | fields firstname, age | head;
fetched rows / total rows = 4/4
+-------------+-------+
| firstname   | age   |
|-------------+-------|
| Jane        | 32    |
| John        | 36    |
| Jorge       | 28    |
| Juan        | 33    |
+-------------+-------+
```

Esempio 2: ottieni i primi N risultati

L'esempio mostra i primi N risultati dell'indice dei conti.

Interrogazione PPL:

```
os> source=accounts | fields firstname, age | head 3;
fetched rows / total rows = 3/3
+-------------+-------+
| firstname   | age   |
```

```
|-------------+-------|
| Jane        | 32    |
| John        | 36    |
| Jorge       | 28    |
+-------------+-------+
```

Esempio 3: ottieni i primi N risultati dopo l'offset M

Questo esempio mostra come recuperare i primi N risultati dopo aver saltato M risultati dall'indice dei conti.

Interrogazione PPL:

```
os> source=accounts | fields firstname, age | head 3 from 1;
fetched rows / total rows = 3/3
+-------------+-------+
| firstname   | age   |
|-------------+-------|
| John        | 36    |
| Jorge       | 28    |
| Juan        | 33    |
+-------------+-------+
```

comando join

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Il comando join consente di combinare dati provenienti da più fonti in base a campi comuni, consentendoti di eseguire analisi complesse e ottenere informazioni più approfondite dai set di dati distribuiti

Schema

Esistono almeno due indici, `otel-v1-apm-span-*` (grande) e `otel-v1-apm-service-map` (piccolo).

Campi pertinenti degli indici:

`otel-v1-apm-span-*`

- TraceID: un identificatore univoco per una traccia. Tutti gli intervalli della stessa traccia condividono lo stesso traceID.
- spanID: un identificatore univoco per uno spazio all'interno di una traccia, assegnato al momento della creazione dello span.
- parentSpanId - Lo spanID dello span principale di questo spazio. Se si tratta di uno span principale, questo campo deve essere vuoto.
- durationInNanos - La differenza in nanosecondi tra StartTime ed EndTime. (questo è nell'interfaccia utente) `latency`
- Servicename: la risorsa da cui proviene lo span.
- TraceGroup: il nome dell'estensione principale della traccia.

`otel-v1-apm-service-map`

- ServiceName - Il nome del servizio che ha emesso lo span.
- destination.domain - Il ServiceName del servizio chiamato da questo client.
- destination.resource - Lo span name (API, operation e così via) chiamato da questo client.
- target.domain - Il ServiceName del servizio chiamato da un client.
- target.resource - Lo span name (API, operation e così via) chiamato da un client.
- traceGroupName - Lo span name di primo livello che ha avviato la catena di richieste.

Requisito

Support join per il calcolo di quanto segue:

Per ogni servizio, unisci span index all'indice della mappa dei servizi per calcolare le metriche in base a diversi tipi di filtri.

Questa query di esempio calcola la latenza quando viene filtrata per gruppo di traccia per il servizio. `client_cancel_order` `order`

```
SELECT avg(durationInNanos)
FROM `otel-v1-apm-span-000001` t1
WHERE t1.serviceName = `order`
  AND ((t1.name in
          (SELECT target.resource
```

```
           FROM `otel-v1-apm-service-map`
           WHERE serviceName = `order`
             AND traceGroupName = `client_cancel_order`)
        AND t1.parentSpanId != NULL)
       OR (t1.parentSpanId = NULL
           AND t1.name = `client_cancel_order`))
  AND t1.traceId in
    (SELECT traceId
     FROM `otel-v1-apm-span-000001`
     WHERE serviceName = `order`)
```

Esegui la migrazione a PPL

Sintassi del comando join

```
SEARCH source=<left-table>
| <other piped command>
| [joinType] JOIN
    [leftAlias]
    ON joinCriteria
    <right-table>
| <other piped command>
```

Riscrittura

```
SEARCH source=otel-v1-apm-span-000001
| WHERE serviceName = 'order'
| JOIN left=t1 right=t2
    ON t1.traceId = t2.traceId AND t2.serviceName = 'order'
    otel-v1-apm-span-000001 -- self inner join
| EVAL s_name = t1.name -- rename to avoid ambiguous
| EVAL s_parentSpanId = t1.parentSpanId -- RENAME command would be better when it is
 supported
| EVAL s_durationInNanos = t1.durationInNanos
| FIELDS s_name, s_parentSpanId, s_durationInNanos -- reduce colunms in join
| LEFT JOIN left=s1 right=t3
    ON s_name = t3.target.resource AND t3.serviceName = 'order' AND t3.traceGroupName =
 'client_cancel_order'
    otel-v1-apm-service-map
| WHERE (s_parentSpanId IS NOT NULL OR (s_parentSpanId IS NULL AND s_name =
 'client_cancel_order'))
| STATS avg(s_durationInNanos) -- no need to add alias if there is no ambiguous
```

Tipo di giunto

- Sintassi: `INNER | LEFT OUTER | CROSS`
- Facoltativo
- Il tipo di join da eseguire. L'impostazione predefinita è `INNER` se non è specificata.

Alias sinistro

- Sintassi: `left = <leftAlias>`
- Facoltativo
- L'alias della subquery da usare con il lato sinistro del join, per evitare denominazioni ambigue.

JoinCriteria

- Sintassi: `<expression>`
- Richiesto
- La sintassi inizia con. `ON` Potrebbe essere qualsiasi espressione di confronto. In genere, i criteri di unione sono i seguenti`<leftAlias>.<leftField>=<rightAlias>.<rightField>`.

  Ad esempio: `l.id = r.id`. Se i criteri di unione contengono più condizioni, è possibile specificare `AND` un `OR` operatore tra ogni espressione di confronto. Ad esempio, `l.id = r.id AND l.email = r.email AND (r.age > 65 OR r.age < 18)`.

Altri esempi

Migrazione da una query SQL (TPC-H Q13):

```
SELECT c_count, COUNT(*) AS custdist
FROM
  ( SELECT c_custkey, COUNT(o_orderkey) c_count
    FROM customer LEFT OUTER JOIN orders ON c_custkey = o_custkey
        AND o_comment NOT LIKE '%unusual%packages%'
    GROUP BY c_custkey
  ) AS c_orders
GROUP BY c_count
ORDER BY custdist DESC, c_count DESC;
```

Riscritto dalla join query PPL:

```
SEARCH source=customer
| FIELDS c_custkey
| LEFT OUTER JOIN
    ON c_custkey = o_custkey AND o_comment NOT LIKE '%unusual%packages%'
    orders
| STATS count(o_orderkey) AS c_count BY c_custkey
| STATS count() AS custdist BY c_count
| SORT - custdist, - c_count
```

Limitazione: le ricerche secondarie non sono supportate nella parte destra di join.

Se le ricerche secondarie sono supportate, è possibile riscrivere la precedente query PPL come segue:

```
SEARCH source=customer
| FIELDS c_custkey
| LEFT OUTER JOIN
   ON c_custkey = o_custkey
   [
      SEARCH source=orders
      | WHERE o_comment NOT LIKE '%unusual%packages%'
      | FIELDS o_orderkey, o_custkey
   ]
| STATS count(o_orderkey) AS c_count BY c_custkey
| STATS count() AS custdist BY c_count
| SORT - custdist, - c_count
```

comando lookup

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `lookup` comando per arricchire i dati di ricerca aggiungendo o sostituendo i dati da un indice di ricerca (tabella delle dimensioni). Questo comando consente di estendere i campi di un indice con i valori di una tabella delle dimensioni. È inoltre possibile utilizzarlo per aggiungere o sostituire valori quando vengono soddisfatte le condizioni di ricerca. Il `lookup` comando è più adatto del `Join` comando per arricchire i dati di origine con un set di dati statico.

Sintassi

Utilizzare la seguente sintassi:

```
SEARCH source=<sourceIndex>
| <other piped command>
| LOOKUP <lookupIndex> (<lookupMappingField> [AS <sourceMappingField>])...
    [(REPLACE | APPEND) (<inputField> [AS <outputField>])...]
| <other piped command>
```

Indice di ricerca

- Obbligatorio.
- Il nome dell'indice di ricerca (tabella delle dimensioni).

lookupMappingField

- Obbligatorio.
- Una chiave di mappatura nell'indice di ricerca, analoga a una chiave di join della tabella destra. È possibile specificare più campi, separati da virgole.

sourceMappingField

- Opzionale.
- Impostazione predefinita: **<** lookupMappingField **>**.
- Una chiave di mappatura dalla query di origine, analoga a una chiave di unione dal lato sinistro.

Campo di input

- Opzionale.
- Predefinito: tutti i campi dell'indice di ricerca in cui vengono trovati i valori corrispondenti.
- Un campo dell'indice di ricerca in cui i valori corrispondenti vengono applicati all'output dei risultati. È possibile specificare più campi, separati da virgole.

Campo di output

- Opzionale.

- Default: <inputField>.
- Un campo nell'output. È possibile specificare più campi di output. Se si specifica un nome di campo esistente dalla query di origine, i relativi valori verranno sostituiti o aggiunti dai valori corrispondenti di InputField. Se si specifica un nuovo nome di campo, questo verrà aggiunto ai risultati.

SOSTITUISCI | AGGIUNGI

- Opzionale.
- Predefinito: REPLACE
- Specifica come gestire i valori corrispondenti. Se si specifica REPLACE, i valori corrispondenti nel <lookupIndex>campo sovrascrivono i valori nel risultato. Se si specificaAPPEND, i valori corrispondenti nel <lookupIndex>campo vengono aggiunti solo ai valori mancanti nel risultato.

Utilizzo

- LOOKUP <lookupIndex>id AS cid SOSTITUISCI mail AS email
- <lookupIndex>NOME CERCA SOSTITUISCI mail COME email
- <lookupIndex>ID LOOKUP AS cid, nome AGGIUNGI indirizzo, mail COME email
- <lookupIndex>ID LOOKUP

Esempio

Fare riferimento agli esempi riportati di seguito.

```
SEARCH source=<sourceIndex>
| WHERE orderType = 'Cancelled'
| LOOKUP account_list, mkt_id AS mkt_code REPLACE amount, account_name AS name
| STATS count(mkt_code), avg(amount) BY name
```

```
SEARCH source=<sourceIndex>
| DEDUP market_id
| EVAL category=replace(category, "-", ".")
| EVAL category=ltrim(category, "dvp.")
| LOOKUP bounce_category category AS category APPEND classification
```

```
SEARCH source=<sourceIndex>
```

```
| LOOKUP bounce_category category
```

comando parse

Il `parse` comando analizza un campo di testo con un'espressione regolare e aggiunge il risultato al risultato della ricerca.

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Sintassi

Utilizzare la seguente sintassi:

```
parse <field> <pattern>
```

`field`

- Obbligatorio.
- Il campo deve essere un campo di testo.

`pattern`

- Stringa obbligatoria.
- Questo è il modello di espressione regolare usato per estrarre nuovi campi dal campo di testo specificato.
- Se esiste già un nuovo nome di campo, sostituirà il campo originale.

Espressione regolare

Il modello di espressione regolare viene utilizzato per abbinare l'intero campo di testo di ogni documento con il motore regex Java. Ogni gruppo di acquisizione denominato nell'espressione diventerà un nuovo `STRING` campo.

Esempio 1: creare un nuovo campo

L'esempio mostra come creare un nuovo campo `host` per ogni documento. `host`sarà il nome host dopo `@` nel `email` campo. L'analisi di un campo nullo restituirà una stringa vuota.

Interrogazione PPL:

```
os> source=accounts | parse email '.+@(?<host>.+)' | fields email, host ;
fetched rows / total rows = 4/4
+-----------------------+-------------+
| email                 | host        |
|-----------------------+-------------|
| jane_doe@example.com  | example.com |
| john_doe@example.net  | example.net |
| null                  |             |
| juan_li@example.org   | example.org |
+-----------------------+-------------+
```

Esempio 2: sovrascrivere un campo esistente

L'esempio mostra come sostituire il `address` campo esistente con il numero civico rimosso.

Interrogazione PPL:

```
os> source=accounts | parse address '\d+ (?<address>.+)' | fields address ;
fetched rows / total rows = 4/4
+------------------+
| address          |
|------------------|
| Example Lane     |
| Example Street   |
| Example Avenue   |
| Example Court    |
+------------------+
```

Esempio 3: filtra e ordina per campo analizzato

L'esempio mostra come ordinare i numeri civici che sono superiori a 500 nel `address` campo.

Interrogazione PPL:

```
os> source=accounts | parse address '(?<streetNumber>\d+) (?<street>.+)' | where
 cast(streetNumber as int) > 500 | sort num(streetNumber) | fields streetNumber,
 street ;
```

```
fetched rows / total rows = 3/3
+----------------+----------------+
| streetNumber   | street         |
|----------------+----------------|
| ***            | Example Street |
| ***            | Example Avenue |
| 880            | Example Lane   |
+----------------+----------------+
```

Limitazioni

Esistono alcune limitazioni con il comando parse:

- I campi definiti da parse non possono essere analizzati nuovamente.

  Il seguente comando non funzionerà:

  ```
  source=accounts | parse address '\d+ (?<street>.+)' | parse street '\w+ (?<road>\w+)'
  ```

- I campi definiti da parse non possono essere sovrascritti con altri comandi.

  `where`non corrisponderà a nessun documento poiché non `street` può essere sovrascritto:

  ```
  source=accounts | parse address '\d+ (?<street>.+)' | eval street='1' | where
   street='1' ;
  ```

- Il campo di testo utilizzato da parse non può essere sovrascritto.

  `street`non verrà analizzato correttamente poiché è stato sovrascritto: `address`

  ```
  source=accounts | parse address '\d+ (?<street>.+)' | eval address='1' ;
  ```

- I campi definiti da parse non possono essere filtrati o ordinati dopo averli utilizzati nel comando. `stats`

  `where`nel seguente comando non funzionerà:

  ```
  source=accounts | parse email '.+@(?<host>.+)' | stats avg(age) by host | where
   host=pyrami.com ;
  ```

comando patterns

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Il `patterns` comando estrae i modelli di registro da un campo di testo e aggiunge i risultati al risultato della ricerca. Il raggruppamento dei log in base ai relativi schemi semplifica l'aggregazione delle statistiche provenienti da grandi volumi di dati di registro per l'analisi e la risoluzione dei problemi.

Sintassi

Utilizzare la seguente sintassi:

```
patterns [new_field=<new-field-name>] [pattern=<pattern>] <field>
```

new-field-name

- Stringa opzionale.
- Questo è il nome del nuovo campo per i modelli estratti.
- Il valore predefinito è `patterns_field`.
- Se il nome esiste già, sostituirà il campo originale.

pattern

- Stringa opzionale.
- Questo è lo schema regex dei caratteri che devono essere filtrati dal campo di testo.
- Se assente, lo schema predefinito è costituito da caratteri alfanumerici (). `[a-zA-Z\d]`

field

- Obbligatorio.
- Il campo deve essere un campo di testo.

Esempio 1: creare il nuovo campo

L'esempio mostra come utilizzare l'estrazione della punteggiatura `email` per ogni documento. L'analisi di un campo nullo restituirà una stringa vuota.

Interrogazione PPL:

```
os> source=accounts | patterns email | fields email, patterns_field ;
fetched rows / total rows = 4/4
+----------------------+------------------+
| email                | patterns_field   |
|----------------------+------------------|
| jane_doe@example.com  | @.               |
| john_doe@example.net  | @.               |
| null                  |                  |
| juan_li@example.org   | @.               |
+----------------------+------------------+
```

Esempio 2: Estrarre i modelli di registro

L'esempio mostra come estrarre la punteggiatura da un campo di log non elaborato utilizzando i modelli predefiniti.

Interrogazione PPL:

```
os> source=apache | patterns message | fields message, patterns_field ;
fetched rows / total rows = 4/4
+-------------------------------------------------------------------------------------------------------
+---------------------------------+
| message
                                       | patterns_field                 |
|-------------------------------------------------------------------------------------------------------
+---------------------------------|
| 177.95.8.74 - upton5450 [28/Sep/2022:10:15:57 -0700] "HEAD /e-business/mindshare
 HTTP/1.0" 404 19927                         | ... -  [//::: -] " /-/ /."      |
| ************ - pouros8756 [28/Sep/2022:10:15:57 -0700] "GET /architectures/
convergence/niches/mindshare HTTP/1.0" 100 28722 | ... -  [//::: -] " //// /."     |
| *************** - - [28/Sep/2022:10:15:57 -0700] "PATCH /strategize/out-of-the-box
 HTTP/1.0" 401 27439                        | ... - - [//::: -] " //--- /."   |
| ************** - - [28/Sep/2022:10:15:57 -0700] "POST /users HTTP/1.1" 301 9481
                                       | ... - - [//::: -] " / /."       |
+-------------------------------------------------------------------------------------------------------
+---------------------------------+
```

Esempio 3: Estrai modelli di log con un modello regex personalizzato

L'esempio mostra come estrarre la punteggiatura da un campo di log non elaborato utilizzando modelli definiti dall'utente.

Interrogazione PPL:

```
os> source=apache | patterns new_field='no_numbers' pattern='[0-9]' message | fields
 message, no_numbers ;
fetched rows / total rows = 4/4
+-----------------------------------------------------------------------------------------------------
+-------------------------------------------------------------------------------
+
| message
                                      | no_numbers
                                      |
|-----------------------------------------------------------------------------------------------------
+-------------------------------------------------------------------------------|
| 177.95.8.74 - upton5450 [28/Sep/2022:10:15:57 -0700] "HEAD /e-business/mindshare
 HTTP/1.0" 404 19927                        | ... - upton [/Sep/::: -] "HEAD /e-
business/mindshare HTTP/."                          |
| 127.45.152.6 - pouros8756 [28/Sep/2022:10:15:57 -0700] "GET /architectures/
convergence/niches/mindshare HTTP/1.0" 100 28722 | ... - pouros [/Sep/::: -] "GET /
architectures/convergence/niches/mindshare HTTP/."   |
| *************** - - [28/Sep/2022:10:15:57 -0700] "PATCH /strategize/out-of-the-box
 HTTP/1.0" 401 27439                       | ... - - [/Sep/::: -] "PATCH /strategize/
out-of-the-box HTTP/."                         |
| ************** - - [28/Sep/2022:10:15:57 -0700] "POST /users HTTP/1.1" 301 9481
                                      | ... - - [/Sep/::: -] "POST /users HTTP/."
                                      |
+-----------------------------------------------------------------------------------------------------
+-------------------------------------------------------------------------------
+
```

Limitazione

Il comando patterns presenta le stesse limitazioni del comando parse.

comando raro

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `rare` comando per trovare la tupla di valori meno comune di tutti i campi nell'elenco dei campi.

> **Note**
>
> Vengono restituiti un massimo di 10 risultati per ogni distinta tupla di valori dei campi raggruppati per.

Sintassi

Utilizzare la seguente sintassi:

```
rare [N] <field-list> [by-clause] rare_approx [N] <field-list> [by-clause]
```

elenco di campi

- Obbligatorio.
- Un elenco di nomi di campo delimitato da virgole.

clausola secondaria

- Opzionale.
- Uno o più campi in base ai quali raggruppare i risultati.

N

- Il numero di risultati da restituire.
- Impostazione predefinita: 10

rare_approx

- Il conteggio approssimativo dei campi rari (n) utilizzando la cardinalità stimata mediante l'algoritmo ++. HyperLogLog

Esempio 1: trova i valori meno comuni in un campo

L'esempio trova il sesso meno comune tra tutti gli account.

Interrogazione PPL:

```
os> source=accounts | rare gender;
os> source=accounts | rare_approx 10 gender;
os> source=accounts | rare_approx gender;
fetched rows / total rows = 2/2
+----------+
| gender   |
|----------|
| F        |
| M        |
+----------+
```

Esempio 2: trova i valori meno comuni organizzati per genere

L'esempio trova l'età minima comune di tutti gli account raggruppati per genere.

Interrogazione PPL:

```
os> source=accounts | rare 5 age by gender;
os> source=accounts | rare_approx 5 age by gender;
fetched rows / total rows = 4/4
+----------+-------+
| gender   | age   |
|----------+-------|
| F        | 28    |
| M        | 32    |
| M        | 33    |
| M        | 36    |
+----------+-------+
```

comando di ridenominazione

Usa il `rename` comando per modificare i nomi di uno o più campi nel risultato della ricerca.

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Sintassi

Utilizzare la seguente sintassi:

```
rename <source-field> AS <target-field>["," <source-field> AS <target-field>]...
```

campo sorgente

- Obbligatorio.
- Questo è il nome del campo che vuoi rinominare.

campo-destinazione

- Obbligatorio.
- Questo è il nome con cui vuoi rinominare.

Esempio 1: Rinomina un campo

Questo esempio mostra come rinominare un singolo campo.

Interrogazione PPL:

```
os> source=accounts | rename account_number as an | fields an;
fetched rows / total rows = 4/4
+------+
| an   |
|------|
| 1    |
| 6    |
| 13   |
| 18   |
+------+
```

Esempio 2: rinominare più campi

Questo esempio mostra come rinominare più campi.

Interrogazione PPL:

```
os> source=accounts | rename account_number as an, employer as emp | fields an, emp;
fetched rows / total rows = 4/4
+------+---------+
| an   | emp     |
|------+---------|
| 1    | Pyrami  |
| 6    | Netagy  |
| 13   | Quility |
| 18   | null    |
+------+---------+
```

Limitazioni

- La sovrascrittura del campo esistente non è supportata:

```
source=accounts | grok address '%{NUMBER} %{GREEDYDATA:address}' | fields address
```

comando di ricerca

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `search` comando per recuperare i documenti da un indice. Il `search` comando può essere utilizzato solo come primo comando in una query PPL.

Sintassi

Utilizzare la seguente sintassi:

```
search source=[<remote-cluster>:]<index> [boolean-expression]
```

cerca

- Opzionale.
- Parole chiave di ricerca, che possono essere omesse.

indice

- Obbligatorio.
- Il comando search deve specificare da quale indice eseguire la ricerca.
- Il nome dell'indice può essere preceduto da `<cluster name>:` per le ricerche tra cluster.

espressione bool

- Opzionale.
- Qualsiasi espressione che restituisce un valore booleano.

Esempio 1: recupera tutti i dati

L'esempio mostra il recupero di tutto il documento dall'indice degli account.

Interrogazione PPL:

```
os> source=accounts;
+------------------+-------------+----------------------+-----------+----------
+--------+----------------+---------+-------+-----------------------+------------+
| account_number   | firstname   | address              | balance   | gender   | city
 | employer       | state   | age   | email                 | lastname   |
|------------------+-------------+----------------------+-----------+----------
+--------+----------------+---------+-------+-----------------------+------------|
| 1                | Jorge       | *** Any Lane         | 39225     | M        | Brogan
 | ExampleCorp    | IL      | 32    | jane_doe@example.com  | Souza      |
| 6                | John        | *** Example Street   | 5686      | M        | Dante
 | AnyCorp        | TN      | 36    | john_doe@example.com  | Doe        |
| 13               | Jane        | *** Any Street       | *****     | F        | Nogal
 | ExampleCompany | VA      | 28    | null                  | Doe        |
| 18               | Juan        | *** Example Court    | 4180      | M        | Orick
 | null           | MD      | 33    | juan_li@example.org   | Li         |
```

```
+------------------+-------------+----------------------+-----------+----------
+--------+----------------+---------+-------+-----------------------+------------+
```

Esempio 2: recupera dati con condizione

L'esempio mostra recuperare tutto il documento dall'indice degli account con.

Interrogazione PPL:

```
os> SEARCH source=accounts account_number=1 or gender="F";
+------------------+-------------+--------------------+-----------+----------+--------
+----------------+---------+-------+-------------------------+------------+
| account_number   | firstname   | address            | balance   | gender   | city   |
 employer       | state   | age   | email                 -  | lastname   |
|------------------+-------------+--------------------+-----------+----------+--------
+----------------+---------+-------+-------------------------+------------|
| 1                | Jorge       | *** Any Lane       | *****     | M        | Brogan |
 ExampleCorp    | IL      | 32    | jorge_souza@example.com | Souza      |
| 13               | Jane        | *** Any Street     | *****     | F        | Nogal  |
 ExampleCompany | VA      | 28    | null                    | Doe        |
+------------------+-------------+--------------------+-----------+----------+--------
+-----------------+---------+-------+-------------------------+------------+
```

comando sort

Usa il `sort` comando per ordinare i risultati della ricerca in base ai campi specificati.

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Sintassi

Utilizzare la seguente sintassi:

```
sort <[+|-] sort-field>...
```

[+|-]

- Opzionale.
- Il segno più [+] indica l'ordine crescente con NULL/MISSING i valori al primo posto.
- Il segno meno [-] indica l'ordine decrescente con NULL/MISSING i valori per ultimi.
- Predefinito: ordine crescente con NULL/MISSING i valori per primi.

campo di ordinamento

- Obbligatorio.
- Il campo utilizzato per l'ordinamento.

Esempio 1: ordina per un campo

L'esempio mostra come ordinare il documento con il campo dell'età in ordine crescente.

Interrogazione PPL:

```
os> source=accounts | sort age | fields account_number, age;
fetched rows / total rows = 4/4
+------------------+-------+
| account_number   | age   |
|------------------+-------|
| 13               | 28    |
| 1                | 32    |
| 18               | 33    |
| 6                | 36    |
+------------------+-------+
```

Esempio 2: ordina per un campo e restituisci tutti i risultati

L'esempio mostra come ordinare il documento con il campo età in ordine crescente.

Interrogazione PPL:

```
os> source=accounts | sort age | fields account_number, age;
fetched rows / total rows = 4/4
+------------------+-------+
| account_number   | age   |
```

```
|------------------+-------|
| 13               | 28    |
| 1                | 32    |
| 18               | 33    |
| 6                | 36    |
+------------------+-------+
```

Esempio 3: ordina per un campo in ordine decrescente

L'esempio mostra come ordinare il documento con il campo età in ordine decrescente.

Interrogazione PPL:

```
os> source=accounts | sort - age | fields account_number, age;
fetched rows / total rows = 4/4
+------------------+-------+
| account_number   | age   |
|------------------+-------|
| 6                | 36    |
| 18               | 33    |
| 1                | 32    |
| 13               | 28    |
+------------------+-------+
```

Esempio 4: ordina per più campi

L'esempio mostra come ordinare il documento con il campo genere in ordine crescente e il campo età in ordine decrescente.

Interrogazione PPL:

```
os> source=accounts | sort + gender, - age | fields account_number, gender, age;
fetched rows / total rows = 4/4
+------------------+----------+-------+
| account_number   | gender   | age   |
|------------------+----------+-------|
| 13               | F        | 28    |
| 6                | M        | 36    |
| 18               | M        | 33    |
| 1                | M        | 32    |
+------------------+----------+-------+
```

Esempio 5: Ordina per campo e include un valore nullo

L'esempio mostra come ordinare il campo del datore di lavoro in base all'opzione predefinita (ordine crescente e prima null). Il risultato mostra che il valore nullo è nella prima riga.

Interrogazione PPL:

```
os> source=accounts | sort employer | fields employer;
fetched rows / total rows = 4/4
+------------+
| employer   |
|------------|
| null       |
| AnyCompany |
| AnyCorp    |
| AnyOrgty   |
+------------+
```

comando stats

Usa il `stats` comando per calcolare l'aggregazione dai risultati della ricerca.

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Gestione dei valori NULL/MANCANTI

Gestione dei valori NULL/MANCANTI

Funzione	NULL	MANCANTE
COUNT	Non contato	Non contato
SUM	Ignorare	Ignorare
AVG	Ignorare	Ignorare
MAX	Ignorare	Ignorare

Funzione	NULL	MANCANTE
MIN	Ignorare	Ignorare

Sintassi

Utilizzare la seguente sintassi:

```
stats <aggregation>... [by-clause]
```

Aggregazione

- Obbligatorio.
- Una funzione di aggregazione applicata a un campo.

clausola secondaria

- Opzionale.
- Sintassi: `by [span-expression,] [field,]...`
- Specifica i campi e le espressioni per il raggruppamento dei risultati dell'aggregazione. La by-clausola consente di raggruppare i risultati dell'aggregazione utilizzando campi ed espressioni. È possibile utilizzare funzioni scalari, funzioni di aggregazione e persino espressioni span per dividere campi specifici in bucket di intervalli uguali.
- Predefinito: se `<by-clause>` viene specificato no, il comando stats restituisce una singola riga che rappresenta l'aggregazione sull'intero set di risultati.

espressione spanx

- Facoltativo, al massimo uno.
- Sintassi: `span(field_expr, interval_expr)`
- L'unità dell'espressione dell'intervallo è l'unità naturale di default. Se il campo è di tipo data e ora e l'intervallo è espresso in date/time unità, è necessario specificare l'unità nell'espressione dell'intervallo.
- Ad esempio, suddividendo il `age` campo in secchi per 10 anni, sembra. `span(age, 10)` Per dividere un campo timestamp in intervalli orari, usa. `span(timestamp, 1h)`

Unità di tempo disponibili

Unità di intervallo di ampiezza
millisecondo (ms)
secondo/i
minuto (m, distinzione tra maiuscole e minuscole)
ora (h)
giorno (d)
settimana (e)
mese (M, distinzione tra maiuscole e minuscole)
trimestre (q)
anno (y)

Funzioni di aggregazione

COUNT

Restituisce un conteggio del numero di espr nelle righe recuperate da un'istruzione SELECT.

Esempio:

```
os> source=accounts | stats count();
fetched rows / total rows = 1/1
+-----------+
| count()   |
|-----------|
| 4         |
+-----------+
```

SUM

Si usa `SUM(expr)` per restituire la somma di expr.

Esempio

```
os> source=accounts | stats sum(age) by gender;
fetched rows / total rows = 2/2
+------------+----------+
| sum(age)   | gender   |
|------------+----------|
| 28         | F        |
| 101        | M        |
+------------+----------+
```

AVG

Usa `AVG(expr)` per restituire il valore medio di expr.

Esempio

```
os> source=accounts | stats avg(age) by gender;
fetched rows / total rows = 2/2
+--------------------+----------+
| avg(age)           | gender   |
|--------------------+----------|
| 28.0               | F        |
| 33.666666666666664 | M        |
+--------------------+----------+
```

MAX

Usa `MAX(expr)` per restituire il valore massimo di expr.

Esempio

```
os> source=accounts | stats max(age);
fetched rows / total rows = 1/1
+------------+
| max(age)   |
|------------|
| 36         |
+------------+
```

MIN

Usa `MIN(expr)` per restituire il valore minimo di expr.

Esempio

```
os> source=accounts | stats min(age);
fetched rows / total rows = 1/1
+------------+
| min(age)   |
|------------|
| 28         |
+------------+
```

STDDEV_SAMP

Si usa `STDDEV_SAMP(expr)` per restituire la deviazione standard campionaria di expr.

Esempio:

```
os> source=accounts | stats stddev_samp(age);
fetched rows / total rows = 1/1
+--------------------+
| stddev_samp(age)   |
|--------------------|
| 3.304037933599835  |
+--------------------+
```

STDDEV_POP

Utilizzare `STDDEV_POP(expr)` per restituire la deviazione standard della popolazione di expr.

Esempio:

```
os> source=accounts | stats stddev_pop(age);
fetched rows / total rows = 1/1
+--------------------+
| stddev_pop(age)    |
|--------------------|
| 2.**************** |
+--------------------+
```

AFFERRARE

Si usa `TAKE(field [, size])` per restituire i valori originali di un campo. Non garantisce l'ordine dei valori.

field

- Obbligatorio.
- Il campo deve essere un campo di testo.

formato

- Numero intero opzionale.
- Il numero di valori deve essere restituito.
- L'impostazione predefinita è 10.

Esempio

```
os> source=accounts | stats take(firstname);
fetched rows / total rows = 1/1
+-----------------------------+
| take(firstname)             |
|-----------------------------|
| [Jane, Mary, Nikki, Juan    |
+-----------------------------+
```

PERCENTILE o PERCENTILE_APPROX

Usa `PERCENTILE(expr, percent)` o `PERCENTILE_APPROX(expr, percent)` per restituire il valore percentile approssimativo di expr alla percentuale specificata.

percentuale

- Il numero deve essere una costante compresa tra 0 e 100.

Esempio

```
os> source=accounts | stats percentile(age, 90) by gender;
```

```
fetched rows / total rows = 2/2
+-----------------------+----------+
| percentile(age, 90)   | gender   |
|-----------------------+----------|
| 28                    | F        |
| 36                    | M        |
+-----------------------+----------+
```

Esempio 1: calcola il conteggio degli eventi

L'esempio mostra come calcolare il conteggio degli eventi nei conti.

```
os> source=accounts | stats count();
fetched rows / total rows = 1/1
+-----------+
| count()   |
|-----------|
| 4         |
+-----------+
```

Esempio 2: calcola la media di un campo

L'esempio mostra come calcolare l'età media per tutti gli account.

```
os> source=accounts | stats avg(age);
fetched rows / total rows = 1/1
+------------+
| avg(age)   |
|------------|
| 32.25      |
+------------+
```

Esempio 3: calcola la media di un campo per gruppo

L'esempio mostra come calcolare l'età media per tutti gli account, raggruppati per sesso.

```
os> source=accounts | stats avg(age) by gender;
fetched rows / total rows = 2/2
+--------------------+----------+
| avg(age)           | gender   |
|--------------------+----------|
```

```
| 28.0               | F        |
| 33.666666666666664 | M        |
+--------------------+----------+
```

Esempio 4: calcola la media, la somma e il conteggio di un campo per gruppo

L'esempio mostra come calcolare l'età media, la somma dell'età e il conteggio degli eventi per tutti gli account, raggruppati per sesso.

```
os> source=accounts | stats avg(age), sum(age), count() by gender;
fetched rows / total rows = 2/2
+--------------------+------------+-----------+----------+
| avg(age)           | sum(age)   | count()   | gender   |
|--------------------+------------+-----------+----------|
| 28.0               | 28         | 1         | F        |
| 33.666666666666664 | 101        | 3         | M        |
+--------------------+------------+-----------+----------+
```

Esempio 5: calcola il massimo di un campo

L'esempio calcola l'età massima per tutti gli account.

```
os> source=accounts | stats max(age);
fetched rows / total rows = 1/1
+------------+
| max(age)   |
|------------|
| 36         |
+------------+
```

Esempio 6: calcola il massimo e il minimo di un campo per gruppo

L'esempio calcola i valori di età massima e minima per tutti gli account, raggruppati per sesso.

```
os> source=accounts | stats max(age), min(age) by gender;
fetched rows / total rows = 2/2
+------------+------------+----------+
| max(age)   | min(age)   | gender   |
|------------+------------+----------|
| 28         | 28         | F        |
| 36         | 32         | M        |
```

```
+------------+-----------+----------+
```

Esempio 7: calcola il conteggio distinto di un campo

Per ottenere il conteggio dei valori distinti di un campo, puoi usare la funzione `DISTINCT_COUNT` (or`DC`) invece di`COUNT`. L'esempio calcola sia il campo di conteggio che quello di conteggio distinto per genere di tutti gli account.

```
os> source=accounts | stats count(gender), distinct_count(gender);
fetched rows / total rows = 1/1
+-----------------+--------------------------+
| count(gender)   | distinct_count(gender)   |
|-----------------+--------------------------|
| 4               | 2                        |
+-----------------+--------------------------+
```

Esempio 8: calcola il conteggio in base a un intervallo

L'esempio ottiene il conteggio dell'età per un intervallo di 10 anni.

```
os> source=accounts | stats count(age) by span(age, 10) as age_span
fetched rows / total rows = 2/2
+--------------+------------+
| count(age)   | age_span   |
|--------------+------------|
| 1            | 20         |
| 3            | 30         |
+--------------+------------+
```

Esempio 9: calcola il conteggio in base al sesso e all'intervallo

Questo esempio conta i record raggruppati per sesso e intervalli di età di 5 anni.

```
os> source=accounts | stats count() as cnt by span(age, 5) as age_span, gender
fetched rows / total rows = 3/3
+-------+------------+----------+
| cnt   | age_span   | gender   |
|-------+------------+----------|
| 1     | 25         | F        |
| 2     | 30         | M        |
| 1     | 35         | M        |
```

```
+-------+-----------+----------+
```

L'espressione span appare sempre come prima chiave di raggruppamento, indipendentemente dall'ordine specificato nel comando.

```
os> source=accounts | stats count() as cnt by gender, span(age, 5) as age_span
fetched rows / total rows = 3/3
+-------+-----------+----------+
| cnt   | age_span   | gender   |
|-------+-----------+----------|
| 1     | 25         | F        |
| 2     | 30         | M        |
| 1     | 35         | M        |
+-------+-----------+----------+
```

Esempio 10: calcola il conteggio e ottieni l'elenco di e-mail in base al sesso e all'intervallo

L'esempio ottiene il conteggio dell'età per l'intervallo di 10 anni e il gruppo per sesso, inoltre per ogni riga ottieni un elenco di al massimo 5 email.

```
os> source=accounts | stats count() as cnt, take(email, 5) by span(age, 5) as age_span,
 gender
fetched rows / total rows = 3/3
+-------+-----------------------------------------------------+-----------+----------+
| cnt   | take(email, 5)                                      | age_span   | gender   |
|-------+-----------------------------------------------------+-----------+----------|
| 1     | []                                                  | 25         | F        |
| 2     | [janedoe@anycompany.com,juanli@examplecompany.org] | 30         | M        |
| 1     | [marymajor@examplecorp.com]                         | 35         | M        |
+-------+-----------------------------------------------------+-----------+----------+
```

Esempio 11: calcola il percentile di un campo

L'esempio mostra come calcolare il percentile di 90° età di tutti gli account.

```
os> source=accounts | stats percentile(age, 90);
fetched rows / total rows = 1/1
+----------------------+
| percentile(age, 90)   |
|----------------------|
| 36                    |
```

```
+-----------------------+
```

Esempio 12: calcola il percentile di un campo per gruppo

L'esempio mostra come calcolare il percentile di 90° età di tutti i gruppi di account per sesso.

```
os> source=accounts | stats percentile(age, 90) by gender;
fetched rows / total rows = 2/2
+-----------------------+----------+
| percentile(age, 90)   | gender   |
|-----------------------+----------|
| 28                    | F        |
| 36                    | M        |
+-----------------------+----------+
```

Esempio 13: calcola il percentile in base al sesso e all'intervallo

L'esempio ottiene il percentile di 90° età per intervallo di 10 anni e il gruppo per sesso.

```
os> source=accounts | stats percentile(age, 90) as p90 by span(age, 10) as age_span,
 gender
fetched rows / total rows = 2/2
+-------+------------+----------+
| p90   | age_span   | gender   |
|-------+------------+----------|
| 28    | 20         | F        |
| 36    | 30         | M        |
+-------+------------+----------+
```

```
- `source = table | stats avg(a) `
- `source = table | where a < 50 | stats avg(c) `
- `source = table | stats max(c) by b`
- `source = table | stats count(c) by b | head 5`
- `source = table | stats distinct_count(c)`
- `source = table | stats stddev_samp(c)`
- `source = table | stats stddev_pop(c)`
- `source = table | stats percentile(c, 90)`
- `source = table | stats percentile_approx(c, 99)`
```

Aggregazioni con span

```
- `source = table  | stats count(a) by span(a, 10) as a_span`
- `source = table  | stats sum(age) by span(age, 5) as age_span | head 2`
- `source = table  | stats avg(age) by span(age, 20) as age_span, country  | sort -
 age_span |  head 2`
```

Aggregazioni con timewindow span (funzione tumble windowing)

```
- `source = table | stats sum(productsAmount) by span(transactionDate, 1d) as age_date
 | sort age_date`
- `source = table | stats sum(productsAmount) by span(transactionDate, 1w) as age_date,
 productId`
```

Le aggregazioni sono raggruppate per più livelli

```
- `source = table | stats avg(age) as avg_state_age by country, state | stats
 avg(avg_state_age) as avg_country_age by country`
- `source = table | stats avg(age) as avg_city_age by country, state, city | eval
 new_avg_city_age = avg_city_age - 1 | stats avg(new_avg_city_age) as avg_state_age
 by country, state | where avg_state_age > 18 | stats avg(avg_state_age) as
 avg_adult_country_age by country`
```

comando subquery

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `subquery` comando per eseguire query complesse e annidate all'interno delle istruzioni PPL (Piped Processing Language).

```
source=logs | where field in [ subquery source=events | where condition | fields
 field ]
```

In questo esempio, la ricerca primaria (`source=logs`) viene filtrata in base ai risultati della subquery (). `source=events`

Il comando subquery supporta più livelli di annidamento per l'analisi di dati complessi.

Esempio di subquery annidata

```
source=logs | where id in [ subquery source=users | where user in [ subquery
 source=actions | where action="login" | fields user] | fields uid ]
```

InSubquery Utilizzo

- `source = outer | where a in [ source = inner | fields b ]`
- `source = outer | where (a) in [ source = inner | fields b ]`
- `source = outer | where (a,b,c) in [ source = inner | fields d,e,f ]`
- `source = outer | where a not in [ source = inner | fields b ]`
- `source = outer | where (a) not in [ source = inner | fields b ]`
- `source = outer | where (a,b,c) not in [ source = inner | fields d,e,f ]`
- `source = outer a in [ source = inner | fields b ]`(filtro di ricerca con sottoquery)
- `source = outer a not in [ source = inner | fields b ]`(filtro di ricerca con subquery)
- `source = outer | where a in [ source = inner1 | where b not in [ source = inner2 | fields c ] | fields b ]`(annidato)
- `source = table1 | inner join left = l right = r on l.a = r.a AND r.a in [ source = inner | fields d ] | fields l.a, r.a, b, c`(come filtro di unione)

Esempi di migrazione SQL con In-Subquery PPL

TPC-H Q4 (nella subquery con aggregazione)

```
select
  o_orderpriority,
  count(*) as order_count
from
  orders
where
  o_orderdate >= date '1993-07-01'
  and o_orderdate < date '1993-07-01' + interval '3' month
  and o_orderkey in (
```

```
    select
      l_orderkey
    from
      lineitem
    where l_commitdate < l_receiptdate
  )
group by
  o_orderpriority
order by
  o_orderpriority
```

Riscritto da InSubquery una query PPL:

```
source = orders
| where o_orderdate >= "1993-07-01" and o_orderdate < "1993-10-01" and o_orderkey IN
  [ source = lineitem
    | where l_commitdate < l_receiptdate
    | fields l_orderkey
  ]
| stats count(1) as order_count by o_orderpriority
| sort o_orderpriority
| fields o_orderpriority, order_count
```

TPC-H Q20 (annidato nella sottoquery)

```
select
  s_name,
  s_address
from
  supplier,
  nation
where
  s_suppkey in (
    select
      ps_suppkey
    from
      partsupp
    where
      ps_partkey in (
        select
          p_partkey
        from
          part
```

```
        where
          p_name like 'forest%'
      )
  )
  and s_nationkey = n_nationkey
  and n_name = 'CANADA'
order by
  s_name
```

Riscritto da una query PPL: InSubquery

```
source = supplier
| where s_suppkey IN [
    source = partsupp
    | where ps_partkey IN [
        source = part
        | where like(p_name, "forest%")
        | fields p_partkey
      ]
    | fields ps_suppkey
  ]
| inner join left=l right=r on s_nationkey = n_nationkey and n_name = 'CANADA'
  nation
| sort s_name
```

ExistsSubquery utilizzo

Ipotesi:a, b sono campi della tabella esternac, d sono campi della tabella internae, f sono campi della tabella inner2.

- `source = outer | where exists [ source = inner | where a = c ]`
- `source = outer | where not exists [ source = inner | where a = c ]`
- `source = outer | where exists [ source = inner | where a = c and b = d ]`
- `source = outer | where not exists [ source = inner | where a = c and b = d ]`
- `source = outer exists [ source = inner | where a = c ]`(filtro di ricerca con sottoquery)
- `source = outer not exists [ source = inner | where a = c ]`(filtro di ricerca con subquery)

- `source = table as t1 exists [ source = table as t2 | where t1.a = t2.a ]`(L'alias della tabella è utile nella sottoquery exists)
- `source = outer | where exists [ source = inner1 | where a = c and exists [ source = inner2 | where c = e ] ]`(annidato)
- `source = outer | where exists [ source = inner1 | where a = c | where exists [ source = inner2 | where c = e ] ]`(annidato)
- `source = outer | where exists [ source = inner | where c > 10 ]`(esiste una correlazione non correlata)
- `source = outer | where not exists [ source = inner | where c > 10 ]`(esiste una correlazione non correlata)
- `source = outer | where exists [ source = inner ] | eval l = "nonEmpty" | fields l`(esistono elementi non correlati speciali)

ScalarSubquery utilizzo

Presupposti:`a`, `b` sono campi della tabella esterni`c`, `d` sono campi della tabella interna`e`, `f` sono campi della tabella annidati

Sottoquery scalare non correlata

In Seleziona:

- `source = outer | eval m = [ source = inner | stats max(c) ] | fields m, a`
- `source = outer | eval m = [ source = inner | stats max(c) ] + b | fields m, a`

In dove:

- `source = outer | where a > [ source = inner | stats min(c) ] | fields a`

Nel filtro di ricerca:

- `source = outer a > [ source = inner | stats min(c) ] | fields a`

Sottoquery scalare correlata

In Seleziona:

- `source = outer | eval m = [ source = inner | where outer.b = inner.d | stats max(c) ] | fields m, a`
- `source = outer | eval m = [ source = inner | where b = d | stats max(c) ] | fields m, a`
- `source = outer | eval m = [ source = inner | where outer.b > inner.d | stats max(c) ] | fields m, a`

In dove:

- `source = outer | where a = [ source = inner | where outer.b = inner.d | stats max(c) ]`
- `source = outer | where a = [ source = inner | where b = d | stats max(c) ]`
- `source = outer | where [ source = inner | where outer.b = inner.d OR inner.d = 1 | stats count() ] > 0 | fields a`

Nel filtro di ricerca:

- `source = outer a = [ source = inner | where b = d | stats max(c) ]`
- `source = outer [ source = inner | where outer.b = inner.d OR inner.d = 1 | stats count() ] > 0 | fields a`

Sottoquery scalare annidata

- `source = outer | where a = [ source = inner | stats max(c) | sort c ] OR b = [ source = inner | where c = 1 | stats min(d) | sort d ]`
- `source = outer | where a = [ source = inner | where c = [ source = nested | stats max(e) by f | sort f ] | stats max(d) by c | sort c | head 1 ]`

(Relazione) Sottoquery

`InSubquery`, `ExistsSubquery` e `ScalarSubquery` sono tutte espressioni di sottoquery. Ma non `RelationSubquery` è un'espressione di sottoquery, è un piano di sottoquery che è comunemente usato nella clausola Join o From.

- `source = table1 | join left = l right = r [ source = table2 | where d > 10 | head 5 ]`(subquery nella parte destra di join)
- `source = [ source = table1 | join left = l right = r [ source = table2 | where d > 10 | head 5 ] | stats count(a) by b ] as outer | head 1`

Contesto aggiuntivo

`InSubqueryExistsSubquery`, e `ScalarSubquery` sono espressioni di sottoquery comunemente utilizzate nelle `where` clausole e nei filtri di ricerca.

Comando Where:

```
| where <boolean expression> | ...
```

Filtro di ricerca:

```
search source=* <boolean expression> | ...
```

Un'espressione di sottoquery può essere utilizzata in un'espressione booleana:

```
| where orders.order_id in [ source=returns | where return_reason="damaged" | field
 order_id ]
```

Il `orders.order_id in [ source=... ]` è un. `<boolean expression>`

In generale, chiamiamo questo tipo di clausola di sottoquery espressione. `InSubquery` È un. `<boolean expression>`

Sottoquery con diversi tipi di join

Esempio utilizzando un `ScalarSubquery`:

```
source=employees
| join source=sales on employees.employee_id = sales.employee_id
| where sales.sale_amount > [ source=targets | where target_met="true" | fields
 target_value ]
```

A differenza di InSubquery ExistsSubquery, e ScalarSubquery, a non RelationSubquery è un'espressione di sottoquery. Si tratta invece di un piano di subquery.

```
SEARCH source=customer
| FIELDS c_custkey
| LEFT OUTER JOIN left = c, right = o ON c.c_custkey = o.o_custkey
   [
      SEARCH source=orders
      | WHERE o_comment NOT LIKE '%unusual%packages%'
      | FIELDS o_orderkey, o_custkey
   ]
| STATS ...
```

comando top

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `top` comando per trovare la tupla di valori più comune di tutti i campi nell'elenco dei campi.

Sintassi

Utilizzare la seguente sintassi:

```
top [N] <field-list> [by-clause] top_approx [N] <field-list> [by-clause]
```

N

- Il numero di risultati da restituire.
- Impostazione predefinita: 10

elenco di campi

- Obbligatorio.
- Un elenco di nomi di campo delimitato da virgole.

clausola secondaria

- Opzionale.
- Uno o più campi in base ai quali raggruppare i risultati.

top_approx

- Un conteggio approssimativo dei campi (n) principali utilizzando la cardinalità stimata mediante l'algoritmo ++. HyperLogLog

Esempio 1: trova i valori più comuni in un campo

L'esempio trova il sesso più comune per tutti gli account.

Interrogazione PPL:

```
os> source=accounts | top gender;
os> source=accounts | top_approx gender;
fetched rows / total rows = 2/2
+----------+
| gender   |
|----------|
| M        |
| F        |
+----------+
```

Esempio 2: trova i valori più comuni in un campo (limitato a 1)

L'esempio individua il genere singolo più comune per tutti gli account.

Interrogazione PPL:

```
os> source=accounts | top_approx 1 gender;
fetched rows / total rows = 1/1
+----------+
| gender   |
|----------|
| M        |
+----------+
```

Esempio 3: trova i valori più comuni, raggruppati per genere

L'esempio trova l'età più comune per tutti gli account, raggruppati per sesso.

Interrogazione PPL:

```
os> source=accounts | top 1 age by gender;
os> source=accounts | top_approx 1 age by gender;
fetched rows / total rows = 2/2
+----------+-------+
| gender   | age   |
|----------+-------|
| F        | 28    |
| M        | 32    |
+----------+-------+
```

comando trendline

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `trendline` comando per calcolare le medie mobili dei campi.

Sintassi

Utilizzate la seguente sintassi

```
TRENDLINE [sort <[+|-] sort-field>] SMA(number-of-datapoints, field) [AS alias]
 [SMA(number-of-datapoints, field) [AS alias]]...
```

[+|-]

- Opzionale.
- Il segno più [+] indica l'ordine crescente con NULL/MISSING i valori al primo posto.
- Il segno meno [-] indica l'ordine decrescente con NULL/MISSING i valori per ultimi.
- Predefinito: ordine crescente con NULL/MISSING i valori per primi.

campo di ordinamento

- Obbligatorio quando si utilizza l'ordinamento.
- Il campo utilizzato per l'ordinamento.

number-of-datapoints

- Obbligatorio.
- Il numero di punti dati che calcolano la media mobile.
- Deve essere maggiore di zero.

field

- Obbligatorio.
- Il nome del campo per cui deve essere calcolata la media mobile.

alias

- Opzionale.
- Il nome della colonna risultante contenente la media mobile.

È supportato solo il tipo Simple Moving Average (SMA). Viene calcolato in questo modo:

```
f[i]: The value of field 'f' in the i-th data-point
n: The number of data-points in the moving window (period)
t: The current time index

SMA(t) = (1/n) * Σ(f[i]), where i = t-n+1 to t
```

Esempio 1: calcola la media mobile semplice per una serie temporale di temperature

L'esempio calcola la media mobile semplice rispetto alle temperature utilizzando due punti dati.

Interrogazione PPL:

```
os> source=t | trendline sma(2, temperature) as temp_trend;
fetched rows / total rows = 5/5
+-----------+---------+--------------------+----------+
```

```
|temperature|device-id|           timestamp|temp_trend|
+-----------+---------+--------------------+----------+
|         12|     1492|2023-04-06 17:07:...|      NULL|
|         12|     1492|2023-04-06 17:07:...|      12.0|
|         13|      256|2023-04-06 17:07:...|      12.5|
|         14|      257|2023-04-06 17:07:...|      13.5|
|         15|      258|2023-04-06 17:07:...|      14.5|
+-----------+---------+--------------------+----------+
```

Esempio 2: calcola medie mobili semplici per una serie temporale di temperature con ordinamento

L'esempio calcola due medie mobili semplici sulle temperature utilizzando due e tre punti dati ordinati in modo decrescente per device-id.

Interrogazione PPL:

```
os> source=t | trendline sort - device-id sma(2, temperature) as temp_trend_2 sma(3,
 temperature) as temp_trend_3;
fetched rows / total rows = 5/5
+-----------+---------+--------------------+------------+------------------+
|temperature|device-id|           timestamp|temp_trend_2|      temp_trend_3|
+-----------+---------+--------------------+------------+------------------+
|         15|      258|2023-04-06 17:07:...|        NULL|              NULL|
|         14|      257|2023-04-06 17:07:...|        14.5|              NULL|
|         13|      256|2023-04-06 17:07:...|        13.5|              14.0|
|         12|     1492|2023-04-06 17:07:...|        12.5|              13.0|
|         12|     1492|2023-04-06 17:07:...|        12.0|12.333333333333334|
+-----------+---------+--------------------+------------+------------------+
```

comando where

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Il `where` comando utilizza un'espressione bool per filtrare i risultati della ricerca. Restituisce il risultato solo quando bool-expression restituisce true.

Sintassi

Utilizzare la seguente sintassi:

```
where <boolean-expression>
```

espressione bool

- Opzionale.
- Qualsiasi espressione che possa essere valutata in base a un valore booleano.

Esempio 1: filtra il set di risultati con condizione

L'esempio mostra come recuperare dall'indice degli account i documenti che soddisfano condizioni specifiche.

Interrogazione PPL:

```
os> source=accounts | where account_number=1 or gender="F" | fields account_number,
 gender;
fetched rows / total rows = 2/2
+------------------+----------+
| account_number   | gender   |
|------------------+----------|
| 1                | M        |
| 13               | F        |
+------------------+----------+
```

Esempi aggiuntivi

Filtri con condizioni logiche

- `source = table | where c = 'test' AND a = 1 | fields a,b,c`
- `source = table | where c != 'test' OR a > 1 | fields a,b,c | head 1`
- `source = table | where c = 'test' NOT a > 1 | fields a,b,c`
- `source = table | where a = 1 | fields a,b,c`
- `source = table | where a >= 1 | fields a,b,c`
- `source = table | where a < 1 | fields a,b,c`

- `source = table | where b != 'test' | fields a,b,c`
- `source = table | where c = 'test' | fields a,b,c | head 3`
- `source = table | where ispresent(b)`
- `source = table | where isnull(coalesce(a, b)) | fields a,b,c | head 3`
- `source = table | where isempty(a)`
- `source = table | where isblank(a)`
- `source = table | where case(length(a) > 6, 'True' else 'False') = 'True'`
- `source = table | where a between 1 and 4`- Nota: questo restituisce un >= 1 e un <= 4, cioè [1, 4]
- `source = table | where b not between '2024-09-10' and '2025-09-10'`- Nota: restituisce b >= '**********' e b <= '2025-09-10'
- `source = table | where cidrmatch(ip, '************/24')`
- `source = table | where cidrmatch(ipv6, '2003:db8::/32')`

```
source = table | eval status_category =
    case(a >= 200 AND a < 300, 'Success',
    a >= 300 AND a < 400, 'Redirection',
    a >= 400 AND a < 500, 'Client Error',
    a >= 500, 'Server Error'
    else 'Incorrect HTTP status code')
    | where case(a >= 200 AND a < 300, 'Success',
    a >= 300 AND a < 400, 'Redirection',
    a >= 400 AND a < 500, 'Client Error',
    a >= 500, 'Server Error'
    else 'Incorrect HTTP status code'
    ) = 'Incorrect HTTP status code'
```

```
source = table
    | eval factor = case(a > 15, a - 14, isnull(b), a - 7, a < 3, a + 1 else 1)
    | where case(factor = 2, 'even', factor = 4, 'even', factor = 6, 'even', factor =
 8, 'even' else 'odd') = 'even'
    |  stats count() by factor
```

riepilogo del campo

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questo comando PPL, vedi. the section called “Comandi”

Utilizzate il `fieldsummary` comando per calcolare le statistiche di base per ogni campo (count, distinct count, min, max, avg, stddev, mean) e determinare il tipo di dati di ogni campo. Questo comando può essere usato con qualsiasi pipe precedente e ne terrà conto.

Sintassi

Utilizzare la seguente sintassi. Per i casi d'uso di CloudWatch Logs, è supportato solo un campo in una query.

```
... | fieldsummary <field-list> (nulls=true/false)
```

includere campi

- Elenco di tutte le colonne da raccogliere con le statistiche in un set di risultati unificato.

Nulls

- Opzionale.
- Se impostato su true, includi i valori nulli nei calcoli di aggregazione (sostituisci null con zero per i valori numerici).

Esempio 1

Interrogazione PPL:

```
os> source = t | where status_code != 200 | fieldsummary includefields= status_code
 nulls=true
+------------------+-------------+------------+------------+------------+------------
+------------+------------+----------------|
| Fields           | COUNT       | COUNT_DISTINCT    |  MIN  |  MAX   |  AVG   |  MEAN
  |        STDDEV       | NUlls | TYPEOF |
```

```
|-----------------+-------------+-----------+-----------+-----------+-----------
+-----------+-----------+----------------|
| "status_code"    |      2      |         2         | 301   |   403   |  352.0 |  352.0
  |  72.12489168102785  |  0    | "int"  |
+-----------------+-------------+-----------+-----------+-----------+-----------
+-----------+-----------+----------------|
```

Esempio 2

Interrogazione PPL:

```
os> source = t | fieldsummary includefields= id, status_code, request_path nulls=true
+-----------------+-------------+-----------+-----------+-----------+-----------
+-----------+-----------+----------------|
| Fields           | COUNT       | COUNT_DISTINCT    |  MIN  |  MAX   |  AVG   |  MEAN
  |        STDDEV       | NUlls | TYPEOF |
|-----------------+-------------+-----------+-----------+-----------+-----------
+-----------+-----------+----------------|
|       "id"       |      6      |         6         | 1     |   6    |  3.5   |   3.5
 |  1.8708286933869707  |  0    | "int"  |
+-----------------+-------------+-----------+-----------+-----------+-----------
+-----------+-----------+----------------|
| "status_code"    |      4      |         3         | 200   |   403   |  184.0 |  184.0
  |  161.16699413961905 |  2    | "int"  |
+-----------------+-------------+-----------+-----------+-----------+-----------
+-----------+-----------+----------------|
| "request_path"   |      2      |         2         | /about| /home   |  0.0    |  0.0
    |      0           |  2    |"string"|
+-----------------+-------------+-----------+-----------+-----------+-----------
+-----------+-----------+----------------|
```

comando expand

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

Utilizzate il `expand` comando per appiattire un campo di tipo Array <Any>o Map<Any>, producendo righe singole per ogni elemento o coppia chiave-valore.

Sintassi

Utilizzare la seguente sintassi:

```
expand <field> [As alias]
```

field

- Il campo da espandere (esplodere).
- Il campo deve essere di un tipo supportato.

alias

- Opzionale.
- Il nome da utilizzare al posto del nome del campo originale.

Linee guida per l'utilizzo

Il comando expand produce una riga per ogni elemento nell'array o nel campo della mappa specificato, dove:

- Gli elementi dell'array diventano righe singole.
- Le coppie chiave-valore della mappa sono suddivise in righe separate, con ogni valore chiave rappresentato come una riga.
- Quando viene fornito un alias, i valori esplosi vengono rappresentati sotto l'alias anziché nel nome del campo originale.

È possibile utilizzare questo comando in combinazione con altri comandi, come stats, eval e parse, per manipolare o estrarre i dati dopo l'espansione.

Esempi

- `source = table | expand employee | stats max(salary) as max by state, company`
- `source = table | expand employee as worker | stats max(salary) as max by state, company`

- `source = table | expand employee as worker | eval bonus = salary * 3 | fields worker, bonus`
- `source = table | expand employee | parse description '(?<email>.+@.+)' | fields employee, email`
- `source = table | eval array=json_array(1, 2, 3) | expand array as uid | fields name, occupation, uid`
- `source = table | expand multi_valueA as multiA | expand multi_valueB as multiB`

È possibile utilizzare il comando expand in combinazione con altri comandi come eval, stats e altri. L'utilizzo di più comandi di espansione creerà un prodotto cartesiano di tutti gli elementi interni all'interno di ogni matrice o mappa composita.

Query push-down SQL efficace

Il comando expand viene tradotto in un'operazione SQL equivalente utilizzando LATERAL VIEW explode, che consente l'esplosione efficiente di array o mappe a livello di query SQL.

```
SELECT customer exploded_productId
FROM table
LATERAL VIEW explode(productId) AS exploded_productId
```

Il comando explode offre le seguenti funzionalità:

- È un'operazione su colonna che restituisce una nuova colonna.
- Crea una nuova riga per ogni elemento nella colonna esplosa.
- I valori null interni vengono ignorati come parte del campo esploso (nessuna riga indica null). created/exploded

Funzioni PPL

Argomenti

- Funzioni di condizione PPL
- funzioni hash crittografiche PPL
- Funzioni di data e ora PPL
- espressioni PPL

- Funzioni degli indirizzi IP PPL
- Funzioni PPL JSON
- Funzioni PPL Lambda
- Funzioni matematiche PPL
- Funzioni di stringa PPL
- Funzioni di conversione di tipo PPL

Funzioni di condizione PPL

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

ISNULL

Descrizione: `isnull(field)` restituisce true se il campo è nullo.

Tipo di argomento:

- Tutti i tipi di dati supportati.

Tipo di restituzione:

- BOOLEAN

Esempio:

```
os> source=accounts | eval result = isnull(employer) | fields result, employer,
 firstname
fetched rows / total rows = 4/4
+----------+-------------+-------------+
| result   | employer    | firstname   |
|----------+-------------+-------------|
| False    | AnyCompany  | Mary        |
| False    | ExampleCorp | Jane        |
```

```
| False    | ExampleOrg  | Nikki       |
| True     | null        | Juan        |
+----------+-------------+-------------+
```

NON È NULLO

Descrizione: `isnotnull(field)` restituisce true se il campo non è nullo.

Tipo di argomento:

- Tutti i tipi di dati supportati.

Tipo di restituzione:

- BOOLEAN

Esempio:

```
os> source=accounts | where not isnotnull(employer) | fields account_number, employer
fetched rows / total rows = 1/1
+------------------+------------+
| account_number   | employer   |
|------------------+------------|
| 18               | null       |
+------------------+------------+
```

EXISTS

Esempio:

```
os> source=accounts | where exists(email) | fields account_number, email
fetched rows / total rows = 1/1
```

SE NULL

Descrizione: `ifnull(field1, field2)` restituisce `field2` se `field1` è nullo.

Tipo di argomento:

- Tutti i tipi di dati supportati.

- Se i due parametri hanno tipi diversi, la funzione non supererà il controllo semantico.

Tipo di restituzione:

- Qualsiasi

Esempio:

```
os> source=accounts | eval result = ifnull(employer, 'default') | fields result,
 employer, firstname
fetched rows / total rows = 4/4
+------------+------------+-------------+
| result     | employer   | firstname   |
|------------+------------+-------------|
| AnyCompany | AnyCompany | Mary        |
| ExampleCorp| ExampleCorp| Jane        |
| ExampleOrg | ExampleOrg | Nikki       |
| default    | null       | Juan        |
+------------+------------+-------------+
```

NULLIF

Descrizione: `nullif(field1, field2)` restituisce null se due parametri sono uguali, altrimenti restituisce field1.

Tipo di argomento:

- Tutti i tipi di dati supportati.
- Se i due parametri hanno tipi diversi, la funzione non supererà il controllo semantico.

Tipo di restituzione:

- Qualsiasi

Esempio:

```
os> source=accounts | eval result = nullif(employer, 'AnyCompany') | fields result,
 employer, firstname
fetched rows / total rows = 4/4
```

```
+----------------+----------------+-------------+
| result         | employer       | firstname   |
|----------------+----------------+-------------|
| null           | AnyCompany     | Mary        |
| ExampleCorp    | ExampleCorp    | Jane        |
| ExampleOrg     | ExampleOrg     | Nikki       |
| null           | null           | Juan        |
+----------------+----------------+-------------+
```

IF

Descrizione: `if(condition, expr1, expr2)` restituisce `expr1` se la condizione è vera, altrimenti restituisce`expr2`.

Tipo di argomento:

- Tutti i tipi di dati supportati.
- Se i due parametri hanno tipi diversi, la funzione non supererà il controllo semantico.

Tipo di restituzione:

- Qualsiasi

Esempio:

```
os> source=accounts | eval result = if(true, firstname, lastname) | fields result,
 firstname, lastname
fetched rows / total rows = 4/4
+----------+-------------+----------+
| result   | firstname | lastname   |
|----------+-------------+----------|
| Jane     | Jane      | Doe        |
| Mary     | Mary      | Major      |
| Pat      | Pat       | Candella   |
| Dale     | Jorge     | Souza      |
+----------+-----------+------------+

os> source=accounts | eval result = if(false, firstname, lastname) | fields result,
 firstname, lastname
fetched rows / total rows = 4/4
+----------+-------------+------------+
```

```
| result   | firstname   | lastname   |
|----------+-------------+------------|
| Doe      | Jane        | Doe        |
| Major    | Mary        | Major      |
| Candella | Pat         | Candella   |
| Souza    | Jorge       | Souza      |
+----------+-------------+------------+

os> source=accounts | eval is_vip = if(age > 30 AND isnotnull(employer), true, false) |
 fields is_vip, firstname, lastname
fetched rows / total rows = 4/4
+----------+-------------+------------+
| is_vip   | firstname   | lastname   |
|----------+-------------+------------|
| True     | Jane        | Doe        |
| True     | Mary        | Major      |
| False    | Pat         | Candella   |
| False    | Jorge       | Souza      |
+----------+-------------+------------+
```

funzioni hash crittografiche PPL

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

MD5

MD5 calcola il MD5 digest e restituisce il valore come stringa esadecimale di 32 caratteri.

Utilizzo: `md5('hello')`

Tipo di argomento:

- STRING

Tipo di restituzione:

- STRING

Esempio:

```
os> source=people | eval `MD5('hello')` = MD5('hello') | fields `MD5('hello')`
fetched rows / total rows = 1/1
+----------------------------------+
| MD5('hello')                     |
|----------------------------------|
| <32 character hex string>        |
+----------------------------------+
```

SHA1

SHA1 restituisce il risultato della stringa esadecimale di SHA-1.

Utilizzo: `sha1('hello')`

Tipo di argomento:

- STRING

Tipo di restituzione:

- STRING

Esempio:

```
os> source=people | eval `SHA1('hello')` = SHA1('hello') | fields `SHA1('hello')`
fetched rows / total rows = 1/1
+------------------------------------------+
| SHA1('hello')                            |
|------------------------------------------|
| <40-character SHA-1 hash result>         |
+------------------------------------------+
```

SHA2

SHA2 restituisce il risultato della stringa esadecimale della famiglia di funzioni hash SHA-2 (SHA-224, SHA-256, SHA-384 e SHA-512). NumBits indica la lunghezza di bit desiderata del risultato, che deve avere un valore di 224, 256, 384, 512

Utilizzo:

- `sha2('hello',256)`
- `sha2('hello',512)`

Tipo di argomento:

- STRINGA, NUMERO INTERO

Tipo di restituzione:

- STRING

Esempio:

```
os> source=people | eval `SHA2('hello',256)` = SHA2('hello',256) | fields
 `SHA2('hello',256)`
fetched rows / total rows = 1/1
+------------------------------------------------------------------+
| SHA2('hello',256)                                                |
|------------------------------------------------------------------|
| <64-character SHA-256 hash result>                               |
+------------------------------------------------------------------+

os> source=people | eval `SHA2('hello',512)` = SHA2('hello',512) | fields
 `SHA2('hello',512)`
fetched rows / total rows = 1/1
+------------------------------------------------------------------+
| SHA2('hello',512)                                                |
                                          |
|------------------------------------------------------------------|
| <128-character SHA-512 hash result>                              |
+------------------------------------------------------------------+
```

Funzioni di data e ora PPL

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

DAY

Utilizzo: `DAY(date)` estrae il giorno del mese per una data compresa tra 1 e 31.

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi:`DAYOFMONTH`, `DAY_OF_MONTH`

Esempio:

```
os> source=people | eval `DAY(DATE('2020-08-26'))` = DAY(DATE('2020-08-26')) | fields
 `DAY(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+---------------------------+
| DAY(DATE('2020-08-26'))   |
|---------------------------|
| 26                        |
+---------------------------+
```

DAYOFMONTH

Utilizzo: `DAYOFMONTH(date)` estrae il giorno del mese per una data compresa tra 1 e 31.

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi:`DAY`, `DAY_OF_MONTH`

Esempio:

```
os> source=people | eval `DAYOFMONTH(DATE('2020-08-26'))` =
 DAYOFMONTH(DATE('2020-08-26')) | fields `DAYOFMONTH(DATE('2020-08-26'))`
```

```
fetched rows / total rows = 1/1
+---------------------------------+
| DAYOFMONTH(DATE('2020-08-26'))   |
|---------------------------------|
| 26                              |
+---------------------------------+
```

DAY_OF_MONTH

Utilizzo: DAY_OF_MONTH(DATE) estrae il giorno del mese per una data compresa tra 1 e 31.

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi:DAY, DAYOFMONTH

Esempio:

```
os> source=people | eval `DAY_OF_MONTH(DATE('2020-08-26'))` =
 DAY_OF_MONTH(DATE('2020-08-26')) | fields `DAY_OF_MONTH(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+-----------------------------------+
| DAY_OF_MONTH(DATE('2020-08-26'))   |
|-----------------------------------|
| 26                                |
+-----------------------------------+
```

DAYOFWEEK

Utilizzo: DAYOFWEEK(DATE) restituisce l'indice dei giorni feriali per una data (1 = domenica, 2 = lunedì,..., 7 = sabato).

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: DAY_OF_WEEK

Esempio:

```
os> source=people | eval `DAYOFWEEK(DATE('2020-08-26'))` =
 DAYOFWEEK(DATE('2020-08-26')) | fields `DAYOFWEEK(DATE('2020-08-26'))`
```

```
fetched rows / total rows = 1/1
+---------------------------------+
| DAYOFWEEK(DATE('2020-08-26'))   |
|---------------------------------|
| 4                               |
+---------------------------------+
```

DAY_OF_WEEK

Utilizzo: `DAY_OF_WEEK(DATE)` restituisce l'indice dei giorni feriali per una data (1 = domenica, 2 = lunedì,..., 7 = sabato).

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `DAYOFWEEK`

Esempio:

```
os> source=people | eval `DAY_OF_WEEK(DATE('2020-08-26'))` =
 DAY_OF_WEEK(DATE('2020-08-26')) | fields `DAY_OF_WEEK(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+----------------------------------+
| DAY_OF_WEEK(DATE('2020-08-26'))   |
|----------------------------------|
| 4                                |
+----------------------------------+
```

DAYOFYEAR

Utilizzo: `DAYOFYEAR(DATE)` restituisce il giorno dell'anno per una data compresa tra 1 e 366.

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `DAY_OF_YEAR`

Esempio:

```
os> source=people | eval `DAYOFYEAR(DATE('2020-08-26'))` =
 DAYOFYEAR(DATE('2020-08-26')) | fields `DAYOFYEAR(DATE('2020-08-26'))`
```

```
fetched rows / total rows = 1/1
+---------------------------------+
| DAYOFYEAR(DATE('2020-08-26'))   |
|---------------------------------|
| 239                             |
+---------------------------------+
```

DAY_OF_YEAR

Utilizzo: DAY_OF_YEAR(DATE) restituisce il giorno dell'anno per una data compresa tra 1 e 366.

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: DAYOFYEAR

Esempio:

```
os> source=people | eval `DAY_OF_YEAR(DATE('2020-08-26'))` =
 DAY_OF_YEAR(DATE('2020-08-26')) | fields `DAY_OF_YEAR(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+----------------------------------+
| DAY_OF_YEAR(DATE('2020-08-26'))   |
|----------------------------------|
| 239                              |
+----------------------------------+
```

DAYNAME

Utilizzo: DAYNAME(DATE) restituisce il nome del giorno della settimana per una data, tra cui lunedì, martedì, mercoledì, giovedì, venerdì, sabato e domenica.

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: STRING

Esempio:

```
os> source=people | eval `DAYNAME(DATE('2020-08-26'))` = DAYNAME(DATE('2020-08-26')) |
 fields `DAYNAME(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
```

```
+------------------------------+
| DAYNAME(DATE('2020-08-26'))  |
|------------------------------|
| Wednesday                    |
+------------------------------+
```

FROM_UNIXTIME

Utilizzo: FROM_UNIXTIME restituisce una rappresentazione dell'argomento fornito come timestamp o valore di stringa di caratteri. Questa funzione esegue una conversione inversa della UNIX_TIMESTAMP funzione.

Se fornite un secondo argomento, lo FROM_UNIXTIME utilizza per formattare il risultato in modo simile alla DATE_FORMAT funzione.

Se il timestamp non rientra nell'intervallo dal 1970-01-01 00:00:00 al 3001-01-18 23:59:59.999 999 (da 0 a 32536771199.999999 epoca), la funzione restituisce. NULL

Tipo di argomento: DOUBLE, STRING

Mappa del tipo di ritorno:

DOUBLE -> TIMESTAMP

DOUBLE, STRING -> STRING

Esempi:

```
os> source=people | eval `FROM_UNIXTIME(1220249547)` = FROM_UNIXTIME(1220249547) |
 fields `FROM_UNIXTIME(1220249547)`
fetched rows / total rows = 1/1
+------------------------------+
| FROM_UNIXTIME(1220249547)    |
|------------------------------|
| 2008-09-01 06:12:27          |
+------------------------------+

os> source=people | eval `FROM_UNIXTIME(1220249547, 'HH:mm:ss')` =
 FROM_UNIXTIME(1220249547, 'HH:mm:ss') | fields `FROM_UNIXTIME(1220249547, 'HH:mm:ss')`
fetched rows / total rows = 1/1
+---------------------------------------+
| FROM_UNIXTIME(1220249547, 'HH:mm:ss')   |
|---------------------------------------|
```

```
| 06:12:27                              |
+---------------------------------------+
```

HOUR

Utilizzo: `HOUR(TIME)` estrae il valore dell'ora in base al tempo.

A differenza di un'ora del giorno standard, il valore dell'ora in questa funzione può avere un intervallo maggiore di 23. Di conseguenza, il valore restituito da `HOUR(TIME)` può essere maggiore di 23.

Tipo di argomento: STRING/TIME/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `HOUR_OF_DAY`

Esempio:

```
os> source=people | eval `HOUR(TIME('01:02:03'))` = HOUR(TIME('01:02:03')) | fields
 `HOUR(TIME('01:02:03'))`
fetched rows / total rows = 1/1
+--------------------------+
| HOUR(TIME('01:02:03'))   |
|--------------------------|
| 1                        |
+--------------------------+
```

HOUR_OF_DAY

Utilizzo: `HOUR_OF_DAY(TIME)` estrae il valore dell'ora dall'ora specificata.

A differenza di un'ora del giorno standard, il valore dell'ora in questa funzione può avere un intervallo maggiore di 23. Di conseguenza, il valore restituito da `HOUR_OF_DAY(TIME)` può essere maggiore di 23.

Tipo di argomento: STRING/TIME/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `HOUR`

Esempio:

```
os> source=people | eval `HOUR_OF_DAY(TIME('01:02:03'))` =
 HOUR_OF_DAY(TIME('01:02:03')) | fields `HOUR_OF_DAY(TIME('01:02:03'))`
fetched rows / total rows = 1/1
+--------------------------------+
| HOUR_OF_DAY(TIME('01:02:03'))  |
|--------------------------------|
| 1                              |
+--------------------------------+
```

LAST_DAY

Utilizzo: LAST_DAY restituisce l'ultimo giorno del mese come valore DATE per l'argomento data specificato.

Tipo di argomento: DATE/STRING/TIMESTAMP/TIME

Tipo di restituzione: DATA

Esempio:

```
os> source=people | eval `last_day('2023-02-06')` = last_day('2023-02-06') | fields
 `last_day('2023-02-06')`
fetched rows / total rows = 1/1
+-------------------------+
| last_day('2023-02-06')  |
|-------------------------|
| 2023-02-28              |
+-------------------------+
```

LOCALTIMESTAMP

Utilizzo: LOCALTIMESTAMP() è sinonimo diNOW().

Esempio:

```
> source=people | eval `LOCALTIMESTAMP()` = LOCALTIMESTAMP() | fields
 `LOCALTIMESTAMP()`
fetched rows / total rows = 1/1
+--------------------+
| LOCALTIMESTAMP()   |
|--------------------|
| 2022-08-02 15:54:19 |
```

```
+---------------------+
```

LOCALTIME

Utilizzo: `LOCALTIME()` è sinonimo di`NOW()`.

Esempio:

```
> source=people | eval `LOCALTIME()` = LOCALTIME() | fields `LOCALTIME()`
fetched rows / total rows = 1/1
+---------------------+
| LOCALTIME()         |
|---------------------|
| 2022-08-02 15:54:19 |
+---------------------+
```

MAKE_DATE

Utilizzo: `MAKE_DATE` restituisce un valore di data basato sui valori di anno, mese e giorno specificati. Tutti gli argomenti sono arrotondati a numeri interi.

Specifiche: 1. MAKE_DATE (INTERO, INTERO, INTERO) -> DATA

Tipo di argomento: INTEGER, INTEGER, INTEGER

Tipo di restituzione: DATE

Esempio:

```
os> source=people | eval `MAKE_DATE(1945, 5, 9)` = MAKEDATE(1945, 5, 9) | fields
 `MAKEDATE(1945, 5, 9)`
fetched rows / total rows = 1/1
+------------------------+
| MAKEDATE(1945, 5, 9)   |
|------------------------|
| 1945-05-09             |
+------------------------+
```

MINUTE

Utilizzo: `MINUTE(TIME)` restituisce la componente in minuti dell'ora specificata, come numero intero compreso tra 0 e 59.

Tipo di argomento: STRING/TIME/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `MINUTE_OF_HOUR`

Esempio:

```
os> source=people | eval `MINUTE(TIME('01:02:03'))` =  MINUTE(TIME('01:02:03')) |
 fields `MINUTE(TIME('01:02:03'))`
fetched rows / total rows = 1/1
+----------------------------+
| MINUTE(TIME('01:02:03'))   |
|----------------------------|
| 2                          |
+----------------------------+
```

`MINUTE_OF_HOUR`

Utilizzo: `MINUTE_OF_HOUR(TIME)` restituisce il componente minuto del tempo specificato, come numero intero compreso tra 0 e 59.

Tipo di argomento: STRING/TIME/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `MINUTE`

Esempio:

```
os> source=people | eval `MINUTE_OF_HOUR(TIME('01:02:03'))` =
 MINUTE_OF_HOUR(TIME('01:02:03')) | fields `MINUTE_OF_HOUR(TIME('01:02:03'))`
fetched rows / total rows = 1/1
+-----------------------------------+
| MINUTE_OF_HOUR(TIME('01:02:03'))   |
|-----------------------------------|
| 2                                 |
+-----------------------------------+
```

`MONTH`

Utilizzo: `MONTH(DATE)` restituisce il mese della data specificata come numero intero, nell'intervallo da 1 a 12 (dove 1 rappresenta gennaio e 12 rappresenta dicembre).

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `MONTH_OF_YEAR`

Esempio:

```
os> source=people | eval `MONTH(DATE('2020-08-26'))` =  MONTH(DATE('2020-08-26')) |
 fields `MONTH(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+-----------------------------+
| MONTH(DATE('2020-08-26'))   |
|-----------------------------|
| 8                           |
+-----------------------------+
```

MONTHNAME

Utilizzo: `MONTHNAME(DATE)` restituisce il mese della data specificata come numero intero, nell'intervallo da 1 a 12 (dove 1 rappresenta gennaio e 12 rappresenta dicembre).

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `MONTH_OF_YEAR`

Esempio:

```
os> source=people | eval `MONTHNAME(DATE('2020-08-26'))` =
 MONTHNAME(DATE('2020-08-26')) | fields `MONTHNAME(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+---------------------------------+
| MONTHNAME(DATE('2020-08-26'))   |
|---------------------------------|
| August                          |
+---------------------------------+
```

MONTH_OF_YEAR

Utilizzo: `MONTH_OF_YEAR(DATE)` restituisce il mese della data specificata come numero intero, nell'intervallo da 1 a 12 (dove 1 rappresenta gennaio e 12 rappresenta dicembre).

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `MONTH`

Esempio:

```
os> source=people | eval `MONTH_OF_YEAR(DATE('2020-08-26'))` =
 MONTH_OF_YEAR(DATE('2020-08-26')) | fields `MONTH_OF_YEAR(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+------------------------------------+
| MONTH_OF_YEAR(DATE('2020-08-26'))   |
|------------------------------------|
| 8                                  |
+------------------------------------+
```

`NOW`

Utilizzo: `NOW` restituisce la data e l'ora correnti come `TIMESTAMP` valore nel formato YYYY-MM-DD 'hh:mm:ss'. Il valore è espresso nel fuso orario del cluster.

> **Note**
>
> `NOW()` restituisce un tempo costante che indica quando l'istruzione ha iniziato l'esecuzione. Questo differisce da`SYSDATE()`, che restituisce l'ora esatta di esecuzione.

Tipo di restituzione: TIMESTAMP

Specificazione: NOW () -> TIMESTAMP

Esempio:

```
os> source=people | eval `value_1` = NOW(), `value_2` = NOW() | fields `value_1`,
 `value_2`
fetched rows / total rows = 1/1
+---------------------+---------------------+
| value_1             | value_2             |
|---------------------+---------------------|
| 2022-08-02 15:39:05 | 2022-08-02 15:39:05 |
```

```
+--------------------+---------------------+
```

QUARTER

Utilizzo: `QUARTER(DATE)` restituisce il trimestre dell'anno per la data specificata come numero intero, nell'intervallo da 1 a 4.

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Esempio:

```
os> source=people | eval `QUARTER(DATE('2020-08-26'))` = QUARTER(DATE('2020-08-26')) |
 fields `QUARTER(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+-------------------------------+
| QUARTER(DATE('2020-08-26'))   |
|-------------------------------|
| 3                             |
+-------------------------------+
```

SECOND

Utilizzo: `SECOND(TIME)` restituisce il secondo componente dell'ora specificata come numero intero, nell'intervallo da 0 a 59.

Tipo di argomento: STRING/TIME/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `SECOND_OF_MINUTE`

Esempio:

```
os> source=people | eval `SECOND(TIME('01:02:03'))` = SECOND(TIME('01:02:03')) | fields
 `SECOND(TIME('01:02:03'))`
fetched rows / total rows = 1/1
+---------------------------+
| SECOND(TIME('01:02:03'))  |
|---------------------------|
| 3                         |
```

```
+---------------------------+
```

SECOND_OF_MINUTE

Utilizzo: `SECOND_OF_MINUTE(TIME)` restituisce il secondo componente del tempo specificato come numero intero, nell'intervallo da 0 a 59.

Tipo di argomento: STRING/TIME/TIMESTAMP

Tipo restituito: INTEGER

Sinonimi: `SECOND`

Esempio:

```
os> source=people | eval `SECOND_OF_MINUTE(TIME('01:02:03'))` =
 SECOND_OF_MINUTE(TIME('01:02:03')) | fields `SECOND_OF_MINUTE(TIME('01:02:03'))`
fetched rows / total rows = 1/1
+-------------------------------------+
| SECOND_OF_MINUTE(TIME('01:02:03'))  |
|-------------------------------------|
| 3                                   |
+-------------------------------------+
```

SUBDATE

Utilizzo: `SUBDATE(DATE, DAYS)` sottrae il secondo argomento (ad esempio `DATE` o`DAYS`) dalla data specificata.

Tipo di argomento: DATE/TIMESTAMP, LONG

Mappa del tipo di ritorno: (DATE, LONG) -> DATE

Contrari: `ADDDATE`

Esempio:

```
os> source=people | eval `'2008-01-02' - 31d` = SUBDATE(DATE('2008-01-02'), 31),
 `'2020-08-26' - 1` = SUBDATE(DATE('2020-08-26'), 1), `ts '2020-08-26 01:01:01' -
 1` = SUBDATE(TIMESTAMP('2020-08-26 01:01:01'), 1) | fields `'2008-01-02' - 31d`,
 `'2020-08-26' - 1`, `ts '2020-08-26 01:01:01' - 1`
fetched rows / total rows = 1/1
```

```
+----------------------+--------------------+--------------------------------+
| '2008-01-02' - 31d   | '2020-08-26' - 1   | ts '2020-08-26 01:01:01' - 1   |
|----------------------+--------------------+--------------------------------|
| 2007-12-02 00:00:00  | 2020-08-25         | 2020-08-25 01:01:01            |
+----------------------+--------------------+--------------------------------+
```

SYSDATE

Utilizzo: `SYSDATE()` restituisce la data e l'ora correnti come `TIMESTAMP` valore nel formato 'YYYY-MM-DD hh:mm:ss.nnnnnn'.

`SYSDATE()`restituisce l'ora esatta in cui viene eseguito. Ciò differisce da NOW (), che restituisce un tempo costante che indica quando l'istruzione ha iniziato l'esecuzione.

Tipo di argomento opzionale: INTEGER (da 0 a 6): specifica il numero di cifre per frazioni di secondo nel valore restituito.

Tipo restituito: TIMESTAMP

Esempio:

```
os> source=people | eval `SYSDATE()` = SYSDATE() | fields `SYSDATE()`
fetched rows / total rows = 1/1
+----------------------------+
| SYSDATE()                  |
|----------------------------|
| 2022-08-02 15:39:05.123456 |
+----------------------------+
```

TIMESTAMP

Utilizzo: `TIMESTAMP(EXPR)` costruisce un tipo di timestamp con la stringa di input come timestamp. `expr`

Con un solo argomento, `TIMESTAMP(expr)` costruisce un timestamp dall'input. Se `expr` è una stringa, viene interpretata come un timestamp. Per argomenti non di tipo stringa, la funzione esegue il cast su un timestamp `expr` utilizzando il fuso orario UTC. Quando `expr` è un `TIME` valore, la funzione applica la data odierna prima del casting.

Se utilizzata con due argomenti, `TIMESTAMP(expr1, expr2)` aggiunge l'espressione time (`expr2`) all'espressione date o timestamp (`expr1`) e restituisce il risultato come valore di timestamp.

Tipo di argomento: STRING/DATE/TIME/TIMESTAMP

Mappa del tipo di ritorno:

(STRING/DATE/TIME/TIMESTAMP) -> TIMESTAMP

() -> STRING/DATE/TIME/TIMESTAMP, STRING/DATE/TIME/TIMESTAMP TIMESTAMP

Esempio:

```
os> source=people | eval `TIMESTAMP('2020-08-26 13:49:00')` = TIMESTAMP('2020-08-26
 13:49:00'), `TIMESTAMP('2020-08-26 13:49:00', TIME('12:15:42'))` =
 TIMESTAMP('2020-08-26 13:49:00', TIME('12:15:42')) | fields `TIMESTAMP('2020-08-26
 13:49:00')`, `TIMESTAMP('2020-08-26 13:49:00', TIME('12:15:42'))`
fetched rows / total rows = 1/1
+-----------------------------------
+-----------------------------------------------------+
| TIMESTAMP('2020-08-26 13:49:00')   | TIMESTAMP('2020-08-26 13:49:00',
 TIME('12:15:42'))   |
|-----------------------------------
+-----------------------------------------------------|
| 2020-08-26 13:49:00               | 2020-08-27 02:04:42
     |
+-----------------------------------
+-----------------------------------------------------+
```

UNIX_TIMESTAMP

Utilizzo: `UNIX_TIMESTAMP` converte un determinato argomento di data in ora Unix (secondi dall'Epoca, iniziata all'inizio del 1970). Se non viene fornito alcun argomento, restituisce l'ora Unix corrente.

L'argomento della data può essere un`DATE`, una `TIMESTAMP` stringa o un numero in uno dei seguenti formati:`YYMMDD`, `YYMMDDhhmmss``YYYYMMDD`, o`YYYYMMDDhhmmss`. Se l'argomento include un componente temporale, può facoltativamente includere frazioni di secondi.

Se l'argomento è in un formato non valido o non rientra nell'intervallo compreso tra 00:00:00 e 3001-01-18 23:59:59.999 (da 0 a 32536771199.999999 nell'ora dell'epoca), la funzione restituisce. `NULL`

La `DOUBLE` funzione accetta o come tipi di argomento `DATE` o nessun argomento. `TIMESTAMP` Restituisce sempre un `DOUBLE` valore che rappresenta il timestamp Unix.

Per la conversione inversa, puoi usare la funzione FROM_UNIXTIME.

Tipo di argomento:/DOUBLE/DATE/TIMESTAMP<NONE>

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `UNIX_TIMESTAMP(double)` = UNIX_TIMESTAMP(20771122143845),
 `UNIX_TIMESTAMP(timestamp)` = UNIX_TIMESTAMP(TIMESTAMP('1996-11-15 17:05:42')) |
 fields `UNIX_TIMESTAMP(double)`, `UNIX_TIMESTAMP(timestamp)`
fetched rows / total rows = 1/1
+-------------------------+----------------------------+
| UNIX_TIMESTAMP(double)   | UNIX_TIMESTAMP(timestamp)   |
|-------------------------+----------------------------|
| 3404817525.0             | 848077542.0                |
+-------------------------+----------------------------+
```

WEEK

Utilizzo: `WEEK(DATE)` restituisce il numero della settimana per una determinata data.

Tipo di argomento: DATE/TIMESTAMP/STRING

Tipo restituito: INTEGER

Sinonimi: `WEEK_OF_YEAR`

Esempio:

```
os> source=people | eval `WEEK(DATE('2008-02-20'))` = WEEK(DATE('2008-02-20')) | fields
 `WEEK(DATE('2008-02-20'))`
fetched rows / total rows = 1/1
+---------------------------+
| WEEK(DATE('2008-02-20'))   |
|---------------------------|
| 8                          |
+---------------------------+
```

WEEKDAY

Utilizzo: `WEEKDAY(DATE)` restituisce l'indice dei giorni feriali per la data (0 = lunedì, 1 = martedì,..., 6 = domenica).

È simile alla `dayofweek` funzione, ma restituisce indici diversi per ogni giorno.

Tipo di argomento: STRING/DATE/TIME/TIMESTAMP

Tipo restituito: INTEGER

Esempio:

```
os> source=people | eval `weekday(DATE('2020-08-26'))` = weekday(DATE('2020-08-26'))
 | eval `weekday(DATE('2020-08-27'))` = weekday(DATE('2020-08-27')) | fields
 `weekday(DATE('2020-08-26'))`, `weekday(DATE('2020-08-27'))`
fetched rows / total rows = 1/1
+------------------------------+------------------------------+
| weekday(DATE('2020-08-26'))  | weekday(DATE('2020-08-27'))  |
|------------------------------+------------------------------|
| 2                            | 3                            |
+------------------------------+------------------------------+
```

WEEK_OF_YEAR

Utilizzo: `WEEK_OF_YEAR(DATE)` restituisce il numero della settimana per la data specificata.

Tipo di argomento: DATE/TIMESTAMP/STRING

Tipo restituito: INTEGER

Sinonimi: `WEEK`

Esempio:

```
os> source=people | eval `WEEK_OF_YEAR(DATE('2008-02-20'))` = WEEK(DATE('2008-02-20'))|
 fields `WEEK_OF_YEAR(DATE('2008-02-20'))`
fetched rows / total rows = 1/1
+-----------------------------------+
| WEEK_OF_YEAR(DATE('2008-02-20'))  |
|-----------------------------------|
| 8                                 |
+-----------------------------------+
```

YEAR

Utilizzo: `YEAR(DATE)` restituisce l'anno come data, nell'intervallo da 1000 a 9999 o 0 per la data «zero».

Tipo di argomento: STRING/DATE/TIMESTAMP

Tipo restituito: INTEGER

Esempio:

```
os> source=people | eval `YEAR(DATE('2020-08-26'))` = YEAR(DATE('2020-08-26')) | fields
 `YEAR(DATE('2020-08-26'))`
fetched rows / total rows = 1/1
+----------------------------+
| YEAR(DATE('2020-08-26'))   |
|----------------------------|
| 2020                       |
+----------------------------+
```

DATE_ADD

Utilizzo: `DATE_ADD(date, INTERVAL expr unit)` aggiunge l'intervallo specificato alla data specificata.

Tipo di argomento: DATE, INTERVAL

Tipo di restituzione: DATE

Contrari: `DATE_SUB`

Esempio:

```
os> source=people | eval `'2020-08-26' + 1d` = DATE_ADD(DATE('2020-08-26'), INTERVAL 1
 DAY) | fields `'2020-08-26' + 1d`
fetched rows / total rows = 1/1
+---------------------+
| '2020-08-26' + 1d   |
|---------------------|
| 2020-08-27          |
+---------------------+
```

DATE_SUB

Utilizzo: `DATE_SUB(date, INTERVAL expr unit)` sottrae l'intervallo expr dalla data.

Tipo di argomento: DATE, INTERVAL

Tipo di restituzione: DATE

Contrari: `DATE_ADD`

Esempio:

```
os> source=people | eval `'2008-01-02' - 31d` = DATE_SUB(DATE('2008-01-02'), INTERVAL
 31 DAY) | fields `'2008-01-02' - 31d`
fetched rows / total rows = 1/1
+---------------------+
| '2008-01-02' - 31d  |
|---------------------|
| 2007-12-02          |
+---------------------+
```

`TIMESTAMPADD`

Utilizzo: restituisce un `TIMESTAMP` valore dopo aver aggiunto un intervallo di tempo specificato a una determinata data.

Argomenti:

- INTERVALLO: INTERVALLO (SECONDO, MINUTO, ORA, GIORNO, SETTIMANA, MESE, TRIMESTRE, ANNO)
- numero intero: INTERO
- data: DATE, TIMESTAMP o STRING

Se fornisci un `STRING` come argomento per la data, formattalo come valido. `TIMESTAMP` La funzione converte automaticamente un `DATE` argomento in un`TIMESTAMP`.

Esempi:

```
os> source=people | eval `TIMESTAMPADD(DAY, 17, '2000-01-01 00:00:00')` =
 TIMESTAMPADD(DAY, 17, '2000-01-01 00:00:00') | eval `TIMESTAMPADD(QUARTER, -1,
 '2000-01-01 00:00:00')` = TIMESTAMPADD(QUARTER, -1, '2000-01-01 00:00:00') | fields
 `TIMESTAMPADD(DAY, 17, '2000-01-01 00:00:00')`, `TIMESTAMPADD(QUARTER, -1, '2000-01-01
 00:00:00')`
fetched rows / total rows = 1/1
```

```
+----------------------------------------------
+-------------------------------------------------+
| TIMESTAMPADD(DAY, 17, '2000-01-01 00:00:00') | TIMESTAMPADD(QUARTER, -1, '2000-01-01
 00:00:00') |
|----------------------------------------------
+-------------------------------------------------|
| 2000-01-18 00:00:00                          | 1999-10-01 00:00:00
           |
+----------------------------------------------
+-------------------------------------------------+
```

`TIMESTAMPDIFF`

Utilizzo: `TIMESTAMPDIFF(interval, start, end)` restituisce la differenza tra l'inizio e la fine date/times in unità di intervallo specificate.

Argomenti:

- INTERVALLO: INTERVALLO (SECONDO, MINUTO, ORA, GIORNO, SETTIMANA, MESE, TRIMESTRE, ANNO)
- inizio: DATE, TIMESTAMP o STRING
- fine: DATE, TIMESTAMP o STRING

La funzione converte automaticamente gli argomenti in modo appropriato. `TIMESTAMP` Formatta `STRING` gli argomenti come file validi`TIMESTAMP`.

Esempi:

```
os> source=people | eval `TIMESTAMPDIFF(YEAR, '1997-01-01 00:00:00', '2001-03-06
 00:00:00')` = TIMESTAMPDIFF(YEAR, '1997-01-01 00:00:00', '2001-03-06 00:00:00') |
 eval `TIMESTAMPDIFF(SECOND, timestamp('1997-01-01 00:00:23'), timestamp('1997-01-01
 00:00:00'))` = TIMESTAMPDIFF(SECOND, timestamp('1997-01-01 00:00:23'),
 timestamp('1997-01-01 00:00:00')) | fields `TIMESTAMPDIFF(YEAR, '1997-01-01 00:00:00',
 '2001-03-06 00:00:00')`, `TIMESTAMPDIFF(SECOND, timestamp('1997-01-01 00:00:23'),
 timestamp('1997-01-01 00:00:00'))`
fetched rows / total rows = 1/1
+-----------------------------------------------------------------
+-----------------------------------------------------------------------------------------
+
| TIMESTAMPDIFF(YEAR, '1997-01-01 00:00:00', '2001-03-06 00:00:00') |
 TIMESTAMPDIFF(SECOND, timestamp('1997-01-01 00:00:23'), timestamp('1997-01-01
 00:00:00')) |
```

```
|-------------------------------------------------------------------
+-------------------------------------------------------------------------------------------|
| 4                                                                 | -23
                                                                        |
+-------------------------------------------------------------------
+-------------------------------------------------------------------------------------------
+
```

UTC_TIMESTAMP

Utilizzo: UTC_TIMESTAMP restituisce il timestamp UTC corrente come valore in 'YYYY-MM-DD hh:mm:ss'.

Tipo di restituzione: TIMESTAMP

Specificazione: UTC_TIMESTAMP () -> TIMESTAMP

Esempio:

```
> source=people | eval `UTC_TIMESTAMP()` = UTC_TIMESTAMP() | fields `UTC_TIMESTAMP()`
fetched rows / total rows = 1/1
+---------------------+
| UTC_TIMESTAMP()     |
|---------------------|
| 2022-10-03 17:54:28 |
+---------------------+
```

CURRENT_TIMEZONE

Utilizzo: restituisce il fuso orario locale corrente. CURRENT_TIMEZONE

Tipo restituito: STRING

Esempio:

```
> source=people | eval `CURRENT_TIMEZONE()` = CURRENT_TIMEZONE() | fields
 `CURRENT_TIMEZONE()`
fetched rows / total rows = 1/1
+------------------------+
| CURRENT_TIMEZONE()     |
|------------------------|
| America/Chicago        |
+------------------------+
```

espressioni PPL

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

Le espressioni, in particolare le espressioni di valore, restituiscono un valore scalare. Le espressioni hanno tipi e forme diversi. Ad esempio, esistono valori letterali come espressioni atomiche ed espressioni aritmetiche, predicative e funzionali basate su di essi. È possibile utilizzare le espressioni in diverse clausole, ad esempio l'utilizzo di espressioni aritmetiche nei comandi and. `Filter` `Stats`

Operatori

Un'espressione aritmetica è un'espressione formata da valori letterali numerici e operatori aritmetici binari come segue:

1. `+`: Aggiungi.
2. `-`: Sottrai.
3. `*`: Moltiplica.
4. `/`: Dividi (per i numeri interi, il risultato è un numero intero con la parte frazionaria scartata)
5. `%`: Modulo (da utilizzare solo con numeri interi; il risultato è il resto della divisione)

Precedenza

Utilizzate le parentesi per controllare la precedenza degli operatori aritmetici. Altrimenti, gli operatori con precedenza più alta vengono eseguiti per primi.

Conversione del tipo

La conversione implicita dei tipi viene eseguita durante la ricerca delle firme degli operatori. Ad esempio, un numero intero corrisponde `+` a una firma `+(double,double)` che dà come risultato un numero reale. Questa regola si applica anche alle chiamate di funzione.

Esempio per diversi tipi di espressioni aritmetiche:

```
os> source=accounts | where age > (25 + 5) | fields age ;
```

```
fetched rows / total rows = 3/3
+-------+
| age   |
|-------|
| 32    |
| 36    |
| 33    |
+-------+
```

Operatori predicati

Un operatore predicato è un'espressione che risulta vera. Il confronto dei `NULL` valori `MISSING` e segue queste regole:

- Un `MISSING` valore è solo uguale a un `MISSING` valore ed è inferiore agli altri valori.
- Un `NULL` valore è uguale a un `NULL` valore, è maggiore di un `MISSING` valore, ma è minore di tutti gli altri valori.

Operatori

Operatori predicati

Nome	Description
>	Maggiore dell'operatore
>=	Operatore maggiore o uguale
<	Meno dell'operatore
!=	Operatore non uguale
<=	Operatore minore o uguale
=	Operatore uguale
LIKE	Corrispondenza semplice dei modelli
IN	Test del valore NULL
AND	Operatore AND

Nome	Description
OR	Operatore OR
XOR	Operatore XOR
NOT	Test del valore NOT NULL

Puoi confrontare le date e le ore. Quando si confrontano diversi tipi di datetime (ad esempio DATE eTIME), entrambi vengono convertiti in. DATETIME Le seguenti regole si applicano alla conversione:

- TIMEsi applica alla data odierna.
- DATEviene interpretato a mezzanotte.

Operatore di predicato di base

Esempio di operatori di confronto:

```
os> source=accounts | where age > 33 | fields age ;
fetched rows / total rows = 1/1
+-------+
| age   |
|-------|
| 36    |
+-------+
```

IN

Esempio del campo di test IN dell'operatore nelle liste di valori:

```
os> source=accounts | where age in (32, 33) | fields age ;
fetched rows / total rows = 2/2
+-------+
| age   |
|-------|
| 32    |
| 33    |
+-------+
```

OR

Esempio dell'0Roperatore:

```
os> source=accounts | where age = 32 OR age = 33 | fields age ;
fetched rows / total rows = 2/2
+-------+
| age   |
|-------|
| 32    |
| 33    |
+-------+
```

NOT

Esempio dell'N0Toperatore:

```
os> source=accounts | where age not in (32, 33) | fields age ;
fetched rows / total rows = 2/2
+-------+
| age   |
|-------|
| 36    |
| 28    |
+-------+
```

Funzioni degli indirizzi IP PPL

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

CIDRMATCH

Utilizzo: `CIDRMATCH(ip, cidr)` verifica se l'indirizzo IP specificato rientra nell'intervallo cidr specificato.

Tipo di argomento:

- STRINGA, STRINGA
- Tipo restituito: BOOLEAN

Esempio:

```
os> source=ips | where cidrmatch(ip, '***********/24') | fields ip
fetched rows / total rows = 1/1
+--------------+
| ip           |
|--------------|
| ***********  |
+--------------+

os> source=ipsv6 | where cidrmatch(ip, '2003:db8::/32') | fields ip
fetched rows / total rows = 1/1
+----------------------------------------+
| ip                                     |
|----------------------------------------|
| 2003:0db8:****:****:****:****:****:0000 |
+----------------------------------------+
```

> **Note**
>
> - `ip`può essere un indirizzo IPv4 o. IPv6
> - `cidr`può essere un IPv4 o un IPv6 blocco.
> - `ip`e `cidr` deve essere entrambi IPv4 o entrambi IPv6.
> - `ip`e `cidr` devono essere entrambi validi e non vuoti/non nulli.

Funzioni PPL JSON

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

`JSON`

Utilizzo: `json(value)` valuta se una stringa può essere analizzata come formato JSON. La funzione restituisce la stringa originale se è JSON valido o null se non è valida.

Tipo di argomento: STRING

Tipo restituito: STRING/NULL. Un'espressione STRING di un formato di oggetto JSON valido.

Esempi:

```
os> source=people | eval `valid_json()` = json('[1,2,3,{"f1":1,"f2":[5,6]},4]') |
 fields valid_json
fetched rows / total rows = 1/1
+-------------------------------+
| valid_json                    |
+-------------------------------+
| [1,2,3,{"f1":1,"f2":[5,6]},4]   |
+-------------------------------+

os> source=people | eval `invalid_json()` = json('{"invalid": "json"') | fields
 invalid_json
fetched rows / total rows = 1/1
+----------------+
| invalid_json   |
+----------------+
| null           |
+----------------+
```

JSON_OBJECT

Utilizzo: `json_object(<key>, <value>[, <key>, <value>]...)` restituisce un oggetto JSON dai membri di coppie chiave-valore.

Tipo di argomento:

- A <key>deve essere STRING.
- A <value>può essere qualsiasi tipo di dati.

Tipo restituito: JSON_OBJECT. Un' StructType espressione di un oggetto JSON valido.

Esempi:

```
os> source=people | eval result = json_object('key', 123.45) | fields result
fetched rows / total rows = 1/1
+------------------+
| result           |
+------------------+
```

```
| {"key":123.45}   |
+------------------+

os> source=people | eval result = json_object('outer', json_object('inner', 123.45)) |
 fields result
fetched rows / total rows = 1/1
+-----------------------------+
| result                      |
+-----------------------------+
| {"outer":{"inner":123.45}}  |
+-----------------------------+
```

JSON_ARRAY

Utilizzo: `json_array(<value>...)` crea un JSON ARRAY utilizzando un elenco di valori.

Tipo di argomento: A `<value>` può essere qualsiasi tipo di valore come stringa, numero o booleano.

Tipo restituito: ARRAY. Un array di qualsiasi tipo di dati supportato per un array JSON valido.

Esempi:

```
os> source=people | eval `json_array` = json_array(1, 2, 0, -1, 1.1, -0.11)
fetched rows / total rows = 1/1
+-----------------------------+
| json_array                  |
+-----------------------------+
| [1.0,2.0,0.0,-1.0,1.1,-0.11] |
+-----------------------------+

os> source=people | eval `json_array_object` = json_object("array", json_array(1, 2, 0,
 -1, 1.1, -0.11))
fetched rows / total rows = 1/1
+---------------------------------------+
| json_array_object                     |
+---------------------------------------+
| {"array":[1.0,2.0,0.0,-1.0,1.1,-0.11]} |
+---------------------------------------+
```

TO_JSON_STRING

Utilizzo: `to_json_string(jsonObject)` restituisce una stringa JSON con un determinato valore di oggetto json.

Tipo di argomento: JSON_OBJECT

Tipo restituito: STRING

Esempi:

```
os> source=people | eval `json_string` = to_json_string(json_array(1, 2, 0, -1, 1.1,
 -0.11)) | fields json_string
fetched rows / total rows = 1/1
+--------------------------------+
| json_string                    |
+--------------------------------+
| [1.0,2.0,0.0,-1.0,1.1,-0.11]   |
+--------------------------------+

os> source=people | eval `json_string` = to_json_string(json_object('key', 123.45)) |
 fields json_string
fetched rows / total rows = 1/1
+-----------------+
| json_string     |
+-----------------+
| {'key', 123.45} |
+-----------------+
```

ARRAY_LENGTH

Utilizzo: `array_length(jsonArray)` restituisce il numero di elementi nell'array più esterno.

Tipo di argomento: ARRAY. Un oggetto ARRAY o JSON_ARRAY.

Tipo restituito: INTEGER

Esempio:

```
os> source=people | eval `json_array` = json_array_length(json_array(1,2,3,4)),
 `empty_array` = json_array_length(json_array())
fetched rows / total rows = 1/1
+--------------+---------------+
| json_array   | empty_array   |
+--------------+---------------+
| 4            | 0             |
+--------------+---------------+
```

JSON_EXTRACT

Utilizzo: `json_extract(jsonStr, path)` estrae un oggetto JSON da una stringa JSON in base al percorso JSON specificato. La funzione restituisce null se la stringa JSON di input non è valida.

Tipo di argomento: STRING, STRING

Tipo restituito: STRING

- Un'espressione STRING di un formato di oggetto JSON valido.
- `NULL`viene restituito in caso di JSON non valido.

Esempi:

```
os> source=people | eval `json_extract('{"a":"b"}', '$.a')` = json_extract('{"a":"b"}',
 '$a')
fetched rows / total rows = 1/1
+----------------------------------+
| json_extract('{"a":"b"}', 'a')   |
+----------------------------------+
| b                                |
+----------------------------------+

os> source=people | eval `json_extract('{"a":[{"b":1},{"b":2}]}', '$.a[1].b')` =
 json_extract('{"a":[{"b":1},{"b":2}]}', '$.a[1].b')
fetched rows / total rows = 1/1
+-------------------------------------------------------------+
| json_extract('{"a":[{"b":1.0},{"b":2.0}]}', '$.a[1].b')   |
+-------------------------------------------------------------+
| 2.0                                                         |
+-------------------------------------------------------------+

os> source=people | eval `json_extract('{"a":[{"b":1},{"b":2}]}', '$.a[*].b')` =
 json_extract('{"a":[{"b":1},{"b":2}]}', '$.a[*].b')
fetched rows / total rows = 1/1
+-------------------------------------------------------------+
| json_extract('{"a":[{"b":1.0},{"b":2.0}]}', '$.a[*].b')   |
+-------------------------------------------------------------+
| [1.0,2.0]                                                   |
+-------------------------------------------------------------+

os> source=people | eval `invalid_json` = json_extract('{"invalid": "json"')
```

```
fetched rows / total rows = 1/1
+----------------+
| invalid_json   |
+----------------+
| null           |
+----------------+
```

JSON_KEYS

Utilizzo: `json_keys(jsonStr)` restituisce tutte le chiavi dell'oggetto JSON più esterno come array.

Tipo di argomento: STRING. Un'espressione STRING di un formato di oggetto JSON valido.

Tipo restituito: ARRAY [STRING]. La funzione restituisce NULL qualsiasi altra stringa JSON valida, una stringa vuota o un codice JSON non valido.

Esempi:

```
os> source=people | eval `keys` = json_keys('{"f1":"abc","f2":{"f3":"a","f4":"b"}}')
fetched rows / total rows = 1/1
+------------+
| keus       |
+------------+
| [f1, f2]   |
+------------+

os> source=people | eval `keys` = json_keys('[1,2,3,{"f1":1,"f2":[5,6]},4]')
fetched rows / total rows = 1/1
+--------+
| keys   |
+--------+
| null   |
+--------+
```

JSON_VALID

Utilizzo: `json_valid(jsonStr)` valuta se una stringa JSON utilizza una sintassi JSON valida e restituisce TRUE o FALSE.

Tipo di argomento: STRING

Tipo restituito: BOOLEAN

Esempi:

```
os> source=people | eval `valid_json` = json_valid('[1,2,3,4]'), `invalid_json` =
 json_valid('{"invalid": "json"') | feilds `valid_json`, `invalid_json`
fetched rows / total rows = 1/1
+--------------+----------------+
| valid_json   | invalid_json   |
+--------------+----------------+
| True         | False          |
+--------------+----------------+

os> source=accounts | where json_valid('[1,2,3,4]') and isnull(email) | fields
 account_number, email
fetched rows / total rows = 1/1
+------------------+---------+
| account_number   | email   |
|------------------+---------|
| 13               | null    |
+------------------+---------+
```

Funzioni PPL Lambda

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

`EXISTS`

Utilizzo: `exists(array, lambda)` valuta se un predicato Lambda è valido per uno o più elementi dell'array.

Tipo di argomento: ARRAY, LAMBDA

Tipo restituito: BOOLEAN. Restituisce `TRUE` se almeno un elemento dell'array soddisfa il predicato Lambda, altrimenti. `FALSE`

Esempi:

```
 os> source=people | eval array = json_array(1, -1, 2), result = exists(array, x -> x >
 0) | fields result
fetched rows / total rows = 1/1
```

```
+-----------+
| result    |
+-----------+
| true      |
+-----------+

os> source=people | eval array = json_array(-1, -3, -2), result = exists(array, x -> x
 > 0) | fields result
fetched rows / total rows = 1/1
+-----------+
| result    |
+-----------+
| false     |
+-----------+
```

FILTER

Utilizzo: `filter(array, lambda)` filtra l'array di input utilizzando la funzione Lambda specificata.

Tipo di argomento: ARRAY, LAMBDA

Tipo restituito: ARRAY. Un ARRAY che contiene tutti gli elementi dell'array di input che soddisfano il predicato lambda.

Esempi:

```
 os> source=people | eval array = json_array(1, -1, 2), result = filter(array, x -> x >
 0) | fields result
fetched rows / total rows = 1/1
+-----------+
| result    |
+-----------+
| [1, 2]    |
+-----------+

os> source=people | eval array = json_array(-1, -3, -2), result = filter(array, x -> x
 > 0) | fields result
fetched rows / total rows = 1/1
+-----------+
| result    |
+-----------+
| []        |
+-----------+
```

TRANSFORM

Utilizzo: `transform(array, lambda)` trasforma gli elementi di un array utilizzando la funzione di trasformazione Lambda. Il secondo argomento implica l'indice dell'elemento se si utilizza la funzione binaria Lambda. È simile a una `map` nella programmazione funzionale.

Tipo di argomento: ARRAY, LAMBDA

Tipo restituito: ARRAY. Un ARRAY che contiene il risultato dell'applicazione della funzione di trasformazione lambda a ciascun elemento dell'array di input.

Esempi:

```
os> source=people | eval array = json_array(1, 2, 3), result = transform(array, x -> x
 + 1) | fields result
fetched rows / total rows = 1/1
+--------------+
| result       |
+--------------+
| [2, 3, 4]    |
+--------------+

os> source=people | eval array = json_array(1, 2, 3), result = transform(array, (x, i)
 -> x + i) | fields result
fetched rows / total rows = 1/1
+--------------+
| result       |
+--------------+
| [1, 3, 5]    |
+--------------+
```

REDUCE

Utilizzo: `reduce(array, start, merge_lambda, finish_lambda)` riduce un array a un singolo valore applicando le funzioni lambda. La funzione applica merge_lambda al valore iniziale e a tutti gli elementi dell'array, quindi applica il valore al risultato. `finish_lambda`

Tipo di argomento: ARRAY, ANY, LAMBDA, LAMBDA

Tipo di restituzione: ANY. Il risultato finale dell'applicazione delle funzioni Lambda al valore iniziale e all'array di input.

Esempi:

```
 os> source=people | eval array = json_array(1, 2, 3), result = reduce(array, 0, (acc,
 x) -> acc + x) | fields result
fetched rows / total rows = 1/1
+-----------+
| result    |
+-----------+
| 6         |
+-----------+

os> source=people | eval array = json_array(1, 2, 3), result = reduce(array, 10, (acc,
 x) -> acc + x) | fields result
fetched rows / total rows = 1/1
+-----------+
| result    |
+-----------+
| 16        |
+-----------+

os> source=people | eval array = json_array(1, 2, 3), result = reduce(array, 0, (acc,
 x) -> acc + x, acc -> acc * 10) | fields result
fetched rows / total rows = 1/1
+-----------+
| result    |
+-----------+
| 60        |
+-----------+
```

Funzioni matematiche PPL

> **Note**
>
> Per vedere quali integrazioni di sorgenti AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

ABS

Utilizzo: `ABS(x)` calcola il valore assoluto di x.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: INTEGER/LONG/FLOAT/DOUBLE

Esempio:

```
os> source=people | eval `ABS(-1)` = ABS(-1) | fields `ABS(-1)`
fetched rows / total rows = 1/1
+-----------+
| ABS(-1)   |
|-----------|
| 1         |
+-----------+
```

ACOS

Utilizzo: `ACOS(x)` calcola l'arcoseno di x. Restituisce `NULL` se x non è compreso nell'intervallo da -1 a 1.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `ACOS(0)` = ACOS(0) | fields `ACOS(0)`
fetched rows / total rows = 1/1
+--------------------+
| ACOS(0)            |
|--------------------|
| 1.5707963267948966 |
+--------------------+
```

ASIN

Utilizzo: `asin(x)` calcola l'arcoseno di x. Restituisce `NULL` se x non è compreso nell'intervallo da -1 a 1.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `ASIN(0)` = ASIN(0) | fields `ASIN(0)`
fetched rows / total rows = 1/1
+-----------+
```

```
| ASIN(0)   |
|-----------|
| 0.0       |
+-----------+
```

ATAN

Uso: `ATAN(x)` calcola l'arcotangente di x. `atan(y, x)` Calcola l'arcotangente di y/x, tranne per il fatto che i segni di entrambi gli argomenti determinano il quadrante del risultato.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `ATAN(2)` = ATAN(2), `ATAN(2, 3)` = ATAN(2, 3) | fields
 `ATAN(2)`, `ATAN(2, 3)`
fetched rows / total rows = 1/1
+--------------------+--------------------+
| ATAN(2)            | ATAN(2, 3)         |
|--------------------+--------------------|
| 1.1071487177940904 | 0.5880026035475675 |
+--------------------+--------------------+
```

ATAN2

Utilizzo: `ATAN2(y, x)` calcola l'arcotangente di y/x, tranne per il fatto che i segni di entrambi gli argomenti determinano il quadrante del risultato.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `ATAN2(2, 3)` = ATAN2(2, 3) | fields `ATAN2(2, 3)`
fetched rows / total rows = 1/1
+--------------------+
| ATAN2(2, 3)        |
|--------------------|
| 0.5880026035475675 |
+--------------------+
```

CBRT

Utilizzo: CBRT calcola la radice cubica di un numero.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE:

INTEGER/LONG/FLOAT/DOUBLE-> DOPPIO

Esempio:

```
opensearchsql> source=location | eval `CBRT(8)` = CBRT(8), `CBRT(9.261)` = CBRT(9.261),
 `CBRT(-27)` = CBRT(-27) | fields `CBRT(8)`, `CBRT(9.261)`, `CBRT(-27)`;
fetched rows / total rows = 2/2
+-----------+--------------+-------------+
| CBRT(8)   | CBRT(9.261)   | CBRT(-27)   |
|-----------+--------------+-------------|
| 2.0       | 2.1           | -3.0        |
| 2.0       | 2.1           | -3.0        |
+-----------+--------------+-------------+
```

CEIL

Utilizzo: un alias per la CEILING funzione. CEILING(T)prende il limite del valore T.

Limitazione: funziona CEILING solo come previsto quando il tipo doppio IEEE 754 visualizza un decimale quando memorizzato.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: LUNGO

Esempio:

```
os> source=people | eval `CEILING(0)` = CEILING(0), `CEILING(50.00005)` =
 CEILING(50.00005), `CEILING(-50.00005)` = CEILING(-50.00005) | fields `CEILING(0)`,
 `CEILING(50.00005)`, `CEILING(-50.00005)`
fetched rows / total rows = 1/1
+-------------+--------------------+----------------------+
| CEILING(0)   | CEILING(50.00005)   | CEILING(-50.00005)   |
|-------------+--------------------+----------------------|
| 0           | 51                  | -50                  |
+-------------+--------------------+----------------------+
```

```
os> source=people | eval `CEILING(3147483647.12345)` = CEILING(3147483647.12345),
 `CEILING(113147483647.12345)` = CEILING(113147483647.12345),
 `CEILING(3147483647.00001)` = CEILING(3147483647.00001) | fields
 `CEILING(3147483647.12345)`, `CEILING(113147483647.12345)`,
 `CEILING(3147483647.00001)`
fetched rows / total rows = 1/1
+-----------------------------+-------------------------------
+-----------------------------+
| CEILING(3147483647.12345)   | CEILING(113147483647.12345)   |
 CEILING(3147483647.00001)   |
|-----------------------------+-------------------------------
+-----------------------------|
| 3147483648                  | 113147483648                  | 3147483648
     |
+-----------------------------+-------------------------------
+-----------------------------+
```

CONV

Utilizzo: `CONV(x, a, b)` converte il numero x da una base a una base b.

Tipo di argomento: x: STRING, a: INTEGER, b: INTEGER

Tipo restituito: STRING

Esempio:

```
os> source=people | eval `CONV('12', 10, 16)` = CONV('12', 10, 16), `CONV('2C', 16,
 10)` = CONV('2C', 16, 10), `CONV(12, 10, 2)` = CONV(12, 10, 2), `CONV(1111, 2, 10)` =
 CONV(1111, 2, 10) | fields `CONV('12', 10, 16)`, `CONV('2C', 16, 10)`, `CONV(12, 10,
 2)`, `CONV(1111, 2, 10)`
fetched rows / total rows = 1/1
+----------------------+----------------------+-------------------
+----------------------+
| CONV('12', 10, 16)   | CONV('2C', 16, 10)   | CONV(12, 10, 2)   | CONV(1111, 2, 10)
 |
|----------------------+----------------------+-------------------
+----------------------|
| c                    | 44                   | 1100              | 15
 |
+----------------------+----------------------+-------------------
+----------------------+
```

COS

Utilizzo: `COS(x)` calcola il coseno di x, dove x è espresso in radianti.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `COS(0)` = COS(0) | fields `COS(0)`
fetched rows / total rows = 1/1
+----------+
| COS(0)   |
|----------|
| 1.0      |
+----------+
```

COT

Utilizzo: `COT(x)` calcola la cotangente di x. Restituisce out-of-range un errore se x è uguale a 0.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `COT(1)` = COT(1) | fields `COT(1)`
fetched rows / total rows = 1/1
+--------------------+
| COT(1)             |
|--------------------|
| 0.6420926159343306 |
+--------------------+
```

CRC32

Utilizzo: `CRC32` calcola un valore di controllo di ridondanza ciclico e restituisce un valore senza segno a 32 bit.

Tipo di argomento: STRING

Tipo restituito: LONG

Esempio:

```
os> source=people | eval `CRC32('MySQL')` = CRC32('MySQL') | fields `CRC32('MySQL')`
fetched rows / total rows = 1/1
+------------------+
| CRC32('MySQL')   |
|------------------|
| 3259397556       |
+------------------+
```

DEGREES

Utilizzo: `DEGREES(x)` converte x da radianti a gradi.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `DEGREES(1.57)` = DEGREES(1.57) | fields `DEGREES(1.57)`
fetched rows / total rows  = 1/1
+-------------------+
| DEGREES(1.57)     |
|-------------------|
| 89.95437383553924 |
+-------------------+
```

E

Utilizzo: `E()` restituisce il numero di Eulero.

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `E()` = E() | fields `E()`
fetched rows / total rows = 1/1
+-------------------+
| E()               |
|-------------------|
| 2.718281828459045 |
+-------------------+
```

EXP

Utilizzo: `EXP(x)` restituisce e elevato alla potenza di x.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `EXP(2)` = EXP(2) | fields `EXP(2)`
fetched rows / total rows = 1/1
+------------------+
| EXP(2)           |
|------------------|
| 7.38905609893065 |
+------------------+
```

FLOOR

Utilizzo: `FLOOR(T)` prende il nome dal valore T.

Limitazione: funziona `FLOOR` solo come previsto quando il tipo doppio IEEE 754 visualizza un decimale quando memorizzato.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: LUNGO

Esempio:

```
os> source=people | eval `FLOOR(0)` = FLOOR(0), `FLOOR(50.00005)` = FLOOR(50.00005),
 `FLOOR(-50.00005)` = FLOOR(-50.00005) | fields `FLOOR(0)`, `FLOOR(50.00005)`,
 `FLOOR(-50.00005)`
fetched rows / total rows = 1/1
+------------+-------------------+--------------------+
| FLOOR(0)   | FLOOR(50.00005)   | FLOOR(-50.00005)   |
|------------+-------------------+--------------------|
| 0          | 50                | -51                |
+------------+-------------------+--------------------+

os> source=people | eval `FLOOR(3147483647.12345)` = FLOOR(3147483647.12345),
 `FLOOR(113147483647.12345)` = FLOOR(113147483647.12345), `FLOOR(3147483647.00001)`
```

```
 = FLOOR(3147483647.00001) | fields `FLOOR(3147483647.12345)`,
 `FLOOR(113147483647.12345)`, `FLOOR(3147483647.00001)`
fetched rows / total rows = 1/1
+---------------------------+----------------------------+---------------------------+
| FLOOR(3147483647.12345)   | FLOOR(113147483647.12345)   | FLOOR(3147483647.00001)   |
|---------------------------+----------------------------+---------------------------|
| 3147483647                | 113147483647                | 3147483647                |
+---------------------------+----------------------------+---------------------------+

os> source=people | eval `FLOOR(282474973688888.022)` = FLOOR(282474973688888.022),
 `FLOOR(9223372036854775807.022)` = FLOOR(9223372036854775807.022),
 `FLOOR(9223372036854775807.0000001)` = FLOOR(9223372036854775807.0000001)
 | fields `FLOOR(282474973688888.022)`, `FLOOR(9223372036854775807.022)`,
 `FLOOR(9223372036854775807.0000001)`
fetched rows / total rows = 1/1
+-----------------------------+-----------------------------------
+-------------------------------------+
| FLOOR(282474973688888.022)   | FLOOR(9223372036854775807.022)   |
 FLOOR(9223372036854775807.0000001)   |
|-----------------------------+-----------------------------------
+-------------------------------------|
| 282474973688888              | 9223372036854775807              | 9223372036854775807
                  |
+-----------------------------+-----------------------------------
+-------------------------------------+
```

LN

Utilizzo: `LN(x)` restituisce il logaritmo naturale di x.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `LN(2)` = LN(2) | fields `LN(2)`
fetched rows / total rows = 1/1
+--------------------+
| LN(2)              |
|--------------------|
| 0.6931471805599453 |
+--------------------+
```

LOG

Utilizzo: `LOG(x)` restituisce il logaritmo naturale di x che è il logaritmo di base e di x. log (B, x) è equivalente a log (x) /log (B).

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `LOG(2)` = LOG(2), `LOG(2, 8)` = LOG(2, 8) | fields `LOG(2)`,
 `LOG(2, 8)`
fetched rows / total rows = 1/1
+--------------------+-------------+
| LOG(2)             | LOG(2, 8)   |
|--------------------+-------------|
| 0.6931471805599453 | 3.0         |
+--------------------+-------------+
```

LOG2

Utilizzo: `LOG2(x)` è equivalente `log(x)/log(2)`.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `LOG2(8)` = LOG2(8) | fields `LOG2(8)`
fetched rows / total rows = 1/1
+-----------+
| LOG2(8)   |
|-----------|
| 3.0       |
+-----------+
```

LOG10

Utilizzo: `LOG10(x)` è equivalente `log(x)/log(10)`.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `LOG10(100)` = LOG10(100) | fields `LOG10(100)`
fetched rows / total rows = 1/1
+--------------+
| LOG10(100)   |
|--------------|
| 2.0          |
+--------------+
```

MOD

Utilizzo: `MOD(n, m)` calcola il resto del numero n diviso per m.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo restituito: tipo più ampio tra i tipi n e m se m è un valore diverso da zero. Se m è uguale a 0, restituisce NULL.

Esempio:

```
os> source=people | eval `MOD(3, 2)` = MOD(3, 2), `MOD(3.1, 2)` = MOD(3.1, 2) | fields
 `MOD(3, 2)`, `MOD(3.1, 2)`
fetched rows / total rows = 1/1
+-------------+---------------+
| MOD(3, 2)   | MOD(3.1, 2)   |
|-------------+---------------|
| 1           | 1.1           |
+-------------+---------------+
```

PI

Utilizzo: `PI()` restituisce la costante pi.

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `PI()` = PI() | fields `PI()`
fetched rows / total rows = 1/1
+-------------------+
```

```
| PI()              |
|-------------------|
| 3.141592653589793 |
+-------------------+
```

POW

Utilizzo: POW(x, y) calcola il valore di x elevato alla potenza di y. Gli input errati restituiscono un NULL risultato.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Sinonimi: POWER(_, _)

Esempio:

```
os> source=people | eval `POW(3, 2)` = POW(3, 2), `POW(-3, 2)` = POW(-3, 2), `POW(3,
 -2)` = POW(3, -2) | fields `POW(3, 2)`, `POW(-3, 2)`, `POW(3, -2)`
fetched rows / total rows = 1/1
+-------------+--------------+--------------------+
| POW(3, 2)   | POW(-3, 2)   | POW(3, -2)         |
|-------------+--------------+--------------------|
| 9.0         | 9.0          | 0.1111111111111111 |
+-------------+--------------+--------------------+
```

POWER

Utilizzo: POWER(x, y) calcola il valore di x elevato alla potenza di y. Gli input errati restituiscono un NULL risultato.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Sinonimi: POW(_, _)

Esempio:

```
os> source=people | eval `POWER(3, 2)` = POWER(3, 2), `POWER(-3, 2)` = POWER(-3, 2),
 `POWER(3, -2)` = POWER(3, -2) | fields `POWER(3, 2)`, `POWER(-3, 2)`, `POWER(3, -2)`
fetched rows / total rows = 1/1
```

```
+---------------+----------------+--------------------+
| POWER(3, 2)   | POWER(-3, 2)   | POWER(3, -2)       |
|---------------+----------------+--------------------|
| 9.0           | 9.0            | 0.1111111111111111 |
+---------------+----------------+--------------------+
```

RADIANS

Utilizzo: `RADIANS(x)` converte x da gradi a radianti.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `RADIANS(90)` = RADIANS(90) | fields `RADIANS(90)`
fetched rows / total rows  = 1/1
+--------------------+
| RADIANS(90)        |
|--------------------|
| 1.5707963267948966 |
+--------------------+
```

RAND

Utilizzo:`RAND()`/`RAND(N)`restituisce un valore casuale a virgola mobile nell'intervallo 0 <= valore < 1,0. Se specificate il numero intero N, la funzione inizializza il seme prima dell'esecuzione. Un'implicazione di questo comportamento è che con un argomento identico N, `rand(N)` restituisce lo stesso valore ogni volta, producendo una sequenza ripetibile di valori di colonna.

Tipo di argomento: INTERO

Tipo restituito: FLOAT

Esempio:

```
os> source=people | eval `RAND(3)` = RAND(3) | fields `RAND(3)`
fetched rows / total rows = 1/1
+------------+
| RAND(3)    |
|------------|
| 0.73105735 |
```

```
+------------+
```

ROUND

Utilizzo: `ROUND(x, d)` arrotonda l'argomento x a d cifre decimali. Se non si specifica d, il valore predefinito è 0.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Mappa del tipo di ritorno:

- (INTEGER/LONG [, INTEGER]) -> LONG
- (FLOAT/DOUBLE [, INTERO]) -> LUNGO

Esempio:

```
os> source=people | eval `ROUND(12.34)` = ROUND(12.34), `ROUND(12.34, 1)` =
 ROUND(12.34, 1), `ROUND(12.34, -1)` = ROUND(12.34, -1), `ROUND(12, 1)` = ROUND(12, 1)
 | fields `ROUND(12.34)`, `ROUND(12.34, 1)`, `ROUND(12.34, -1)`, `ROUND(12, 1)`
fetched rows / total rows = 1/1
+----------------+-------------------+--------------------+----------------+
| ROUND(12.34)   | ROUND(12.34, 1)   | ROUND(12.34, -1)   | ROUND(12, 1)   |
|----------------+-------------------+--------------------+----------------|
| 12.0           | 12.3              | 10.0               | 12             |
+----------------+-------------------+--------------------+----------------+
```

SIGN

Utilizzo: `SIGN` restituisce il segno dell'argomento come -1, 0 o 1, a seconda che il numero sia negativo, zero o positivo.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo restituito: INTEGER

Esempio:

```
os> source=people | eval `SIGN(1)` = SIGN(1), `SIGN(0)` = SIGN(0), `SIGN(-1.1)` =
 SIGN(-1.1) | fields `SIGN(1)`, `SIGN(0)`, `SIGN(-1.1)`
fetched rows / total rows = 1/1
+-----------+-----------+--------------+
```

```
| SIGN(1)   | SIGN(0)   | SIGN(-1.1)   |
|-----------+-----------+--------------|
| 1         | 0         | -1           |
+-----------+-----------+--------------+
```

SIN

Utilizzo: `sin(x)` calcola il seno di x, dove x è espresso in radianti.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Tipo di restituzione: DOUBLE

Esempio:

```
os> source=people | eval `SIN(0)` = SIN(0) | fields `SIN(0)`
fetched rows / total rows = 1/1
+----------+
| SIN(0)   |
|----------|
| 0.0      |
+----------+
```

SQRT

Utilizzo: SQRT calcola la radice quadrata di un numero non negativo.

Tipo di argomento: INTEGER/LONG/FLOAT/DOUBLE

Mappa del tipo di ritorno:

- (Non negativo) INTEGER/LONG/FLOAT/DOUBLE -> DOUBLE
- (Negativo) INTEGER/LONG/FLOAT/DOUBLE -> NULL

Esempio:

```
os> source=people | eval `SQRT(4)` = SQRT(4), `SQRT(4.41)` = SQRT(4.41) | fields
 `SQRT(4)`, `SQRT(4.41)`
fetched rows / total rows = 1/1
+-----------+--------------+
| SQRT(4)   | SQRT(4.41)   |
|-----------+--------------|
```

```
| 2.0       | 2.1          |
+-----------+--------------+
```

Funzioni di stringa PPL

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

CONCAT

Utilizzo: `CONCAT(str1, str2, ...., str_9)` somma fino a 9 stringhe insieme.

Tipo di argomento:

- STRINGA, STRINGA,..., STRINGA
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `CONCAT('hello', 'world')` = CONCAT('hello', 'world'),
 `CONCAT('hello ', 'whole ', 'world', '!')` = CONCAT('hello ', 'whole ', 'world', '!')
 | fields `CONCAT('hello', 'world')`, `CONCAT('hello ', 'whole ', 'world', '!')`
fetched rows / total rows = 1/1
+----------------------------+---------------------------------------------+
| CONCAT('hello', 'world')   | CONCAT('hello ', 'whole ', 'world', '!')    |
|----------------------------+---------------------------------------------|
| helloworld                 | hello whole world!                          |
+----------------------------+---------------------------------------------+
```

CONCAT_WS

Utilizzo: `CONCAT_WS(sep, str1, str2)` concatena due o più stringhe utilizzando un separatore specificato tra di loro.

Tipo di argomento:

- STRINGA, STRINGA,..., STRINGA

- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `CONCAT_WS(',', 'hello', 'world')` = CONCAT_WS(',', 'hello',
 'world') | fields `CONCAT_WS(',', 'hello', 'world')`
fetched rows / total rows = 1/1
+-----------------------------------+
| CONCAT_WS(',', 'hello', 'world')   |
|-----------------------------------|
| hello,world                       |
+-----------------------------------+
```

LENGTH

Utilizzo: `length(str)` restituisce la lunghezza della stringa di input misurata in byte.

Tipo di argomento:

- STRING
- Tipo restituito: INTEGER

Esempio:

```
os> source=people | eval `LENGTH('helloworld')` = LENGTH('helloworld') | fields
 `LENGTH('helloworld')`
fetched rows / total rows = 1/1
+------------------------+
| LENGTH('helloworld')   |
|------------------------|
| 10                     |
+------------------------+
```

LOWER

Utilizzo: `lower(string)` converte la stringa di input in lettere minuscole.

Tipo di argomento:

- STRING
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `LOWER('helloworld')` = LOWER('helloworld'),
 `LOWER('HELLOWORLD')` = LOWER('HELLOWORLD') | fields `LOWER('helloworld')`,
 `LOWER('HELLOWORLD')`
fetched rows / total rows = 1/1
+----------------------+----------------------+
| LOWER('helloworld')   | LOWER('HELLOWORLD')   |
|----------------------+----------------------|
| helloworld            | helloworld            |
+----------------------+----------------------+
```

LTRIM

Utilizzo: `ltrim(str)` rimuove i caratteri spaziali iniziali dalla stringa di input.

Tipo di argomento:

- STRING
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `LTRIM('   hello')` = LTRIM('   hello'), `LTRIM('hello   ')` =
 LTRIM('hello   ') | fields `LTRIM('   hello')`, `LTRIM('hello   ')`
fetched rows / total rows = 1/1
+---------------------+---------------------+
| LTRIM('   hello')   | LTRIM('hello   ')   |
|---------------------+---------------------|
| hello               | hello               |
+---------------------+---------------------+
```

POSITION

Utilizzo: `POSITION(substr IN str)` restituisce la posizione della prima occorrenza della sottostringa nella stringa. Restituisce 0 se la sottostringa non è nella stringa. Restituisce NULL se un argomento è NULL.

Tipo di argomento:

- STRINGA, STRINGA

- Tipo restituito INTEGER

Esempio:

```
os> source=people | eval `POSITION('world' IN 'helloworld')` = POSITION('world'
 IN 'helloworld'), `POSITION('invalid' IN 'helloworld')`= POSITION('invalid' IN
 'helloworld')  | fields `POSITION('world' IN 'helloworld')`, `POSITION('invalid' IN
 'helloworld')`
fetched rows / total rows = 1/1
+-------------------------------------+---------------------------------------+
| POSITION('world' IN 'helloworld')   | POSITION('invalid' IN 'helloworld')   |
|-------------------------------------+---------------------------------------|
| 6                                   | 0                                     |
+-------------------------------------+---------------------------------------+
```

`REVERSE`

Utilizzo: `REVERSE(str)` restituisce la stringa invertita della stringa di input.

Tipo di argomento:

- STRING
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `REVERSE('abcde')` = REVERSE('abcde') | fields
 `REVERSE('abcde')`
fetched rows / total rows = 1/1
+--------------------+
| REVERSE('abcde')   |
|--------------------|
| edcba              |
+--------------------+
```

`RIGHT`

Utilizzo: `right(str, len)` restituisce i caratteri più a destra della stringa di input. Restituisce 0 se la sottostringa non è nella stringa. Restituisce NULL se un argomento è NULL.

Tipo di argomento:

- STRINGA, NUMERO INTERO
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `RIGHT('helloworld', 5)` = RIGHT('helloworld', 5),
 `RIGHT('HELLOWORLD', 0)` = RIGHT('HELLOWORLD', 0) | fields `RIGHT('helloworld', 5)`,
 `RIGHT('HELLOWORLD', 0)`
fetched rows / total rows = 1/1
+-------------------------+-------------------------+
| RIGHT('helloworld', 5)  | RIGHT('HELLOWORLD', 0)  |
|-------------------------+-------------------------|
| world                   |                         |
+-------------------------+-------------------------+
```

RTRIM

Utilizzo: `rtrim(str)` elimina i caratteri spaziali finali dalla stringa di input.

Tipo di argomento:

- STRING
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `RTRIM('   hello')` = RTRIM('   hello'), `RTRIM('hello   ')` =
 RTRIM('hello   ') | fields `RTRIM('   hello')`, `RTRIM('hello   ')`
fetched rows / total rows = 1/1
+--------------------+--------------------+
| RTRIM('   hello')  | RTRIM('hello   ')  |
|--------------------+--------------------|
|    hello           | hello              |
+--------------------+--------------------+
```

SUBSTRING

Utilizzo: `substring(str, start)` o `substring(str, start, length)` restituisce una sottostringa della stringa di input. Senza una lunghezza specificata, restituisce l'intera stringa dalla posizione iniziale.

Tipo di argomento:

- STRINGA, NUMERO INTERO, NUMERO INTERO
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `SUBSTRING('helloworld', 5)` = SUBSTRING('helloworld',
 5), `SUBSTRING('helloworld', 5, 3)` = SUBSTRING('helloworld', 5, 3) | fields
 `SUBSTRING('helloworld', 5)`, `SUBSTRING('helloworld', 5, 3)`
fetched rows / total rows = 1/1
+------------------------------+---------------------------------+
| SUBSTRING('helloworld', 5)   | SUBSTRING('helloworld', 5, 3)   |
|------------------------------+---------------------------------|
| oworld                       | owo                             |
+------------------------------+---------------------------------+
```

`TRIM`

Utilizzo: `trim(string)` rimuove gli spazi bianchi iniziali e finali dalla stringa di input.

Tipo di argomento:

- STRING
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `TRIM('   hello')` = TRIM('   hello'), `TRIM('hello   ')` =
 TRIM('hello   ') | fields `TRIM('   hello')`, `TRIM('hello   ')`
fetched rows / total rows = 1/1
+--------------------+--------------------+
| TRIM('   hello')   | TRIM('hello   ')   |
|--------------------+--------------------|
| hello              | hello              |
+--------------------+--------------------+
```

`UPPER`

Utilizzo: `upper(string)` converte la stringa di input in maiuscolo.

Tipo di argomento:

- STRING
- Tipo restituito: STRING

Esempio:

```
os> source=people | eval `UPPER('helloworld')` = UPPER('helloworld'),
 `UPPER('HELLOWORLD')` = UPPER('HELLOWORLD') | fields `UPPER('helloworld')`,
 `UPPER('HELLOWORLD')`
fetched rows / total rows = 1/1
+----------------------+----------------------+
| UPPER('helloworld')   | UPPER('HELLOWORLD')   |
|----------------------+----------------------|
| HELLOWORLD            | HELLOWORLD            |
+----------------------+----------------------+
```

Funzioni di conversione di tipo PPL

> **Note**
>
> Per vedere quali integrazioni di fonti di AWS dati supportano questa funzione PPL, vedi. the section called “Funzioni”

`TRIM`

Utilizzo: `cast(expr as dateType)` trasmette il `expr` a `dataType` e restituisce il valore di. `dataType`

Si applicano le seguenti regole di conversione:

Regole di conversione dei tipi

Src/Target	STRING	NUMBER	BOOLEAN	TIMESTAMP	DATE	TIME
STRING		Nota 1	Nota 1	TIMESTAMP ()	DATE()	TIME()
NUMBER	Nota 1		v! =0	N/D	N/D	N/D

Src/Target	STRING	NUMBER	BOOLEAN	TIMESTAMP	DATE	TIME
BOOLEAN	Nota 1	v? 1:0		N/D	N/D	N/D
TIMESTAMP	Nota 1	N/D	N/D		DATE()	TIME()
DATE	Nota 1	N/D	N/D	N/D		N/D
TIME	Nota 1	N/D	N/D	N/D	N/D	

Esempio di cast su stringa:

```
os> source=people | eval `cbool` = CAST(true as string), `cint` = CAST(1 as string),
 `cdate` = CAST(CAST('2012-08-07' as date) as string) | fields `cbool`, `cint`, `cdate`
fetched rows / total rows = 1/1
+---------+--------+------------+
| cbool   | cint   | cdate      |
|---------+--------+------------|
| true    | 1      | 2012-08-07 |
+---------+--------+------------+
```

Esempio di cast to number:

```
os> source=people | eval `cbool` = CAST(true as int), `cstring` = CAST('1' as int) |
 fields `cbool`, `cstring`
fetched rows / total rows = 1/1
+---------+-----------+
| cbool   | cstring   |
|---------+-----------|
| 1       | 1         |
+---------+-----------+
```

Esempio di cast to date:

```
os> source=people | eval `cdate` = CAST('2012-08-07' as date), `ctime` =
 CAST('01:01:01' as time), `ctimestamp` = CAST('2012-08-07 01:01:01' as timestamp) |
 fields `cdate`, `ctime`, `ctimestamp`
fetched rows / total rows = 1/1
+------------+----------+---------------------+
| cdate      | ctime    | ctimestamp          |
|------------+----------+---------------------|
```

```
| 2012-08-07 | 01:01:01 | 2012-08-07 01:01:01 |
+------------+----------+---------------------+
```

Esempio di cast concatenato:

```
os> source=people | eval `cbool` = CAST(CAST(true as string) as boolean) | fields
 `cbool`
fetched rows / total rows = 1/1
+---------+
| cbool   |
|---------|
| True    |
+---------+
```

Monitoraggio dei domini Amazon OpenSearch Service

Il monitoraggio è una parte importante per mantenere l'affidabilità, la disponibilità e le prestazioni di Amazon OpenSearch Service e delle altre AWS soluzioni. AWS fornisce i seguenti strumenti per monitorare le risorse OpenSearch del Servizio, segnalare problemi e intraprendere azioni automatiche quando necessario:

Amazon CloudWatch

Amazon CloudWatch monitora le risorse OpenSearch del Servizio in tempo reale. È possibile raccogliere i parametri e tenerne traccia, creare pannelli di controllo personalizzati e impostare allarmi che inviino una notifica o intraprendano azioni quando un parametro specificato raggiunge una determinata soglia. Per ulteriori informazioni, consulta la Amazon CloudWatch User Guide.

CloudWatch Registri Amazon

Amazon CloudWatch Logs ti consente di monitorare, archiviare e accedere ai tuoi file di OpenSearch registro. CloudWatch Logs monitora le informazioni nei file di registro e può avvisarti quando vengono raggiunte determinate soglie. Per ulteriori informazioni, consulta la Amazon CloudWatch Logs User Guide.

Amazon EventBridge

Amazon EventBridge offre un flusso quasi in tempo reale di eventi di sistema che descrivono le modifiche nei tuoi domini OpenSearch di servizio. Puoi creare regole che controllano determinati eventi e attivare azioni automatiche in altri AWS servizi quando si verificano tali eventi. Per ulteriori informazioni, consulta la Amazon EventBridge User Guide.

AWS CloudTrail

AWS CloudTrail acquisisce le chiamate API di configurazione effettuate a OpenSearch Service come eventi. Può distribuire questi eventi a un bucket Amazon S3 specificato. Con queste informazioni, è possibile identificare quali utenti e account hanno effettuato richieste, l'indirizzo IP di origine da cui sono state effettuate le richieste e quando sono avvenute. Per ulteriori informazioni, consulta la AWS Guida per l'utente CloudTrail .

Argomenti

- Monitoraggio delle metriche dei OpenSearch cluster con Amazon CloudWatch
- Monitoraggio dei OpenSearch log con Amazon CloudWatch Logs

- Monitoraggio dei log di controllo in Amazon Service OpenSearch
- Monitoraggio degli eventi del OpenSearch servizio con Amazon EventBridge
- Monitoraggio delle chiamate API di Amazon OpenSearch Service con AWS CloudTrail

Monitoraggio delle metriche dei OpenSearch cluster con Amazon CloudWatch

Amazon OpenSearch Service pubblica i dati dei tuoi domini su Amazon. CloudWatch CloudWatch ti consente di recuperare le statistiche su tali punti dati sotto forma di un insieme ordinato di dati di serie temporali, noti come metriche. OpenSearch Il servizio invia la maggior parte delle metriche a CloudWatch intervalli di 60 secondi. Se utilizzi volumi magnetici EBS o per uso generale, i parametri relativi al volume EBS si aggiorneranno ogni cinque minuti. Tutte le metriche cumulative (ad esempio`ThreadpoolSearchRejected`) sono in memoria e `ThreadpoolWriteRejected` perderanno lo stato. Le metriche verranno reimpostate durante la caduta di un nodo, il rimbalzo del nodo, la sostituzione del nodo e la distribuzione. blue/green Per ulteriori informazioni su Amazon CloudWatch, consulta la Amazon CloudWatch User Guide.

La console OpenSearch di servizio mostra una serie di grafici basati sui dati grezzi di CloudWatch. A seconda delle esigenze, potresti preferire visualizzare i dati del cluster CloudWatch anziché i grafici nella console. Il servizio archivia i parametri per due settimane prima di eliminarli. Le metriche vengono fornite senza costi aggiuntivi, ma sono CloudWatch comunque a pagamento per la creazione di dashboard e allarmi. Per ulteriori informazioni, consulta i CloudWatchprezzi di Amazon.

OpenSearch Il servizio pubblica le seguenti metriche su: CloudWatch

- the section called “Parametri cluster”
- the section called “Parametri nodo master dedicato”
- the section called “Parametri volume EBS”
- the section called “Parametri dell'istanza”
- the section called “Metriche calde”
- the section called “Metriche del nodo di coordinatore dedicato”
- the section called “Parametri di archiviazione a freddo”
- the section called “Parametri di avvisi”
- the section called “Parametri di rilevamento delle anomalie”

- the section called “Parametri di ricerca asincrona”
- the section called “Parametri SQL”
- the section called “Parametri k-NN”
- the section called “Parametri di ricerca tra cluster”
- the section called “Parametri di replica tra cluster”
- the section called “Parametri di Learning to Rank”
- the section called “Parametri Piped Processing Language (PPL)”

Visualizzazione delle metriche in CloudWatch

CloudWatch le metriche vengono raggruppate prima in base allo spazio dei nomi del servizio e quindi in base alle varie combinazioni di dimensioni all'interno di ogni spazio dei nomi.

Per visualizzare le metriche utilizzando la console CloudWatch

1. Apri la CloudWatch console all'indirizzo https://console.aws.amazon.com/cloudwatch/.
2. Nel pannello di navigazione a sinistra, scegli Metrics (Parametri), quindi scegli All metrics (Tutti i parametri). Seleziona lo spazio dei OpenSearchService nomi ES/.
3. Scegliere una dimensione per visualizzare i parametri corrispondenti. I parametri per i singoli nodi si trovano nella dimensione `ClientId, DomainName, NodeId`. I parametri del cluster si trovano nella dimensione `Per-Domain, Per-Client Metrics`. Alcuni parametri dei nodi vengono aggregati a livello di cluster e quindi inclusi in entrambe le dimensioni. I parametri delle partizioni si trovano nella dimensione `ClientId, DomainName, NodeId, ShardRole`.

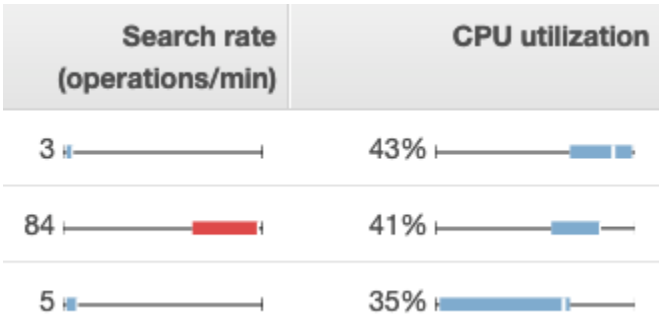
Per visualizzare un elenco di metriche utilizzando il AWS CLI

Esegui il comando seguente:

```
aws cloudwatch list-metrics --namespace "AWS/ES"
```

Interpretazione delle cartelle cliniche in Service OpenSearch

Per visualizzare le metriche in OpenSearch Service, utilizza le schede Cluster Health e Instance Health. La scheda Instance Health utilizza diagrammi a riquadri per fornire at-a-glance visibilità sullo stato di ogni OpenSearch nodo:



- Ogni casella colorata mostra l'intervallo di valori per il nodo nel periodo di tempo specificato.
- Le caselle blu rappresentano i valori che sono compatibili con gli altri nodi. Le caselle rosse rappresentano i valori erratici.
- La linea bianca all'interno di ogni casella mostra il valore corrente del nodo.
- Le "parentesi angolari" su entrambi i lati di ciascuna casella mostrano i valori minimo e massimo per tutti i nodi nel periodo di tempo.

Se si apportano modifiche alla configurazione del dominio, l'elenco delle singole istanze nelle schede Cluster health (Stato cluster) e Instance health (Stato istanza) raddoppierà spesso in dimensione per un breve periodo prima di tornare al numero corretto. Per una spiegazione del comportamento, consulta the section called “Modifiche di configurazione”.

Parametri cluster

Amazon OpenSearch Service fornisce le seguenti metriche per i cluster.

Metrica	Description
`ClusterStatus.green`	Un valore pari a 1 indica che tutte le partizioni di indice sono assegnate a nodi nel cluster. Statistiche rilevanti: Massima
`ClusterStatus.yellow`	Un valore pari a 1 indica che le partizioni principali per tutti gli indici sono allocate a nodi nel cluster, ma che le partizioni di replica per almeno un indice non lo sono. Per ulteriori informazioni, consulta the section called “Stato giallo del cluster”. Statistiche rilevanti: Massima

Metrica	Description
`ClusterStatus.red`	Un valore pari a 1 indica che le partizioni primarie e di replica di almeno un indice non sono allocate ai nodi nel cluster. Per ulteriori informazioni, consultare the section called “Cluster in stato rosso”. Statistiche rilevanti: Massima
`Shards.active`	Il numero totale di partizioni primarie e di replica attive. Statistiche rilevanti: Massima, Somma
`Shards.unassigned`	Il numero di partizioni non allocate ai nodi nel cluster. Statistiche rilevanti: Massima, Somma
`Shards.delayedUnassigned`	Il numero di partizioni la cui allocazione dei nodi è stata ritardata dalle impostazioni di timeout. Statistiche rilevanti: Massima, Somma
`Shards.activePrimary`	Il numero di partizioni primarie attive. Statistiche rilevanti: Massima, Somma
`Shards.initializing`	Il numero di partizioni in fase di inizializzazione. Statistiche rilevanti: Sum (Somma)
`Shards.relocating`	Il numero di partizioni in fase di rilocazione. Statistiche rilevanti: Sum (Somma)
`Nodes`	Il numero di nodi nel cluster di OpenSearch servizio, inclusi i nodi master dedicati e i nodi Warm. Per ulteriori informazioni, consulta the section called “Modifiche di configurazione”. Statistiche rilevanti: Massima

Metrica	Description
`SearchableDocuments`	Il numero totale di documenti disponibili per la ricerca tra tutti i nodi di dati nel cluster. Statistiche rilevanti: Minima, Massima, Media
`DeletedDocuments`	Il numero totale di documenti contrassegnati per l'eliminazione tra tutti i nodi di dati nel cluster. Questi documenti non vengono più visualizzati nei risultati di ricerca, ma rimuovono dal disco OpenSearch solo i documenti eliminati durante l'unione dei segmenti. Questo parametro aumenta dopo le richieste di eliminazione e diminuisce dopo la fusione dei segmenti. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`CPUUtilization`	Percentuale di utilizzo della CPU per i nodi di dati nel cluster. Il numero massimo mostra il nodo con il più alto utilizzo della CPU. La media rappresenta tutti i nodi del cluster. Questo parametro è disponibile anche per singoli nodi. Statistiche rilevanti: Maximum (Massimo), Average (Media)

Metrica	Description
`FreeStorageSpace`	Lo spazio libero per i nodi di dati nel cluster. `Sum` mostra lo spazio libero totale per il cluster, ma è necessario lasciare il periodo a un minuto per ottenere un valore accurato. `Minimum` e `Maximum` mostrano i nodi con lo spazio libero maggiore e minore, rispettiv amente. Questa metrica è disponibile anche per i singoli nodi. OpenSearch Il servizio genera un `ClusterBlockException` quando questa metrica raggiunge. `0` Per risolvere il problema devi eliminare gli indici, aggiungere istanze più grandi oppure aggiunger e archiviazione basata su EBS alle istanze esistenti. Per ulteriori informazioni, consulta the section called “Mancanza di spazio di archiviazione disponibile”. La console OpenSearch di servizio visualizza questo valore in GiB. La CloudWatch console Amazon lo visualizza in MiB. **Note** `FreeStorageSpace` sarà sempre inferiore ai valori `_cat/allocation` APIs forniti da OpenSearch `_cluster/stats` e. OpenSearch Il servizio riserva una percentuale dello spazio di archiviazione su ciascuna istanza per le operazioni interne. Per ulteriori informazioni, consultare Calcolo dei requisiti di archiviazione. Statistiche rilevanti: Minimum (Minimo), Maximum (Massimo), Average (Media), Sum (Somma)
`ClusterUsedSpace`	Lo spazio totale utilizzato per il cluster. È necessario lasciare il periodo su un minuto per ottenere un valore preciso. La console OpenSearch di servizio visualizza questo valore in GiB. La CloudWatch console Amazon lo visualizza in MiB. Statistiche rilevanti: Minimum (Minimo), Maximum (Massimo)

Metrica	Description
`ClusterIndexWritesBlocked`	Indica se il cluster accetta o blocca le richieste di scrittura in entrata. Un valore pari a 0 significa che il cluster accetta le richieste . Un valore pari a 1 significa che il cluster blocca le richieste. Alcuni fattori comuni sono i seguenti: `FreeStorageSpace` è troppo basso oppure `JVMMemoryPressure` è troppo alto. Per contenere questo problema, puoi decidere di aggiungere altro spazio su disco oppure di dimensionare il tuo cluster. Statistiche rilevanti: Massima
`JVMMemoryPressure`	La percentuale massima dell'heap Java utilizzata per tutti i nodi di dati del cluster. OpenSearch Il servizio utilizza metà della RAM di un'istanza per l'heap Java, fino a una dimensione dell'heap di 32 GiB. Puoi scalare le istanze verticalmente fino a 64 GiB di RAM e poi scalare orizzontalmente aggiungendo le istanze. Per informazi oni, consulta the section called “Allarmi consigliati CloudWatch ”. Statistiche rilevanti: Massima **Note** La logica di questo parametro è cambiata nel software del servizio R20220323. Per ulteriori informazioni, consulta le note di rilascio.
`OldGenJVMMemoryPressure`	La percentuale massima dell'heap Java utilizzata per la "vecchia generazione" di tutti i nodi di dati nel cluster. Questo parametro è disponibile anche a livello di nodo. Statistiche rilevanti: Massima

Metrica	Description
`AutomatedSnapshotFailure`	Il numero di snapshot automatici non riusciti per il cluster. Un valore pari a 1 indica che non è stato acquisito alcuno snapshot automatico per il dominio nelle 36 ore precedenti. Statistiche rilevanti: Minimum (Minimo), Maximum (Massimo)
`CPUCreditBalance`	I crediti CPU rimanenti disponibili per i nodi di dati nel cluster. Un credito CPU fornisce le prestazioni di un core CPU completo per un minuto. Per ulteriori informazioni, consultare Crediti CPU nella Guida per gli sviluppatori di Amazon EC2. Questo parametro è disponibile solo per i tipi di istanza T2. Statistiche rilevanti: Minimum (Minimo)
`OpenSearchDashboardsHealthyNodes`	Un controllo dello stato di salute per Dashboards. OpenSearch Se il valore minimo, massimo e medio sono tutti uguali a 1, Dashboards si comporta normalmente. Se si dispone di 10 nodi con un massimo di 1, minimo di 0 e media di 0,7, allora significa che 7 nodi (70%) sono integri e 3 nodi (30%) non lo sono. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`OpensearchDashboardsReportingFailedRequestSysErrCount`	Il numero di richieste di generazione di report di OpenSearch Dashboard che non sono riuscite a causa di problemi del server o limitazioni delle funzionalità. Statistiche rilevanti: Sum (Somma)
`OpensearchDashboardsReportingFailedRequestUserErrCount`	Il numero di richieste di generazione di report di OpenSearch Dashboards che non sono riuscite a causa di problemi del client. Statistiche rilevanti: Sum (Somma)

Metrica	Description
OpensearchDashboardsReportingRequestCount	Il numero totale di richieste per generare OpenSearch report Dashboards. Statistiche rilevanti: Sum (Somma)
OpensearchDashboardsReportingSuccessCount	Il numero di richieste riuscite per generare OpenSearch report Dashboards. Statistiche rilevanti: Sum (Somma)
KMSKeyError	Il valore 1 indica che la AWS KMS chiave utilizzata per crittogra fare i dati inattivi è stata disabilitata. Per ripristinare il dominio sulle operazioni normali, riabilita la chiave. La console visualizza questo parametro solo per i domini che crittografano i dati a riposo. Statistiche rilevanti: Minimum (Minimo), Maximum (Massimo)
KMSKeyInaccessible	Il valore 1 indica che la AWS KMS chiave utilizzata per crittografare i dati inattivi è stata eliminata o le relative concessioni al Servizio sono state revocate. OpenSearch Non è possibile recuperare i domini che sono in questo stato. Se hai una snapshot manuale, puoi utilizzarla per migrare i dati del dominio in un nuovo dominio. La console visualizza questo parametro solo per i domini che crittografano i dati a riposo. Statistiche rilevanti: Minimum (Minimo), Maximum (Massimo)

Metrica	Description
`InvalidHostHeaderRequests`	Il numero di richieste HTTP effettuate al OpenSearch cluster che includevano un'intestazione host non valida (o mancante). Le richieste valide includono il nome host del dominio come valore dell'intestazione dell'host. OpenSearch Il servizio rifiuta le richieste non valide per i domini di accesso pubblico che non dispongono di una politica di accesso restrittiva. Si consiglia di applicare una policy di accesso restrittiva a tutti i domini. Se per questo parametro si rilevano valori di grandi dimensioni, confermare che i client OpenSearch includano il nome host del dominio (e non, ad esempio, l'indirizzo IP) nelle richieste. Statistiche rilevanti: Sum (Somma)
`OpenSearchRequests (previously ElasticsearchRequests)`	Il numero di richieste fatte al cluster. OpenSearch Statistiche rilevanti: Sum (Somma)
`2xx, 3xx, 4xx, 5xx`	Il numero di richieste al dominio che hanno prodotto il codice di risposta HTTP specificato (2xx, 3xx, 4xx, 5xx). Statistiche rilevanti: Sum (Somma)

Metrica	Description
`ThroughputThrottle`	Indica se i dischi sono stati limitati o meno. La limitazione si verifica quando la velocità effettiva combinata di `ReadThroughputMicroBursting` e `WriteThroughputMicroBursting` è superiore alla velocità massima,. `MaxProvisionedThroughput` `MaxProvisionedThroughput` è il valore più basso del throughput dell'istanza o del throughput di volume assegnato. Il valore 1 indica che i dischi sono stati limitati. Un valore 0 indica un comportamento normale. Per informazioni sulla velocità effettiva delle istanze, consulta Amazon EBS: istanze ottimizzate. Per informazioni sulla velocità effettiva dei volumi, consulta i tipi di volume di Amazon EBS. Statistiche rilevanti: Minimum (Minimo), Maximum (Massimo)
`IopsThrottle`	Indica se il numero di input/output operazioni al secondo (IOPS) sul dominio è stato limitato o meno. La limitazione si verifica quando gli IOPS del nodo dati superano il limite massimo consentito del volume EBS o dell'istanza EC2 del nodo dati. Per informazioni sugli IOPS delle istanze, consulta Amazon EBS: istanze ottimizzate. Per informazioni sugli IOPS di volume, consulta i tipi di volume di Amazon EBS. Statistiche rilevanti: Minimum (Minimo), Maximum (Massimo)
`HighSwapUsage`	Il valore 1 indica che lo scambio dovuto a errori di pagina ha potenzialmente causato picchi nell'utilizzo del disco sottostante durante un periodo di tempo specifico. Statistiche rilevanti: Massima

Parametri nodo master dedicato

Amazon OpenSearch Service fornisce le seguenti metriche per i nodi master dedicati.

Metrica	Description
`MasterCPUUtilization`	La percentuale massima di risorse della CPU utilizzate dai nodi master dedicati. È consigliato aumentare le dimensioni del tipo di istanza quando questo parametro raggiunge 60%. Statistiche rilevanti: Massima
`MasterFreeStorageSpace`	Questo parametro non è rilevante e può essere ignorato. Il servizio non utilizza i nodi master come nodi di dati.
`MasterJVMMemoryPressure`	La percentuale massima dell'heap di Java utilizzata per tutti i nodi master dedicati nel cluster. È consigliato passare a un tipo di istanza più grande quando questo parametro raggiunge 85%. Statistiche rilevanti: Massima **Note** La logica di questo parametro è cambiata nel software del servizio R20220323. Per ulteriori informazioni, consulta le note di rilascio.
`MasterOldGenJVMMemoryPressure`	La percentuale massima dell'heap Java utilizzata per la "vecchia generazione" per ciascun nodo principale. Statistiche rilevanti: Massima
`MasterCPUCreditBalance`	I crediti CPU rimanenti disponibili per i nodi master dedicati nel cluster. Un credito CPU fornisce le prestazioni di un core CPU completo per un minuto. Per ulteriori informazioni, consultare Crediti CPU nella Guida per gli sviluppatori di Amazon EC2. Questo parametro è disponibile solo per i tipi di istanza T2. Statistiche rilevanti: Minimum (Minimo)
`MasterReachableFromNode`	Un controllo dello stato per le eccezioni `MasterNotDiscovered` . Un valore pari a 1 indica un comportamento normale. Un

Metrica	Description
	valore di pari a indica che `/_cluster/health/` ha avuto esito negativo. Gli errori indicano che il nodo master non è raggiungibile dal nodo di origine. Di solito sono il risultato di un problema di connettività di rete o di un AWS problema di dipendenza. Statistiche rilevanti: Massima
`MasterSysMemoryUtilization`	La percentuale di memoria del nodo master utilizzata. Statistiche rilevanti: Massima

Metriche del nodo di coordinatore dedicato

Amazon OpenSearch Service fornisce le seguenti metriche per i nodi coordinatori dedicati.

Metrica	Description
`CoordinatorCPUUtilization`	La percentuale massima di risorse CPU utilizzate dai nodi coordinatori dedicati. Ti consigliamo di aumentare la dimensione del tipo di istanza quando questa metrica raggiunge l'80%. Statistiche rilevanti: Massima
`CoordinatorJVMMemoryPressure`	La percentuale massima dell'heap Java utilizzata per tutti i nodi coordinatori dedicati nel cluster. È consigliato passare a un tipo di istanza più grande quando questo parametro raggiunge 85%. Statistiche rilevanti: Massima
`CoordinatorOldGenJVMMemoryPressure`	La percentuale massima dell'heap Java utilizzata per la "vecchia generazione" per ciascun nodo principale. Statistiche rilevanti: Massima
`CoordinatorSysMemoryUtilization`	La percentuale di memoria del nodo coordinatore in uso.

Metrica	Description
	Statistiche rilevanti: Massima
`CoordinatorFreeStorageSpace`	Questa metrica indica che il servizio non utilizza i nodi coordinatori come nodi di dati.

Parametri volume EBS

Amazon OpenSearch Service fornisce le seguenti metriche per i volumi EBS.

Metrica	Description
`ReadLatency`	La latenza, in secondi, per le operazioni di lettura sui volumi di EBS. Questo parametro è disponibile anche per singoli nodi. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`WriteLatency`	La latenza, in secondi, per le operazioni di scrittura sui volumi di EBS. Questo parametro è disponibile anche per singoli nodi. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`ReadThroughput`	Il throughput, in byte al secondo, per le operazioni di lettura sui volumi di EBS. Questo parametro è disponibile anche per singoli nodi. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`ReadThroughputMicroBursting`	Il throughput, in byte al secondo, per le operazioni di lettura sui volumi EBS quando si prende in considerazione il microbursting. Questo parametro è disponibile anche per singoli nodi. Il microbursting si verifica quando un volume EBS raggiunge livelli elevati di IOPS o di throughput per periodi di tempo significativamente più brevi (meno di un minuto). Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)

Metrica	Description
`WriteThroughput`	Il throughput, in byte al secondo, per le operazioni di scrittura sui volumi di EBS. Questo parametro è disponibile anche per singoli nodi. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`WriteThroughputMicroBursting`	La velocità effettiva, in byte al secondo, per le operazioni di scrittura su volumi EBS quando si prende in considerazione il microbursting. Questo parametro è disponibile anche per singoli nodi. Il microbursting si verifica quando un volume EBS raggiunge livelli elevati di IOPS o di throughput per periodi di tempo significativamente più brevi (meno di un minuto). Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`DiskQueueDepth`	Il numero di richieste di I/O in sospeso per un volume di EBS. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`ReadIOPS`	Il numero di operazioni I/O al secondo per le operazioni di lettura sui volumi di EBS. Questo parametro è disponibile anche per singoli nodi. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`ReadIOPSMicroBursting`	Il numero di operazioni di input e output (I/O) al secondo per le operazioni di lettura sui volumi EBS quando si prende in considerazione il microbursting. Questo parametro è disponibile anche per singoli nodi. Il microbursting si verifica quando un volume EBS raggiunge livelli elevati di IOPS o di throughput per periodi di tempo significativamente più brevi (meno di un minuto). Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)

Metrica	Description
`WriteIOPS`	Il numero di operazioni I/O al secondo per le operazioni di scrittura sui volumi di EBS. Questo parametro è disponibile anche per singoli nodi. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`WriteIOPS MicroBursting`	Il numero di operazioni di input e output (I/O) al secondo per le operazioni di scrittura su volumi EBS quando si prende in considerazione il microbursting. Questo parametro è disponibile anche per singoli nodi. Il microbursting si verifica quando un volume EBS raggiunge livelli elevati di IOPS o di throughput per periodi di tempo significativamente più brevi (meno di un minuto). Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`BurstBalance`	La percentuale di crediti di input e output (I/O) che rimangono nel bucket burst per un volume EBS. Un valore pari a 100 indica che il volume ha accumulato il numero massimo di crediti. Se questa percentuale scende al di sotto del 70%, consulta the section called “Saldo di burst EBS basso”. Il saldo di espansione rimane a 0 per i domini con tipi di volumi gp3 e i domini con volumi gp2 con una dimensione del volume superiore a 1.000 GiB. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)
`VolumeStalledIOcheck`	Lo stato dei volumi EBS per determinare quando sono compromessi. La metrica è un valore binario che restituisce lo stato 0 (passaggio) o 1 (errore) a seconda che il volume EBS sia in grado di completare le operazioni di input e output. `VolumeStalledIOcheck` è disponibile anche per singoli nodi. Statistiche rilevanti: Minimum, Maximum, Average (Minimo, Massimo, Medio)

Parametri dell'istanza

Amazon OpenSearch Service fornisce le seguenti metriche per ogni istanza in un dominio. OpenSearch Il servizio aggrega inoltre questi parametri delle istanze per fornire informazioni sullo stato generale del cluster. È possibile verificare questo comportamento utilizzando la statistica Conteggio del campione nella console. Nota che ogni parametro nella tabella seguente dispone di statistiche rilevanti per il nodo e il cluster.

> **Important**
>
> Versioni diverse di Elasticsearch utilizzano pool di thread diversi per elaborare le chiamate all'API `_index`. Elasticsearch 1.5 e 2.3 utilizzano il pool di thread di indice. Elasticsearch 5.x, 6.0 e 6.2 utilizzano il pool di thread in blocco. OpenSearch e Elasticsearch 6.3 e versioni successive utilizzano il pool di thread di scrittura. Attualmente, la console OpenSearch di servizio non include un grafico per il pool di thread in blocco.
> Utilizzare `GET _cluster/settings?include_defaults=true` per controllare le dimensioni del pool di thread e della coda per il cluster.

Metrica	Description
`FetchLatency`	La differenza nel tempo totale, in millisecondi, rilevata da tutte le operazioni di shard fetch in un nodo tra il minuto N e il minuto (N - 1). Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima
`FetchRate`	Il numero totale di operazioni di shard fetch al minuto per tutti gli shard su un nodo di dati. Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima, Somma
`ScrollTotal`	Il numero totale di operazioni di shard scroll al minuto per tutti gli shard su un nodo di dati.

Metrica	Description
	Statistiche rilevanti sui nodi: media, massima Statistiche del cluster rilevanti: Media, Massima, Somma
`ScrollCurrent`	Il numero di operazioni di shard scroll attualmente in esecuzione. Statistiche rilevanti sui nodi: media, massima Statistiche del cluster rilevanti: Media, Massima, Somma
`OpenContexts`	Il numero di contesti di ricerca aperti. Statistiche rilevanti sui nodi: media, massima Statistiche del cluster rilevanti: Media, Massima, Somma
`ThreadCount`	Il numero totale di thread attualmente utilizzati dal OpenSearch processo. Statistiche rilevanti sui nodi: media, massima Statistiche del cluster rilevanti: Media, Massima, Somma
`ShardReactivateCount`	Il numero totale di volte in cui tutti gli shard sono stati attivati da uno stato di inattività. Statistiche relative ai nodi: Sum, Maximum Statistiche pertinenti sui cluster: somma, massima
`ConcurrentSearchRate`	Il numero totale di richieste di ricerca che utilizzano la ricerca simultanea per segmenti al minuto per tutti gli shard su un nodo di dati. Una singola chiamata all'API `_search` potrebbe restituire risultati da diverse partizioni. Se cinque di queste partizioni si trovano in un solo nodo, il nodo indicherà 5 per questo parametro, anche se il client ha effettuato una sola richiesta. Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima, Somma

Metrica	Description
`ConcurrentSearchLatency`	La differenza nel tempo totale, in millisecondi, impiegato da tutte le ricerche utilizzando la ricerca simultanea per segmenti in un nodo tra il minuto N e il minuto (N-1). Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima
`IndexingLatency`	La differenza nel tempo totale, in millisecondi, rilevata da tutte le operazioni di indicizzazione in un nodo tra il minuto N e il minuto (N-1). Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima
`IndexingRate`	Il numero di operazioni di indicizzazione al minuto. Una singola chiamata all'API `_bulk` che aggiunge due documenti e aggiorna due conteggi come quattro operazioni, che possono essere diffuse in uno o più nodi. Se tale indice ha una o più repliche e si trova in un OpenSearch dominio senza istanze ottimizzate, anche gli altri nodi del cluster registrano un totale di quattro operazioni di indicizzazione. Per i OpenSearch domini con istanze ottimizza te, gli altri nodi con repliche non registrano alcuna operazione. L'eliminazione di documenti non viene conteggiata ai fini di questo parametro. Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima, Somma
`SearchLatency`	La differenza nel tempo totale, in millisecondi, rilevato da tutte le ricerche in un nodo tra il minuto N e il minuto (N-1). Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima

Metrica	Description
`SearchRate`	Il numero totale di richieste di ricerca al minuto per tutte le partizioni in un nodo di dati. Una singola chiamata all'API `_search` potrebbe restituire risultati da diverse partizioni. Se cinque di queste partizioni si trovano in un solo nodo, il nodo indicherà 5 per questo parametro, anche se il client ha effettuato una sola richiesta. Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima, Somma
`SegmentCount`	Il numero di segmenti in un nodo di dati. Più segmenti hai, più tempo impiega ogni ricerca. OpenSearch occasionalmente unisce segmenti più piccoli in segmenti più grandi. Statistiche nodo rilevanti: Massima, Media Statistiche del cluster rilevanti: Somma, Massimo, Media
`SysMemoryUtilization`	La percentuale di memoria dell'istanza utilizzata. I valori elevati per questa metrica sono normali e in genere non rappresentano un problema con il cluster. Per un migliore indicatore dei potenziali problemi di prestazioni e stabilità, vedere la metrica `JVMMemoryPressure` . Statistiche di nodo rilevanti: Minima, Massima, Media Statistiche del cluster rilevanti: Minima, Massima, Media
`JVMGCYoungCollectionCount`	Il numero di volte in cui è stata eseguita la garbage collection di "nuova generazione". Un numero elevato e in continua crescita di esecuzioni è una parte normale delle operazioni del cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media

Metrica	Description
JVMGCYoungCollectionTime	La quantità di tempo, in millisecondi, che il cluster ha impiegato per eseguire la garbage collection di "nuova generazione". Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
JVMGCOldCollectionCount	Il numero di volte in cui è stata eseguita la garbage collection "vecchia generazione". In un cluster con risorse sufficienti, questo numero deve rimanere basso e senza frequenti incrementi. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
JVMGCOldCollectionTime	La quantità di tempo, in millisecondi, che il cluster ha impiegato per eseguire la garbage collection "vecchia generazione". Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
OpenSearchDashboardsConcurrentConnections	Il numero di connessioni simultanee attive alle dashboard. OpenSearch Se questo numero cresce costantemente, valutare la possibilità di dimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
OpenSearchDashboardsHealthyNode	Un controllo di integrità per il singolo nodo OpenSearch Dashboards. Un valore pari a 1 indica un comportamento normale. Un valore pari a 0 indica che Dashboards non è accessibile. Statistiche nodo rilevanti: Minima Statistiche del cluster rilevanti: Minima, Massima, Media

Metrica	Description
OpenSearchDashboardsHeapTotal	La quantità di memoria heap allocata alle OpenSearch dashboard in MiB. Diversi tipi di istanza EC2 possono influire sull'esatta allocazione della memoria. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
OpenSearchDashboardsHeapUsed	La quantità assoluta di memoria heap utilizzata dai OpenSearch dashboard in MiB. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
OpenSearchDashboardsHeapUtilization	La percentuale massima di memoria heap disponibile utilizzata dai dashboard. OpenSearch Se questo valore supera l'80%, valutare la possibilità di dimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Minima, Massima, Media
OpenSearchDashboardsOS1MinuteLoad	La media di carico della CPU in un minuto per le dashboard. OpenSearch Il carico della CPU dovrebbe idealmente rimanere al di sotto di 1. Mentre i picchi temporanei vanno bene, se questo parametro è costantemente superiore a 1 si consiglia di aumentare la dimensione del tipo di istanza. Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima

Metrica	Description
`OpenSearchDashboardsRequestTotal`	Il numero totale di richieste HTTP inviate alle OpenSearch dashboard. Se il sistema è lento o viene visualizzato un numero elevato di richieste Dashboards, è consigliabile aumentare le dimensioni del tipo di istanza. Statistiche del nodo rilevanti: Somma Statistiche del cluster rilevanti: Somma
`OpenSearchDashboardsResponseTimesMaxInMillis`	Il tempo massimo, in millisecondi, impiegato dai OpenSearch dashboard per rispondere a una richiesta. Se le richieste richiedon o molto tempo per restituire i risultati, è consigliabile aumentare le dimensioni del tipo di istanza. Statistiche di nodo rilevanti: Massima Statistiche cluster rilevanti: Massima, Media
`SearchTaskCancelled`	Il numero di cancellazioni del nodo coordinatore. Statistiche del nodo rilevanti: Somma Statistiche del cluster rilevanti: Somma
`SearchShardTaskCancelled`	Il numero di cancellazioni dei nodi dati. Statistiche del nodo rilevanti: Somma Statistiche pertinenti sui cluster: somma,
`ThreadpoolForce_mergeQueue`	Il numero di attività in coda nel pool di thread forza unione. Se la dimensione della coda è costantemente elevata, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media

Metrica	Description
`ThreadpoolForce_mergeRejected`	Il numero di attività rifiutate nel pool di thread forza unione. Se questo numero cresce costantemente, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma
`ThreadpoolForce_mergeThreads`	Le dimensioni del pool di thread forza unione. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
`ThreadpoolIndexQueue`	Il numero di attività in coda nel pool di thread di indice. Se la dimensione della coda è costantemente elevata, valutare la possibilità di ridimensionare il cluster. La dimensione massima della coda dell'indice è di 200. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
`ThreadpoolIndexRejected`	Il numero di attività rifiutate nel pool di thread di indice. Se questo numero cresce costantemente, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma
`ThreadpoolIndexThreads`	Le dimensioni del pool di thread di indice. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma

Metrica	Description
ThreadpoolSearchQueue	Il numero di attività in coda nel pool di thread di ricerca. Se la dimensione della coda è costantemente elevata, valutare la possibilità di ridimensionare il cluster. La dimensione massima della coda di ricerca è di 1.000. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
ThreadpoolSearchRejected	Il numero di attività rifiutate nel pool di thread di ricerca. Se questo numero cresce costantemente, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma
ThreadpoolSearchThreads	Le dimensioni del pool di thread di ricerca. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
Threadpoolsql-workerQueue	Il numero di attività in coda nel pool di thread di ricerca SQL. Se la dimensione della coda è costantemente elevata, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
Threadpoolsql-workerRejected	Il numero di attività rifiutate nel pool di thread di ricerca SQL. Se questo numero cresce costantemente, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma

Metrica	Description
`Threadpoolsql-workerThreads`	Le dimensioni del pool di thread di ricerca SQL. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
`ThreadpoolBulkQueue`	Il numero di attività in coda nel pool di thread blocco. Se la dimensione della coda è costantemente elevata, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
`ThreadpoolBulkRejected`	Il numero di attività rifiutate nel pool di thread blocco. Se questo numero cresce costantemente, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma
`ThreadpoolBulkThreads`	Le dimensioni del pool di thread blocco. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
`ThreadpoolIndexSearcherQueue`	Il numero di attività in coda nel pool di thread di index searcher. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
`ThreadpoolIndexSearcherRejected`	Il numero di attività rifiutate nel pool di thread di index searcher. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma

Metrica	Description
`ThreadpoolIndexSearcherThreads`	La dimensione del pool di thread di Index Searcher. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
`ThreadpoolWriteThreads`	La dimensione del pool di thread di scrittura. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
`ThreadpoolWriteQueue`	Il numero di attività in coda nel pool di thread di scrittura. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
`ThreadpoolWriteRejected`	Il numero di attività rifiutate nel pool di thread di scrittura. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma **Note** Poiché la dimensione predefinita della coda di scrittura è stata aumentata da 200 a 10000 nella versione 7.1, questa metrica non è più l'unico indicatore dei rifiuti da parte del Servizio. OpenSearch Utilizzare i parametri `CoordinatingWriteRejected` , `PrimaryWriteRejected` e `ReplicaWriteRejected` per monitorare i rifiuti nelle versioni 7.1 e successive.

Metrica	Description
`CoordinatingWriteRejected`	Il numero totale di rifiuti si è verificato sul nodo di coordinamento a causa della pressione di indicizzazione dall'ultimo avvio del processo di servizio. OpenSearch Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma Questo parametro è disponibile nella versione 7.1 e nelle versioni successive.
`PrimaryWriteRejected`	Il numero totale di rifiuti si è verificato sugli shard primari a causa della pressione di indicizzazione dall'ultimo avvio del processo di servizio. OpenSearch Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma Questo parametro è disponibile nella versione 7.1 e nelle versioni successive.
`ReplicaWriteRejected`	Il numero totale di rifiuti si è verificato sugli shard di replica a causa della pressione di indicizzazione dall'ultimo avvio del processo di servizio. OpenSearch Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma Questo parametro è disponibile nella versione 7.1 e nelle versioni successive.

Metrica	Description
`WorkloadManagementEnabled`	Indica se la funzionalità di gestione del carico di lavoro è abilitata. Il valore 1 indica che è abilitata e il valore 0 indica che è `monitor_only` disattivata. Statistiche relative ai nodi: massimo, minimo Statistiche del cluster rilevanti: Media, Somma Questo parametro è disponibile nella versione 7.1 e nelle versioni successive.
`SoftQueryGroupCount`	Numero di gruppi di query in modalità soft nel dominio. Statistiche rilevanti sui nodi: media, massima Statistiche del cluster rilevanti: Media, Massima, Somma Questo parametro è disponibile nella versione 7.1 e nelle versioni successive.
`EnforcedQueryGroupCount`	Numero di gruppi di query in modalità applicata nel dominio. Statistiche rilevanti sui nodi: media, massima Statistiche del cluster rilevanti: Media, Massima, Somma Questo parametro è disponibile nella versione 7.1 e nelle versioni successive.

Metriche calde

Amazon OpenSearch Service fornisce i seguenti parametri per l'architettura di storage a più livelli e UltraWarm

> **Note**
>
> Le metriche relative all'indicizzazione Warm sono applicabili solo all'architettura di storage a più livelli

Metrica	Description
`WarmIndexingLatency`	La differenza nel tempo totale, in millisecondi, rilevata da tutte le operazioni di indicizzazione in un nodo caldo tra i minuti N e i minuti (N-1). Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima
`WarmIndexingRate`	Il numero di operazioni di indicizzazione a caldo al minuto. Una singola chiamata all'API `_bulk` che aggiunge due documenti e aggiorna due conteggi come quattro operazioni, che possono essere diffuse in uno o più nodi. Se tale indice ha una o più repliche e si trova in un OpenSearch dominio senza istanze ottimizzate, anche gli altri nodi del cluster registrano un totale di quattro operazioni di indicizzazione. Per i OpenSearch domini con istanze ottimizzate, gli altri nodi con repliche non registrano alcuna operazione. L'eliminazione di documenti non viene conteggiata ai fini di questo parametro. Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima, Somma
`WarmThreadpoolIndexingQueue`	Il numero di attività in coda nel pool di thread di indice. Se la dimensione della coda è costantemente elevata, valutare la possibilità di ridimensionare il cluster. La dimensione massima della coda dell'indice è di 200. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Massima, Somma

Metrica	Description
`WarmThreadpoolIndexingRejected`	Il numero di attività rifiutate nel pool di thread di indice. Se questo numero cresce costantemente, valutare la possibilità di ridimensionare il cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma
`WarmThreadpoolIndexingThreads`	Le dimensioni del pool di thread di indice. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
`WarmCPUUtilization`	La percentuale di utilizzo della CPU per i nodi Warm nel cluster. Il numero massimo mostra il nodo con il più alto utilizzo della CPU. La media rappresenta tutti i nodi Warm del cluster. Questa metrica è disponibile anche per i singoli nodi Warm. Statistiche rilevanti: Maximum (Massimo), Average (Media)
`WarmFreeStorageSpace`	La quantità di spazio di archiviazione a caldo gratuito in MiB. Poiché Warm utilizza Amazon S3 anziché dischi collegati, `Sum` è l'unica statistic a rilevante. È necessario lasciare il periodo su un minuto per ottenere un valore preciso. Statistiche rilevanti: Sum (Somma)
`WarmSearchableDocuments`	Il numero totale di documenti disponibili per la ricerca tra tutti gli indici a caldo nel cluster. È necessario lasciare il periodo su un minuto per ottenere un valore preciso. Statistiche rilevanti: Sum (Somma)

Metrica	Description
`WarmSearchLatency`	La differenza nel tempo totale, in millisecondi, rilevato da tutte le ricerche in un Warm tra il minuto N e il minuto (N-1). Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima
`WarmSearchRate`	Il numero totale di richieste di ricerca al minuto per tutti gli shard su un nodo Warm. Una singola chiamata all'API `_search` potrebbe restituire risultati da diverse partizioni. Se cinque di queste partizioni si trovano in un solo nodo, il nodo indicherà 5 per questo parametro, anche se il client ha effettuato una sola richiesta. Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Media, Massima, Somma
`WarmStorageSpaceUtilization`	La quantità totale di spazio di archiviazione a caldo, in MiB, che sta utilizzando il cluster. Statistiche rilevanti: Massima
`HotStorageSpaceUtilization`	La quantità totale di spazio di archiviazione ad accesso frequente utilizzata dal cluster. Statistiche rilevanti: Massima
`WarmSysMemoryUtilization`	La percentuale di memoria del nodo Warm utilizzata. Statistiche rilevanti: Massima
`HotToWarmMigrationQueueSize`	Il numero di indici attualmente in attesa di migrazione dall'archiviazione ad accesso frequente a quella a caldo. Statistiche rilevanti: Massima

Metrica	Description
`WarmToHot Migration QueueSize`	Il numero di indici attualmente in attesa di migrazione dall'archiviazione a caldo a quella ad accesso frequente. Statistiche rilevanti: Massima
`HotToWarm Migration FailureCount`	Il numero totale di migrazioni da "ad accesso frequente" a "a caldo" non riuscite. Questa metrica è disponibile solo per i UltraWarm nodi. Statistiche rilevanti: Sum (Somma)
`HotToWarm Migration ForceMerg eLatency`	La latenza media della fase di unione forzata del processo di migrazion e. Se questa fase richiede sempre troppo tempo, valuta la possibilità di aumentarla. Questa metrica è disponibile solo per UltraWarm i nodi. `index.ultrawarm.migration.force_merge.max_num _segments` . Statistiche rilevanti: Average (Media)
`HotToWarm Migration SnapshotL atency`	La latenza media della fase di snapshot del processo di migrazione. Se questa fase richiede troppo tempo, assicurarsi che le partizioni siano dimensionate e distribuite in modo appropriato in tutto il cluster. Questa metrica è disponibile solo per UltraWarm i nodi. Statistiche rilevanti: Average (Media)
`HotToWarm Migration Processin gLatency`	La latenza media delle migrazioni riuscite da "ad accesso frequente " a "a caldo", senza includere il tempo trascorso nella coda. Questo valore è la somma del tempo necessario per completare le fasi di unione forzata, snapshot e rilocazione delle partizioni del processo di migrazion e. Questa metrica è disponibile solo per i UltraWarm nodi. Statistiche rilevanti: Average (Media)
`HotToWarm Migration SuccessCount`	Il numero totale di migrazioni riuscite da "ad accesso frequente" a "a caldo". Statistiche rilevanti: Sum (Somma)

Metrica	Description
`HotToWarmMigrationSuccessLatency`	La latenza media delle migrazioni riuscite da "ad accesso frequente" a "a caldo", compreso il tempo trascorso nella coda. Statistiche rilevanti: Average (Media)
`WarmThreadpoolSearchThreads`	La dimensione del pool di thread di ricerca di Warm. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Media, Somma
`WarmThreadpoolSearchRejected`	Il numero di attività rifiutate nel pool di thread di ricerca di Warm. Se questo numero aumenta continuamente, valuta la possibilità di aggiungere altri nodi Warm. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma
`WarmThreadpoolSearchQueue`	Il numero di attività in coda nel pool di thread di ricerca di Warm. Se la dimensione della coda è costantemente elevata, valuta la possibilità di aggiungere altri nodi Warm. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
`WarmJVMMemoryPressure`	La percentuale massima dell'heap Java utilizzata per i nodi Warm. Statistiche rilevanti: Massima **Note** La logica di questo parametro è cambiata nel software del servizio R20220323. Per ulteriori informazioni, consulta le note di rilascio.

Metrica	Description
`WarmOldGenJVMMemoryPressure`	La percentuale massima dell'heap Java utilizzata per la «vecchia generazione» per nodo Warm. Statistiche rilevanti: Massima
`WarmJVMGCYoungCollectionCount`	Il numero di volte in cui la raccolta dei rifiuti delle «giovani generazioni» è stata eseguita sui nodi Warm. Un numero elevato e in continua crescita di esecuzioni è una parte normale delle operazioni del cluster. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
`WarmJVMGCYoungCollectionTime`	La quantità di tempo, in millisecondi, impiegata dal cluster per eseguire la raccolta dei rifiuti di «nuova generazione» sui nodi Warm. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
`WarmJVMGCOldCollectionCount`	Il numero di volte in cui la raccolta dei rifiuti di «vecchia generazione» è stata eseguita sui nodi Warm. In un cluster con risorse sufficienti, questo numero deve rimanere basso e senza frequenti incrementi. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
`WarmConcurrentSearchRate`	Il numero totale di richieste di ricerca che utilizzano la ricerca simultane a per segmenti al minuto per tutti gli shard su un nodo Warm. Una singola chiamata all'API `_search` potrebbe restituire risultati da diverse partizioni. Se cinque di queste partizioni si trovano in un solo nodo, il nodo indicherà 5 per questo parametro, anche se il client ha effettuato una sola richiesta. Statistiche di nodo rilevanti: Media Statistiche del cluster rilevanti: Somma, Massimo, Media

Metrica	Description
`WarmConcurrentSearchLatency`	La differenza nel tempo totale, in millisecondi, impiegato da tutte le ricerche utilizzando la ricerca simultanea per segmenti in un nodo Warm tra il minuto N e il minuto (N-1). Statistiche di nodo rilevanti: Media Statistiche cluster rilevanti: Massima, Media
`WarmThreadpoolIndexSearcherQueue`	Il numero di attività in coda nel pool di thread di Warm index searcher. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma, Massimo, Media
`WarmThreadpoolIndexSearcherRejected`	Il numero di attività rifiutate nel pool di thread di Warm index searcher. Statistiche di nodo rilevanti: Massima Statistiche del cluster rilevanti: Somma
`WarmThreadpoolIndexSearcherThreads`	La dimensione del pool di thread di Warm index searcher. Statistiche di nodo rilevanti: Massima Statistiche pertinenti sui cluster: somma, media

Parametri di archiviazione a freddo

Amazon OpenSearch Service fornisce le seguenti metriche per la conservazione a freddo.

Metrica	Description
`ColdStorageSpaceUtilization`	La quantità totale di spazio di archiviazione a freddo, in MiB, utilizzato dal cluster. Statistiche rilevanti: Max (Massimo)

Metrica	Description
`ColdToWarmMigrationFailureCount`	Il numero totale di migrazioni da freddo a caldo non riuscite. Statistiche rilevanti: Sum (Somma)
`ColdToWarmMigrationLatency`	Il tempo necessario per completare le migrazioni da freddo a caldo riuscite. Statistiche rilevanti: Average (Media)
`ColdToWarmMigrationQueueSize`	Il numero di indici attualmente in attesa di migrazione dall'archiviazione a freddo a quella a caldo. Statistiche rilevanti: Massima
`ColdToWarmMigrationSuccessCount`	Il numero totale di migrazioni da freddo a caldo riuscite. Statistiche rilevanti: Sum (Somma)
`WarmToColdMigrationFailureCount`	Il numero totale di migrazioni da caldo a freddo non riuscite. Statistiche rilevanti: Sum (Somma)
`WarmToColdMigrationLatency`	Il tempo necessario per completare le migrazioni da caldo a freddo riuscite. Statistiche rilevanti: Average (Media)
`WarmToColdMigrationQueueSize`	Il numero di indici attualmente in attesa di migrazione dall'archiviazione a caldo a quella a freddo. Statistiche rilevanti: Massima
`WarmToColdMigrationSuccessCount`	Il numero totale di migrazioni da caldo a freddo riuscite. Statistiche rilevanti: Sum (Somma)

OpenSearch Metriche ottimizzate per le istanze () OR1

Amazon OpenSearch Service fornisce le seguenti metriche per le OR1istanze.

Metrica	Description
`RemoteStorageUsedSpace`	La quantità totale di spazio Amazon S3, in MiB, utilizzata dal cluster. Statistiche rilevanti: Sum (Somma)
`RemoteStorageWriteRejected`	Il numero totale di richieste rifiutate sugli shard primari a causa della pressione di storage e replica remoti. Viene calcolato a partire dall'ultimo avvio del processo OpenSearch di servizio. Statistiche rilevanti: Sum (Somma)
`ReplicationLagMaxTime`	La quantità di tempo, in millisecondi, in cui i frammenti di replica restano indietro rispetto agli shard primari. Statistiche rilevanti: Massima

Parametri di avvisi

Amazon OpenSearch Service fornisce le seguenti metriche per gli avvisi.

Metrica	Description
`AlertingDegraded`	Il valore 1 indica che l'indice di allerta è rosso oppure uno o più nodi non sono pianificati. Un valore 0 indica un comportamento normale. Statistiche rilevanti: Massima
`AlertingIndexExists`	Un valore pari a 1 significa che l'indice `.opensearch-alerting-config` esiste. Un valore pari a 0 significa che non esiste. Fino a quando non si utilizza la funzione di allarme per la prima volta, questo valore rimane 0. Statistiche rilevanti: Massima
`AlertingIndexStatus.green`	La salute dell'indice. Un valore pari a 1 significa verde. Un valore pari a 0 significa che l'indice non esiste o non è verde.

Metrica	Description
	Statistiche rilevanti: Massima
`AlertingIndexStatus.red`	La salute dell'indice. Un valore pari a 1 significa rosso. Un valore pari a 0 significa che l'indice non esiste o non è rosso. Statistiche rilevanti: Massima
`AlertingIndexStatus.yellow`	La salute dell'indice. Un valore pari a 1 significa giallo. Un valore pari a 0 significa che l'indice non esiste o non è giallo. Statistiche rilevanti: Massima
`AlertingNodesNotOnSchedule`	Il valore 1 indica che alcuni processi non sono in esecuzione nei tempi previsti. Il valore 0 indica che tutti i processi di allerta sono in esecuzione nella pianificazione (o che non esistono processi di avvisi). Controlla la console OpenSearch di servizio o fai una `_nodes/stats` richiesta per vedere se alcuni nodi mostrano un elevato utilizzo delle risorse. Statistiche rilevanti: Massima
`AlertingNodesOnSchedule`	Il valore 1 indica che tutti i processi di allerta sono in esecuzione nella pianificazione (o che non esistono processi di avvisi). Un valore pari a 0 indica che alcuni processi non sono in esecuzione nella pianificazione. Statistiche rilevanti: Massima
`AlertingScheduledJobEnabled`	Il valore 1 indica che l'impostazione del cluster `opensearch.scheduled_jobs.enabled` è true. Il valore 0 indica che è falsa e i processi pianificati sono disabilitati. Statistiche rilevanti: Massima

Parametri di rilevamento delle anomalie

Amazon OpenSearch Service fornisce le seguenti metriche per il rilevamento delle anomalie.

Metrica	Description
`ADPluginUnhealthy`	Il valore 1 indica che il plug-in di rilevamento delle anomalie non funziona correttamente, a causa di un numero elevato di errori o perché uno degli indici utilizzati è rosso. Il valore 0 indica che il plugin funziona come previsto. Statistiche rilevanti: Massima
`ADExecuteRequestCount`	Numero di richieste per il rilevamento delle anomalie. Statistiche rilevanti: Sum (Somma)
`ADExecuteFailureCount`	Numero di richieste non riuscite per il rilevamento delle anomalie. Statistiche rilevanti: Sum (Somma)
`ADHCExecuteFailureCount`	Il numero di richieste non riuscite per il rilevamento delle anomalie per i rilevatori ad alta cardinalità. Statistiche rilevanti: Sum (Somma)
`ADHCExecuteRequestCount`	Il numero di richieste per il rilevamento delle anomalie per i rilevatori ad alta cardinalità. Statistiche rilevanti: Sum (Somma)
`ADAnomalyResultsIndexStatusIndexExists`	Il valore 1 indica che l'indice a cui punta l'alias `.opensearch-anomaly-results` esiste. Fino a quando non si utilizza la funzionalità di rilevamento delle anomalie per la prima volta, questo valore rimane 0. Statistiche rilevanti: Massima
`ADAnomalyResultsIndexStatus.red`	Il valore 1 indica che l'indice a cui punta l'alias `.opensearch-anomaly-results` è rosso. Un valore pari a 0 significa che non lo è. Fino a quando non si utilizza la funzionalità di rilevamento delle anomalie per la prima volta, questo valore rimane 0. Statistiche rilevanti: Massima

Metrica	Description
`ADAnomaly Detectors IndexStat usIndexExists`	Un valore pari a 1 significa che l'indice `.opensearch-anomaly-detectors` esiste. Un valore pari a 0 significa che non esiste. Fino a quando non si utilizza la funzionalità di rilevamento delle anomalie per la prima volta, questo valore rimane 0. Statistiche rilevanti: Massima
`ADAnomaly Detectors IndexStat us.red`	Un valore pari a 1 indica che l'indice `.opensearch-anomaly-detectors` è rosso. Un valore pari a 0 significa che non lo è. Fino a quando non si utilizza la funzionalità di rilevamento delle anomalie per la prima volta, questo valore rimane 0. Statistiche rilevanti: Massima
`ADModelsC heckpoint IndexStat usIndexExists`	Un valore pari a 1 significa che l'indice `.opensearch-anomaly-checkpoints` esiste. Un valore pari a 0 significa che non esiste. Fino a quando non si utilizza la funzionalità di rilevamento delle anomalie per la prima volta, questo valore rimane 0. Statistiche rilevanti: Massima
`ADModelsC heckpoint IndexStat us.red`	Un valore pari a 1 indica che l'indice `.opensearch-anomaly-checkpoints` è rosso. Un valore pari a 0 significa che non lo è. Fino a quando non si utilizza la funzionalità di rilevamento delle anomalie per la prima volta, questo valore rimane 0. Statistiche rilevanti: Massima

Parametri di ricerca asincrona

Amazon OpenSearch Service fornisce le seguenti metriche per la ricerca asincrona.

Statistiche del nodo coordinatore di ricerca asincrona (per nodo coordinatore)

Metrica	Description
`AsynchronousSearchSubmissionRate`	Il numero di ricerche asincrone inviate nell'ultimo minuto.
`AsynchronousSearchInitializedRate`	Il numero di ricerche asincrone inizializzate nell'ultimo minuto.
`AsynchronousSearchRunningCurrent`	Il numero di ricerche asincrone correntemente in esecuzione.
`AsynchronousSearchCompletionRate`	Il numero di ricerche asincrone completate correttamente nell'ultimo minuto.
`AsynchronousSearchFailureRate`	Il numero di ricerche asincrone completate e non riuscite nell'ultimo minuto.
`AsynchronousSearchPersistRate`	Il numero di ricerche asincrone conservate nell'ultimo minuto.
`AsynchronousSearchPersistFailedRate`	Il numero di ricerche asincrone che non sono state conservate nell'ultimo minuto.
`AsynchronousSearchRejected`	Il numero totale di ricerche asincrone rifiutate dall'attivazione del nodo.

Metrica	Description
`AsynchronousSearchCancelled`	Il numero totale di ricerche asincrone cancellate dall'attivazione del nodo.
`AsynchronousSearchMaxRunningTime`	La durata della ricerca asincrona più lunga in esecuzione su un nodo nell'ultimo minuto.

Statistiche del cluster di ricerca asincrona

Metrica	Description
`AsynchronousSearchStoreHealth`	Lo stato dell'archiviazione nell'indice persistente (rosso/non rosso) nell'ultimo minuto.
`AsynchronousSearchStoreSize`	La dimensione dell'indice di sistema su tutte le partizioni nell'ultimo minuto.
`AsynchronousSearchStoredResponseCount`	Il numero di risposte memorizzate nell'indice di sistema nell'ultimo minuto.

Metriche Auto-Tune

Amazon OpenSearch Service fornisce le seguenti metriche per Auto-Tune.

Metrica	Description
`AutoTuneChangesHistoryHeapSize`	La cronologia delle modifiche in MiB per i valori di ottimizzazione della dimensione dell'heap.

Metrica	Description
`AutoTuneChangesHistoryJVMYoungGenArgs`	La cronologia delle modifiche per gli argomenti JVM. YongGen
`AutoTuneFailed`	Un valore booleano che indica se la modifica Auto-Tune non è riuscita.
`AutoTuneSucceeded`	Un valore booleano che indica se la modifica Auto-Tune è stata completata.
`AutoTuneValue`	La cronologia delle modifiche alla coda (count) e le ottimizzazioni della cache modificano la cronologia (in MiB) per modifiche senza interruzioni.

Multi-AZ con metriche Standby

Amazon OpenSearch Service fornisce le seguenti metriche per Multi-AZ with Standby.

Metriche a livello di nodo per i nodi di dati nelle zone di disponibilità attive

Metrica	Description
`CPUUtilization`	Percentuale di utilizzo della CPU per i nodi di dati nel cluster. Il numero massimo mostra il nodo con il più alto utilizzo della CPU. La media rappresenta tutti i nodi del cluster. Questo parametro è disponibile anche per singoli nodi.
`FreeStorageSpace`	Lo spazio libero per i nodi di dati nel cluster. `Sum` mostra lo spazio libero totale per il cluster, ma è necessario lasciare il periodo a un minuto per ottenere un valore accurato. `Minimum` e `Maximum` mostrano i nodi con lo spazio libero maggiore e minore, rispettivamente. Questa metrica è disponibile anche per i singoli nodi. OpenSearch Il servizio genera un `ClusterBlockException` quando questa metrica raggiunge. `0` Per risolvere il problema devi eliminare gli indici, aggiunger e istanze più grandi oppure aggiungere archiviazione basata su EBS alle istanze esistenti. Per ulteriori informazioni, consulta the section called “Mancanza di spazio di archiviazione disponibile”.

Metrica	Description
	La console OpenSearch di servizio visualizza questo valore in GiB. La CloudWatch console Amazon lo visualizza in MiB.
`JVMMemory Pressure`	La percentuale massima dell'heap Java utilizzata per tutti i nodi di dati del cluster. OpenSearch Il servizio utilizza metà della RAM di un'istanza per l'heap Java, fino a una dimensione dell'heap di 32 GiB. Puoi scalare le istanze verticalmente fino a 64 GiB di RAM e poi scalare orizzonta lmente aggiungendo le istanze. Per informazioni, consulta the section called “Allarmi consigliati CloudWatch ”.
`SysMemory Utilization`	La percentuale di memoria dell'istanza utilizzata. I valori elevati per questa metrica sono normali e in genere non rappresentano un problema con il cluster. Per un migliore indicatore dei potenziali problemi di prestazioni e stabilità, vedere la metrica `JVMMemoryPressure` .
`IndexingLatency`	La differenza nel tempo totale, in millisecondi, rilevata da tutte le operazioni di indicizzazione in un nodo tra i minuti N e i minuti (N-1).
`IndexingRate`	Il numero di operazioni di indicizzazione al minuto.
`SearchLatency`	La differenza nel tempo totale, in millisecondi, rilevato da tutte le ricerche in un nodo tra il minuto N e il minuto (N-1).
`SearchRate`	Il numero totale di richieste di ricerca al minuto per tutte le partizioni in un nodo di dati.
`Threadpoo lSearchQueue`	Il numero di attività in coda nel pool di thread di ricerca. Se la dimension e della coda è costantemente elevata, valutare la possibilità di ridimensi onare il cluster. La dimensione massima della coda di ricerca è di 1.000.
`Threadpoo lWriteQueue`	Il numero di attività in coda nel pool di thread di scrittura.
`Threadpoo lSearchRe jected`	Il numero di attività rifiutate nel pool di thread di ricerca. Se questo numero cresce costantemente, valutare la possibilità di ridimensionare il cluster.

Metrica	Description
`Threadpool WriteRejected`	Il numero di attività rifiutate nel pool di thread di scrittura.

Metriche a livello di cluster per i cluster nelle zone di disponibilità attive

Metrica	Description
`DataNodes`	Il numero totale di shard attivi e in standby.
`DataNodes Shards.active`	Il numero totale di partizioni primarie e di replica attive.
`DataNodes Shards.unassigned`	Il numero di partizioni non allocate ai nodi nel cluster.
`DataNodes Shards.initializing`	Il numero di partizioni in fase di inizializzazione.
`DataNodes Shards.relocating`	Il numero di partizioni in fase di rilocazione.

Metriche di rotazione della zona di disponibilità

`SeActiveReads.`*`Availability-Zone`* `= 1`, allora la zona è attiva.
`SeActiveReads.`*`Availability-Zone`* `= 0`, allora la zona è in standby.

Metriche puntuali

Amazon OpenSearch Service fornisce le seguenti metriche per le ricerche point-in-time (PIT).

Statistiche sul nodo coordinatore PIT (per nodo coordinatore)

Metrica	Description
`CurrentPointInTime`	Il numero di contesti di ricerca PIT attivi nel nodo.
`TotalPointInTime`	Il numero di contesti di ricerca PIT scaduti dal momento dell'attività del nodo.
`AvgPointInTimeAliveTime`	Il mantenimento medio dei contesti di ricerca PIT dal momento dell'attività del nodo.
`HasActivePointInTime`	Il valore 1 indica che ci sono contesti PIT attivi sui nodi sin dal momento in cui il nodo è attivo. Un valore pari a 0 significa che non ce ne sono.
`HasUsedPointInTime`	Il valore 1 indica che ci sono contesti PIT scaduti sui nodi dal momento in cui il nodo è attivo. Un valore pari a 0 significa che non ce ne sono.

Parametri SQL

Amazon OpenSearch Service fornisce le seguenti metriche per il supporto SQL.

Metrica	Description
`SQLFailedRequestCountByCusErr`	Numero di richieste all'API `_sql` non riuscite a causa di un problema client. Ad esempio, una richiesta potrebbe restituire il codice di stato HTTP 400 a causa di un `IndexNotFoundException` . Statistiche rilevanti: Sum (Somma)
`SQLFailedRequestCountBySysErr`	Numero di richieste all'API `_sql` non riuscite a causa di un problema del server o di una limitazione della funzionalità. Ad esempio, una richiesta potrebbe restituire il codice di stato HTTP 503 a causa di un `VerificationException` . Statistiche rilevanti: Sum (Somma)
`SQLRequestCount`	Il numero di richieste all'API `_sql`.

Metrica	Description
	Statistiche rilevanti: Sum (Somma)
`SQLDefaultCursorRequestCount`	Simile a `SQLRequestCount`, ma conta solo le richieste di impaginazione. Statistiche rilevanti: Sum (Somma)
`SQLUnhealthy`	Un valore pari a 1 indica che, in risposta a determinate richieste, il plugin SQL restituisce 5 codici di rispostaxx o passa query DSL non valida a OpenSearch. Altre richieste dovrebbero continuare ad avere esito positivo. Un valore pari a 0 indica nessun errore recente. Se viene visualizzato un valore sostenuto pari a 1, risolvere i problemi relativi alle richieste che i client stanno facendo al plugin. Statistiche rilevanti: Massima

Parametri k-NN

Amazon OpenSearch Service include le seguenti metriche per il plugin k-Nearest Neighbor (k-NN).

Metrica	Description
`KNNCacheCapacityReached`	Parametro per nodo per stabilire se è stata raggiunta la capacità della cache. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Massima
`KNNCircuitBreakerTriggered`	Parametro per cluster per indicare se l'interruttore è attivato. Se alcuni nodi restituiscono un valore pari a 1 per `KNNCacheCapacityReached`, anche questo valore restituirà 1. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Massima
`KNNEvictionCount`	Parametro per nodo per il numero di grafici rimossi dalla cache a causa di vincoli di memoria o tempo di inattività. Le rimozioni

Metrica	Description
	esplicite che si verificano a causa dell'eliminazione dell'indice non vengono conteggiate. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Sum (Somma)
`KNNGraphIndexErrors`	Parametro per nodo per il numero di richieste da aggiungere il campo `knn_vector` di un documento a un grafico che ha generato un errore. Statistiche rilevanti: Sum (Somma)
`KNNGraphIndexRequests`	Parametro per nodo per il numero di richieste per aggiungere il campo `knn_vector` di un documento a un grafico. Statistiche rilevanti: Sum (Somma)
`KNNGraphMemoryUsage`	Parametro per nodo per la dimensione della cache corrente (dimensione totale di tutti i grafici in memoria) in kilobyte. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Average (Media)
`KNNGraphQueryErrors`	Parametro per nodo per il numero di query del grafico che hanno generato un errore. Statistiche rilevanti: Sum (Somma)
`KNNGraphQueryRequests`	Parametro per nodo per il numero di query del grafico. Statistiche rilevanti: Sum (Somma)
`KNNHitCount`	Parametro per nodo per il numero di occorrenze della cache. Una occorrenza della cache si verifica quando un utente esegue una query su un grafico già caricato in memoria. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Sum (Somma)

Metrica	Description
KNNLoadExceptionCount	Parametro per nodo per il numero di volte in cui si è verificat a un'eccezione durante il tentativo di caricare un grafico nella cache. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Sum (Somma)
KNNLoadSuccessCount	Parametro per nodo per il numero di volte in cui il plug-in ha caricato correttamente un grafico nella cache. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Sum (Somma)
KNNMissCount	Parametro per nodo per il numero di mancati riscontri nella cache. Un mancato riscontro nella cache si verifica quando un utente esegue una query su un grafico non ancora caricato in memoria. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Sum (Somma)
KNNQueryRequests	Parametro per nodo per il numero di richieste di query ricevute dal plugin k-NN. Statistiche rilevanti: Sum (Somma)
KNNRemoteBuildEnabled	Valore binario che specifica se la funzionalità è abilitata. Statistiche pertinenti: binario
KNNRemoteIndexBuil dFailureCount	Numero totale di errori di compilazione. Statistiche rilevanti: Sum (Somma)
KNNRemoteIndexBuil dSuccessCount	Numero totale di build riuscite. Statistiche rilevanti: Sum (Somma)

Metrica	Description
KNNScriptCompilationErrors	Parametro per nodo per il numero di errori durante la compilazione dello script. Questa statistica è rilevante solo per la ricerca di script di punteggio k-NN. Statistiche rilevanti: Sum (Somma)
KNNScriptCompilations	Parametro per nodo per il numero di volte in cui lo script k-NN è stato compilato. Questo valore dovrebbe in genere essere 1 o 0, ma se la cache contenente gli script compilati viene riempita, lo script k-NN potrebbe essere ricompilato. Questa statistica è rilevante solo per la ricerca di script di punteggio k-NN. Statistiche rilevanti: Sum (Somma)
KNNScriptQueryErrors	Parametro per nodo per il numero di errori durante le query dello script. Questa statistica è rilevante solo per la ricerca di script di punteggio k-NN. Statistiche rilevanti: Sum (Somma)
KNNScriptQueryRequests	Parametro per nodo per il numero totale di query dello script. Questa statistica è rilevante solo per la ricerca di script di punteggio k-NN. Statistiche rilevanti: Sum (Somma)
KNNTotalLoadTime	Il tempo in nanosecondi impiegato da k-NN per caricare i grafici nella cache. Questo parametro è rilevante solo per approssimare la ricerca k-NN. Statistiche rilevanti: Sum (Somma)
VectorIndexBuildAccelerationOCU	Il numero di unità di OpenSearch calcolo (OCUs) utilizzate per accelerare l'indicizzazione vettoriale. Statistiche rilevanti: Sum (Somma)

Parametri di ricerca tra cluster

Amazon OpenSearch Service fornisce le seguenti metriche per la ricerca tra cluster.

Parametri del dominio di origine

Metrica	Dimensione	Description
`CrossClusterOutboundConnections`	`ConnectionId`	Numero di nodi connessi. Se la risposta include uno o più domini ignorati, utilizzare questo parametro per tracciare eventuali connessioni non integre. Se questo numero scende a 0, la connessione non è integra.
`CrossClusterOutboundRequests`	`ConnectionId`	Numero di richieste di ricerca inviate al dominio di destinazione. Utilizzare per verificare se il carico di richieste di ricerca tra cluster sta sovraccaricando il dominio; correlare eventuali picchi in questo parametro con eventuali picchi JVM/CPU.

Parametri del dominio di destinazione

Metrica	Dimensione	Description
`CrossClusterInboundRequests`	`ConnectionId`	Numero di richieste di connessione in ingresso ricevute dal dominio di origine.

Aggiungi un CloudWatch allarme nel caso in cui perdi una connessione in modo imprevisto. Per i passaggi per creare un allarme, vedi Creare un CloudWatch allarme basato su una soglia statica.

Parametri di replica tra cluster

Amazon OpenSearch Service fornisce le seguenti metriche per la replica tra cluster.

Metrica	Description
`ReplicationRate`	La percentuale media di operazioni di replica al secondo. Questo parametro è analogo al parametro `IndexingRate` .
`LeaderCheckPoint`	Per una connessione specifica, la somma dei valori del checkpoint leader in tutti gli indici di replica. Puoi utilizzare questo parametro per misurare la latenza di replica.
`FollowerCheckPoint`	Per una connessione specifica, la somma dei valori del checkpoint follower in tutti gli indici di replica. Puoi utilizzare questo parametro per misurare la latenza di replica.
`ReplicationNumSyncingIndices`	Il numero di indici con uno stato di replica di `SYNCING`.
`ReplicationNumBootstrappingIndices`	Il numero di indici con uno stato di replica di `BOOTSTRAPPING` .
`ReplicationNumPausedIndices`	Il numero di indici con uno stato di replica di `PAUSED`.
`ReplicationNumFailedIndices`	Il numero di indici con uno stato di replica di `FAILED`.
`CrossClusterOutboundReplicationRequests`	Il numero di richieste di trasporto di replica sul dominio del follower. Le richieste di trasporto sono interne e si verificano ogni volta che viene chiamata un'operazione API di replica. Si verificano anche quando il polling del dominio follower cambia rispetto al dominio leader.
`CrossClusterInboun`	Il numero di richieste di trasporto di replica sul dominio leader. Le richieste di trasporto sono interne e si verificano ogni volta che viene chiamata un'operazione API di replica.

Metrica	Description
dReplicationRequests	
AutoFollowNumSuccessStartReplication	Il numero di indici follower creati correttamente da una regola di replica per una connessione specifica.
AutoFollowNumFailedStartReplication	Il numero di indici follower che non sono stati creati da una regola di replica in presenza di un modello corrispondente. Questo problema potrebbe sorgere a causa di un problema di rete sul cluster remoto o di un problema di sicurezza (ad esempio, il ruolo associato non ha l'autorizzazione per avviare la replica).
AutoFollowLeaderCallFailure	Se ci sono state query non riuscite dall'indice follower all'indice leader per estrarre nuovi dati. Un valore pari a 1 significa che ci sono state una o più chiamate non riuscite nell'ultimo minuto.

Parametri di Learning to Rank

Amazon OpenSearch Service fornisce le seguenti metriche per Learning to Rank.

Metrica	Description
LTRRequestTotalCount	Conteggio totale delle richieste di classificazione.
LTRRequestErrorCount	Conteggio totale delle richieste non riuscite.
LTRStatus.red	Traccia se uno degli indici necessari per eseguire il plug-in è rosso.
LTRMemoryUsage	La memoria totale utilizzata dal plug-in.

Metrica	Description
`LTRFeatureMemoryUsageInBytes`	La quantità di memoria, espressa in byte, utilizzata dai campi della funzionalità Learning to Rank.
`LTRFeaturesetMemoryUsageInBytes`	La quantità di memoria, espressa in byte, utilizzata dai set di funzionalità Learning to Rank.
`LTRModelMemoryUsageInBytes`	La quantità di memoria, espressa in byte, utilizzata da tutti i modelli Learning to Rank.

Parametri Piped Processing Language (PPL)

Amazon OpenSearch Service fornisce le seguenti metriche per Piped Processing Language.

Metrica	Description
`PPLFailedRequestCountByCusErr`	Numero di richieste all'API `_ppl` non riuscite a causa di un problema client. Ad esempio, una richiesta potrebbe restituire il codice di stato HTTP 400 a causa di un `IndexNotFoundException` .
`PPLFailedRequestCountBySysErr`	Numero di richieste all'API `_ppl` non riuscite a causa di un problema del server o di una limitazione della funzionalità. Ad esempio, una richiesta potrebbe restituire il codice di stato HTTP 503 a causa di un `VerificationException` .
`PPLRequestCount`	Il numero di richieste all'API `_ppl`.

Monitoraggio dei OpenSearch log con Amazon CloudWatch Logs

Amazon OpenSearch Service espone i seguenti OpenSearch log tramite Amazon CloudWatch Logs:

- Log di errore
- Registri lenti delle richieste di ricerca

- Condividi gli slow log
- Log di verifica

Gli slow log degli shard di ricerca, gli slow log di indicizzazione degli shard e i log degli errori sono utili per la risoluzione dei problemi di prestazioni e stabilità. I log di verifica tengono traccia dell'attività degli utenti a fini di conformità Tutti i log sono disabilitati per impostazione predefinita. Se abilitata, si applica la tariffa standard. CloudWatch

> **Note**
>
> I log degli errori sono disponibili solo per le versioni 5.1 OpenSearch e successive di Elasticsearch. Gli slow log sono disponibili per tutte OpenSearch le versioni e per quelle di Elasticsearch.

Per i suoi log, OpenSearch utilizza Apache Log4j 2 e i suoi livelli di registro incorporati (dal più basso al più severo) di`TRACE`,,, e. `DEBUG` `INFO` `WARN` `ERROR` `FATAL`

Se abiliti i log degli errori, OpenSearch Service pubblica le righe di registro di, e to. `WARN` `ERROR` `FATAL` CloudWatch OpenSearch Il servizio pubblica anche diverse eccezioni dal `DEBUG` livello, tra cui le seguenti:

- `org.opensearch.index.mapper.MapperParsingException`
- `org.opensearch.index.query.QueryShardException`
- `org.opensearch.action.search.SearchPhaseExecutionException`
- `org.opensearch.common.util.concurrent.OpenSearchRejectedExecutionException`
- `java.lang.IllegalArgumentException`

I log di errore possono aiutarti con la risoluzione dei problemi in molti casi, tra cui i seguenti:

- Problemi di compilazione di script Painless
- Query non valide
- Problemi di indicizzazione
- Snapshot non riuscite
- Errori di migrazione di Index State Management

> **Note**
>
> Non tutti gli errori vengono riportati nei log degli errori.

> **Note**
>
> OpenSearch Il servizio non registra tutti gli errori che si verificano.

Argomenti

- Abilitazione della pubblicazione di log (console)
- Abilitazione della pubblicazione di log (AWS CLI)
- Abilitazione della pubblicazione di log (AWS SDKs)
- Abilitazione della pubblicazione di log (CloudFormation)
- Impostazione delle soglie di slow log delle richieste di ricerca
- Impostazione delle soglie di slow log degli shard
- Test degli slow log
- Visualizzazione dei log di

Abilitazione della pubblicazione di log (console)

La console OpenSearch di servizio è il modo più semplice per abilitare la pubblicazione dei log su. CloudWatch

Per abilitare la pubblicazione dei log su CloudWatch (console)

1. Vai su aws.amazon.com, quindi scegli Accedi e fornisci le tue credenziali.
2. In Analytics, scegli Amazon OpenSearch Service.
3. Selezionare il dominio da aggiornare.
4. Nella scheda Log, seleziona un tipo di log e scegli Abilita.
5. Crea un nuovo gruppo di CloudWatch log o scegline uno esistente.

> **Note**
>
> Se si pianifica di abilitare molteplici log, consigliamo di pubblicarli ognuno sul proprio gruppo di log. Questa separazione rende più semplice analizzare i log.

6. Scegliere una policy d'accesso che contiene le autorizzazioni appropriate oppure creare una policy utilizzando il formato JSON che la console offre:

   JSON

   ```
   {
     "Version":"2012-10-17",
     "Statement": [
       {
         "Effect": "Allow",
         "Principal": {
           "Service": "es.amazonaws.com"
         },
         "Action": [
           "logs:PutLogEvents",
           "logs:CreateLogStream"
         ],
         "Resource": "arn:aws:logs:us-east-1:111122223333:log-
   group:cw_log_group_name:*"
       }
     ]
   }
   ```

   Si consiglia di aggiungere le chiavi di condizione `aws:SourceAccount` e `aws:SourceArn` alla policy per proteggersi dal problema del "confused deputy". L'account fonte è il proprietario del flusso di log e l'ARN fonte è l'ARN del dominio. Per aggiungere queste chiavi di condizione, il dominio deve trovarsi sul software di servizio R20211203 o versioni successive.

   Ad esempio, è possibile aggiungere il seguente blocco di condizione alla policy:

   ```
   "Condition": {
       "StringEquals": {
           "aws:SourceAccount": "account-id"
       },
   ```

```
    "ArnLike": {
        "aws:SourceArn": "arn:aws:es:region:account-id:domain/domain-name"
    }
}
```

> **Important**
>
> CloudWatch Logs supporta 10 politiche relative alle risorse per regione. Se si prevede di abilitare i log per diversi domini di OpenSearch servizio, è necessario creare e riutilizzare una politica più ampia che includa più gruppi di log per evitare di raggiungere questo limite. Per le fasi dell'aggiornamento della policy, consulta the section called “Abilitazione della pubblicazione di log (AWS CLI)”.

7. Scegli Abilita .

   Lo stato del dominio cambia da Attivo a Elaborazione. Lo stato deve tornare su Attivo prima che venga abilitata la pubblicazione dei log. Questa modifica richiede in genere 30 minuti, ma può richiedere più tempo a seconda della configurazione del dominio.

Se hai abilitato uno degli shard slow log, vedi. the section called “Impostazione delle soglie di slow log degli shard” Se sono stati abilitati i log di verifica, consultare the section called “Passaggio 2: attiva i registri di controllo nelle dashboard OpenSearch”. Se sono stati abilitati solo i log di errore, non è necessario eseguire operazioni di configurazione aggiuntive.

Abilitazione della pubblicazione di log (AWS CLI)

Prima di poter abilitare la pubblicazione dei log, è necessario un gruppo di CloudWatch log. Se non se ne dispone già di uno, è possibile crearlo utilizzando il comando seguente:

```
aws logs create-log-group --log-group-name my-log-group
```

Immettere il comando successivo per individuare l'ARN del gruppo di log e quindi annotarlo:

```
aws logs describe-log-groups --log-group-name my-log-group
```

Ora puoi concedere al OpenSearch servizio le autorizzazioni di scrittura nel gruppo di log. È necessario indicare l'ARN del gruppo di log vicino al termine del comando:

```
aws logs put-resource-policy \
```

```
  --policy-name my-policy \
  --policy-document '{ "Version": "2012-10-17",        "Statement": [{ "Sid":
 "", "Effect": "Allow", "Principal": { "Service": "es.amazonaws.com"}, "Action":
[ "logs:PutLogEvents","logs:CreateLogStream"],"Resource": "cw_log_group_arn:*"}]}'
```

> **Important**
>
> CloudWatch Logs supporta 10 politiche di risorse per regione. Se prevedi di abilitare gli shard slow log per diversi domini di OpenSearch servizio, dovresti creare e riutilizzare una politica più ampia che includa più gruppi di log per evitare di raggiungere questo limite.

Se è necessario rivedere questa policy in un secondo momento, utilizza il comando `aws logs describe-resource-policies`. Per aggiornare la policy, esegui lo stesso comando `aws logs put-resource-policy` con un nuovo documento di policy.

Infine, è possibile utilizzare l'opzione `--log-publishing-options` per abilitare la pubblicazione. La sintassi dell'opzione è identica per entrambi i comandi `create-domain` e `update-domain-config`.

Parametro	Valori validi
`--log-publishing-options`	`SEARCH_SLOW_LOGS={CloudWatchLogsLogGroupArn= cw_log_group_arn ,Enabled=true|false}`
	`INDEX_SLOW_LOGS={CloudWatchLogsLogGroupArn= cw_log_group_arn ,Enabled=true|false}`
	`ES_APPLICATION_LOGS={CloudWatchLogsLogGroupArn= cw_log_group_arn ,Enabled=true|false}`
	`AUDIT_LOGS={CloudWatchLogsLogGroupArn= cw_log_group_arn ,Enabled=true|false}`

> **Note**
>
> Se si pianifica di abilitare molteplici log, consigliamo di pubblicarli ognuno sul proprio gruppo di log. Questa separazione rende più semplice analizzare i log.

Esempio

L'esempio seguente consente la pubblicazione di slow log degli shard di ricerca e indicizzazione per il dominio specificato:

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --log-publishing-options
 "SEARCH_SLOW_LOGS={CloudWatchLogsLogGroupArn=arn:aws:logs:us-east-1:123456789012:log-group:my-log-
group,Enabled=true},INDEX_SLOW_LOGS={CloudWatchLogsLogGroupArn=arn:aws:logs:us-east-1:123456789012:log-group:my-other-log-group,Enabled=true}"
```

Per disabilitare la pubblicazione su CloudWatch, esegui lo stesso comando con. `Enabled=false`

Se hai abilitato uno degli slow log dello shard, vedi. the section called “Impostazione delle soglie di slow log degli shard” Se sono stati abilitati i log di verifica, consultare the section called “Passaggio 2: attiva i registri di controllo nelle dashboard OpenSearch”. Se sono stati abilitati solo i log di errore, non è necessario eseguire operazioni di configurazione aggiuntive.

Abilitazione della pubblicazione di log (AWS SDKs)

Prima di poter abilitare la pubblicazione dei log, è necessario creare un gruppo di CloudWatch log, ottenerne l'ARN e concedere al OpenSearch Service le autorizzazioni di scrittura su di esso. Le operazioni pertinenti sono documentate nell'Amazon CloudWatch Logs API Reference:

- `CreateLogGroup`
- `DescribeLogGroup`
- `PutResourcePolicy`

Puoi accedere a queste operazioni utilizzando. AWS SDKs

AWS SDKs (eccetto Android e iOS SDKs) supportano tutte le operazioni definite nell'Amazon OpenSearch Service API Reference, inclusa l'`--log-publishing-options`opzione per `CreateDomain` e`UpdateDomainConfig`.

Se hai abilitato uno degli slow log dello shard, vedi. the section called “Impostazione delle soglie di slow log degli shard” Se sono stati abilitati solo i log di errore, non è necessario eseguire operazioni di configurazione aggiuntive.

Abilitazione della pubblicazione di log (CloudFormation)

In questo esempio, creiamo CloudFormation un gruppo di log chiamato`opensearch-logs`, assegniamo le autorizzazioni appropriate e quindi creiamo un dominio con la pubblicazione dei log abilitata per i log delle applicazioni, gli slow log degli shard di ricerca e gli slow log di indicizzazione.

Prima di poter abilitare la pubblicazione dei log, è necessario creare un gruppo di log: CloudWatch

```
Resources:
  OpenSearchLogGroup:
    Type: AWS::Logs::LogGroup
    Properties:
      LogGroupName: opensearch-logs
Outputs:
  Arn:
    Value:
      'Fn::GetAtt':
        - OpenSearchLogGroup
        - Arn
```

Il modello emette l'ARN del gruppo di log. In questo caso, l'ARN è `arn:aws:logs:us-east-1:123456789012:log-group:opensearch-logs`.

Utilizzando l'ARN, create una politica delle risorse che dia al OpenSearch servizio le autorizzazioni di scrittura nel gruppo di log:

```
Resources:
 OpenSearchLogPolicy:
   Type: AWS::Logs::ResourcePolicy
   Properties:
     PolicyName: my-policy
     PolicyDocument: "{ \"Version\": \"2012-10-17\", \"Statement\": [{ \"Sid\": \"\",
 \"Effect\": \"Allow\", \"Principal\": { \"Service\": \"es.amazonaws.com\"}, \"Action
```

```
\":[ \"logs:PutLogEvents\",\"logs:CreateLogStream\"],\"Resource\": \"arn:aws:logs:us-
east-1:123456789012:log-group:opensearch-logs:*\"}]}"
```

Infine, crea il seguente CloudFormation stack, che genera un dominio di OpenSearch servizio con pubblicazione dei log. La politica di accesso consente all'utente di Account AWS effettuare tutte le richieste HTTP al dominio.

```
Resources:
  OpenSearchServiceDomain:
    Type: "AWS::OpenSearchService::Domain"
    Properties:
      DomainName: my-domain
      EngineVersion: "OpenSearch_1.0"
      ClusterConfig:
        InstanceCount: 2
        InstanceType: "r6g.xlarge.search"
        DedicatedMasterEnabled: true
        DedicatedMasterCount: 3
        DedicatedMasterType: "r6g.xlarge.search"
      EBSOptions:
        EBSEnabled: true
        VolumeSize: 10
        VolumeType: "gp2"
      AccessPolicies:
        Version: "2012-10-17"
        Statement:
            Effect: "Allow"
            Principal:
                AWS: "arn:aws:iam::123456789012:user/es-user"
            Action: "es:*"
            Resource: "arn:aws:es:us-east-1:123456789012:domain/my-domain/*"
      LogPublishingOptions:
        ES_APPLICATION_LOGS:
          CloudWatchLogsLogGroupArn: "arn:aws:logs:us-east-1:123456789012:log-
group:opensearch-logs"
          Enabled: true
        SEARCH_SLOW_LOGS:
          CloudWatchLogsLogGroupArn: "arn:aws:logs:us-east-1:123456789012:log-
group:opensearch-logs"
          Enabled: true
        INDEX_SLOW_LOGS:
          CloudWatchLogsLogGroupArn: "arn:aws:logs:us-east-1:123456789012:log-
group:opensearch-logs"
```

```
          Enabled: true
```

Per informazioni dettagliate sulla sintassi, consultare Opzioni di pubblicazione dei log nella Guida per l'utente di CloudFormation

Impostazione delle soglie di slow log delle richieste di ricerca

Gli slow log delle richieste di ricerca sono disponibili per la ricerca nei domini di OpenSearch servizio in esecuzione nella versione 2.13 e successive. Le soglie di slow log delle richieste di ricerca sono configurate per il tempo totale impiegato dalla richiesta. Ciò è diverso dagli slow log delle richieste di shard, che sono configurati in base al tempo impiegato dai singoli shard.

È possibile specificare gli slow log delle richieste di ricerca con le impostazioni del cluster. Ciò è diverso dagli shard slow log, che puoi abilitare con le impostazioni dell'indice. Ad esempio, puoi specificare le seguenti impostazioni tramite l' OpenSearch API REST:

```
PUT domain-endpoint/_cluster/settings
{
  "transient": {
    "cluster.search.request.slowlog.threshold.warn": "5s",
    "cluster.search.request.slowlog.threshold.info": "2s"
  }
}
```

> **Note**
>
> L'PUT `/_cluster/settings`endpoint è limitato in Amazon OpenSearch Service. L'invio di questa richiesta restituirà l'errore: `Your request: '/_cluster/settings' payload is not allowed.` per configurare le soglie di slow log degli shard, utilizza invece le impostazioni a livello di indice. Esempio:
>
> ```
> PUT domain-endpoint/index/_settings
> {
>   "index.search.slowlog.threshold.query.warn": "5s",
>   "index.search.slowlog.threshold.query.info": "2s"
> }
> ```

Impostazione delle soglie di slow log degli shard

OpenSearch disabilita gli slow log degli shard per impostazione predefinita. Dopo aver abilitato la pubblicazione degli shard slow log su CloudWatch, è comunque necessario specificare le soglie di registrazione per ogni indice. OpenSearch Queste soglie definiscono esattamente cosa deve essere registrato e a quale livello di log.

Ad esempio, puoi specificare queste impostazioni tramite l'API REST: OpenSearch

```
PUT domain-endpoint/index/_settings
{
  "index.search.slowlog.threshold.query.warn": "5s",
  "index.search.slowlog.threshold.query.info": "2s"
}
```

Test degli slow log

Per verificare che sia la richiesta di ricerca che gli shard slow log vengano pubblicati correttamente, è consigliabile iniziare con valori molto bassi per verificare che i log vengano visualizzati CloudWatch, e quindi aumentare le soglie portandole a livelli più utili.

Se i log non vengono visualizzati, verificare quanto segue:

- Il gruppo di log esiste? CloudWatch Controlla la CloudWatch console.
- Il OpenSearch servizio dispone delle autorizzazioni per scrivere nel gruppo di log? Controlla la console OpenSearch di servizio.
- Il dominio del OpenSearch servizio è configurato per la pubblicazione nel gruppo di log? Controlla la console di OpenSearch servizio, utilizza l' AWS CLI `describe-domain-config`opzione o chiama `DescribeDomainConfig` utilizzando una delle SDKs.
- Le soglie OpenSearch di registrazione sono sufficientemente basse da consentire alle tue richieste di superarle?

  Per esaminare le soglie di slow log della richiesta di ricerca per un dominio, usa il seguente comando:

  ```
  GET domain-endpoint/_cluster/settings?flat_settings
  ```

  Per esaminare le soglie di slow log dello shard per un indice, utilizzate il seguente comando:

```
GET domain-endpoint/index/_settings?pretty
```

Se desideri disattivare i log di query lente per un indice, ripristinare le soglie modificate sul valore predefinito di `-1`.

La disattivazione della pubblicazione per l' CloudWatch utilizzo della console di OpenSearch servizio o AWS CLI non OpenSearch impedisce la generazione dei log; interrompe solo la pubblicazione di tali registri. Assicurati di controllare le impostazioni dell'indice se non hai più bisogno degli shard slow log e le impostazioni del dominio se non hai più bisogno degli slow log della richiesta di ricerca.

Visualizzazione dei log di

Visualizzare l'applicazione e gli slow log in CloudWatch è come visualizzare qualsiasi altro registro. CloudWatch Per ulteriori informazioni, consulta View Log Data nella Amazon CloudWatch Logs User Guide.

Di seguito sono elencate alcune considerazioni per la visualizzazione dei log:

- OpenSearch Il servizio pubblica solo i primi 255.000 caratteri di ogni riga su. CloudWatch Il contenuto rimanente viene troncato. Per i log di verifica, sono possibili 10.000 caratteri per messaggio.
- In CloudWatch, i nomi dei flussi di registro hanno il suffisso`-index-slow-logs`, `-search-slow-logs-application-logs`, e `-audit-logs` per facilitare l'identificazione del contenuto.

Monitoraggio dei log di controllo in Amazon Service OpenSearch

Se il tuo dominio Amazon OpenSearch Service utilizza un controllo granulare degli accessi, puoi abilitare i log di controllo per i tuoi dati. I log di controllo sono altamente personalizzabili e ti consentono di tenere traccia delle attività degli utenti sui tuoi OpenSearch cluster, compresi i successi e gli errori di autenticazione, le richieste, le modifiche all'indice e le query di ricerca in OpenSearch arrivo. La configurazione predefinita tiene traccia di una serie comune di azioni utente, ma si consiglia di personalizzare le impostazioni in base alle proprie esigenze.

Proprio come i log OpenSearch delle applicazioni e gli slow log, Service pubblica i log di controllo su Logs. OpenSearch CloudWatch Se abilitato, si applica il prezzo standard. CloudWatch

> **Note**
>
> Per abilitare i log di controllo, il ruolo utente deve essere mappato al `security_manager` ruolo, che consente di accedere all'API OpenSearch `plugins/_security` REST. Per ulteriori informazioni, consulta the section called “Modifica dell'utente principale”.

Argomenti

- Limitazioni
- Abilitazione dei log di verifica
- Abilita la registrazione di controllo utilizzando il AWS CLI
- Abilitare la registrazione di controllo tramite l'API di configurazione
- Livelli e categorie dei log di verifica
- Impostazioni dei log di verifica
- Esempi di log di verifica
- Configurazione dei log di verifica tramite la REST API

Limitazioni

I log di verifica hanno le seguenti limitazioni:

- I log di verifica non includono le richieste della ricerca tra cluster rifiutate dalla policy di accesso al dominio della destinazione.
- La dimensione massima di ogni messaggio del log di verifica è 10.000 caratteri. Il messaggio del log di verifica viene troncato se supera questo limite.

Abilitazione dei log di verifica

La procedura per abilitare i log di verifica prevede due fasi: Innanzitutto, configuri il tuo dominio per pubblicare i log di controllo in Logs. CloudWatch Quindi, abiliti i registri di controllo nelle OpenSearch dashboard e li configuri in base alle tue esigenze.

> **Important**
>
> Se si verifica un errore durante la procedura descritta di seguito, consultare the section called "Impossibile abilitare i log di verifica" per informazioni sulla risoluzione dei problemi.

Passaggio 1: abilitare i log di verifica e configurare una policy di accesso

In questi passaggi viene descritto come abilitare i log di verifica tramite la console. Puoi anche abilitarli utilizzando o AWS CLI l'API di OpenSearch servizio.

Per abilitare i registri di controllo per un dominio OpenSearch di servizio (console)

1. Scegliere il dominio per aprire la sua configurazione, quindi andare alla scheda Log.
2. Selezionare Audit logs (Log di verifica) e poi Enable (Abilita).
3. Crea un gruppo di CloudWatch log o scegline uno esistente.
4. Scegliere una policy d'accesso che contiene le autorizzazioni appropriate oppure creare una policy utilizzando il formato JSON che la console offre:

   JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "es.amazonaws.com"
      },
      "Action": [
        "logs:PutLogEvents",
        "logs:CreateLogStream"
      ],
      "Resource": "arn:aws:logs:us-east-1:111122223333:log-group:/aws/
opensearch/domains/domain-name/*"
    }
  ]
}
```

Si consiglia di aggiungere le chiavi di condizione `aws:SourceAccount` e `aws:SourceArn` alla policy per proteggersi dal problema del "confused deputy". L'account fonte è il proprietario del flusso di log e l'ARN fonte è l'ARN del dominio. Per aggiungere queste chiavi di condizione, il dominio deve trovarsi sul software di servizio R20211203 o versioni successive.

Ad esempio, è possibile aggiungere il seguente blocco di condizione alla policy:

```
"Condition": {
    "StringEquals": {
        "aws:SourceAccount": "account-id"
    },
    "ArnLike": {
        "aws:SourceArn": "arn:aws:es:region:account-id:domain/domain-name"
    }
}
```

5. Scegli Abilita .

Passaggio 2: attiva i registri di controllo nelle dashboard OpenSearch

Dopo aver abilitato i registri di controllo nella console di OpenSearch servizio, devi abilitarli anche nelle OpenSearch dashboard e configurarli in base alle tue esigenze.

1. Apri OpenSearch Dashboard e scegli Sicurezza dal menu a sinistra.
2. Scegliere Log di verifica.
3. Scegliere Abilita registrazione di verifica.

L'interfaccia utente di Dashboards offre il controllo completo delle impostazioni del log di verifica in Impostazioni generali e Impostazioni di conformità. Per una descrizione di tutte le opzioni di configurazione, consultare Impostazioni dei log di verifica.

Abilita la registrazione di controllo utilizzando il AWS CLI

Il AWS CLI comando seguente abilita i registri di controllo su un dominio esistente:

```
aws opensearch update-domain-config --domain-name my-domain --log-publishing-options
 "AUDIT_LOGS={CloudWatchLogsLogGroupArn=arn:aws:logs:us-east-1:123456789012:log-
group:my-log-group,Enabled=true}"
```

È inoltre possibile abilitare i log di verifica quando si crea un dominio. Per ulteriori informazioni, consultare Riferimento ai comandi AWS CLI.

Abilitare la registrazione di controllo tramite l'API di configurazione

La seguente richiesta all'API di configurazione abilita i log di verifica su un dominio esistente:

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/config
{
  "LogPublishingOptions": {
    "AUDIT_LOGS": {
      "CloudWatchLogsLogGroupArn":"arn:aws:logs:us-east-1:123456789012:log-
group1:sample-domain",
      "Enabled":true
    }
  }
}
```

Per ulteriori informazioni, consulta il riferimento all'API OpenSearch di Amazon Service.

Livelli e categorie dei log di verifica

La comunicazione del cluster avviene su due livelli separati: il livello REST e il livello di trasporto.

- Il livello REST copre la comunicazione con i client HTTP come curl, Logstash, OpenSearch Dashboards, il client REST di alto livello Java, la libreria Python Requests, tutte le richieste HTTP che arrivano al cluster.
- Il livello di trasporto copre la comunicazione tra i nodi. Ad esempio, dopo che una richiesta di ricerca arriva al cluster (sul livello REST), il nodo di coordinamento che serve la richiesta invia la query ad altri nodi, riceve le risposte, raccoglie i documenti necessari e li unisce nella risposta finale. Operazioni come l'allocazione delle partizioni e il ribilanciamento si verificano anche sul livello di trasporto.

È possibile abilitare o disabilitare i log di verifica per interi livelli, nonché per singole categorie di controllo per un livello. La tabella seguente contiene un riepilogo delle categorie di verifica e dei livelli per i quali sono disponibili.

Categoria	Description	Disponibile per REST	Disponibile per il trasporto
FALLITO_LOGIN	Una richiesta conteneva credenziali non valide e l'autenticazione non è riuscita.	Sì	Sì
MISSING_PRIVILEGES	Un utente non disponeva dei privilegi per effettuare la richiesta.	Sì	Sì
GRANTED_PRIVILEGES	Un utente disponeva dei privilegi per effettuare la richiesta.	Sì	Sì
OPENSEARCH_SECURITY_INDEX_ATTEMPT	Una richiesta ha provato a modificare l'indice `.opendistro_security` .	No	Sì
AUTHENTICATED	Una richiesta conteneva credenziali valide e l'autenticazione è riuscita.	Sì	Sì
INDEX_EVENT	Una richiesta ha eseguito un'operazione amministrativa su un indice, ad esempio la creazione , l'impostazione di un alias o l'esecuzione di un'unione forzata.	No	Sì

Categoria	Description	Disponibile per REST	Disponibile per il trasporto
	L'elenco completo delle `indices:admin/` azioni incluse in questa categoria è disponibile nella documentazione. OpenSearch		

Oltre a queste categorie standard, il controllo granulare degli accessi offre diverse categorie aggiuntive progettate per soddisfare i requisiti di conformità dei dati.

Categoria	Description
COMPLIANCE_DOC_READ	Una richiesta ha eseguito un evento di lettura su un documento in un indice.
COMPLIANCE_DOC_WRITE	Una richiesta ha eseguito un evento di scrittura su un documento in un indice.
COMPLIANCE_INTERNAL_CONFIG_READ	Una richiesta ha eseguito un evento di lettura sull'indice `.opendistro_security` .
COMPLIANCE_INTERNAL_CONFIG_WRITE	Una richiesta ha eseguito un evento di scrittura sull'indice `.opendistro_security` .

È possibile disporre di una combinazione di categorie e attributi di messaggio. Ad esempio, se si invia una richiesta REST per indicizzare un documento, è possibile che vengano visualizzate le seguenti righe nei log di verifica:

- AUTHENTICATED sul livello REST (autenticazione)
- GRANTED_PRIVILEGE a livello di trasporto (autorizzazione)

- COMPLIANCE_DOC_WRITE (documento scritto su un indice)

Impostazioni dei log di verifica

I log di verifica dispongono di numerose opzioni di configurazione.

Impostazioni generali

Le impostazioni generali consentono di abilitare o disabilitare singole categorie o interi livelli. Si consiglia vivamente di lasciare GRANTED_PRIVILES e AUTHENTICATED come categorie escluse. In caso contrario, queste categorie vengono registrate per ogni richiesta valida al cluster.

Nome	Impostazione del back-end	Description
Livello REST	enable_rest	Abilitare o disabilitare gli eventi che si verificano sul livello REST.
Categorie REST disabilitate	disabled_rest_categories	Specificare le categorie di verifica da ignorare sul livello REST. La modifica di queste categorie può aumentare notevolmente le dimensioni dei log di verifica.
Livello di trasporto	enable_transport	Abilitare o disabilitare gli eventi che si verificano sul livello di trasporto.
Categorie di trasporto disabilitate	disabled_transport_categories	Specificare le categorie di verifica che devono essere ignorate sul livello di trasporto. La modifica di queste categorie può aumentare notevolmente le dimensioni dei log di verifica.

Le impostazioni degli attributi consentono di personalizzare la quantità di dettagli in ogni riga di log.

Nome	Impostazione del back-end	Description
Richieste in blocco	resolve_bulk_requests	L'abilitazione di questa impostazione genera un log per ogni documento in una richiesta in blocco, che può aumentare notevolmente le dimensioni dei log di verifica.
Corpo della richiesta	log_request_body	Includere il corpo della richiesta delle richieste.
Risoluzione di indici	resolve_indices	Risolvere gli alias sugli indici.

Utilizzare Ignora impostazioni per escludere un set di utenti o percorsi API:

Nome	Impostazione del back-end	Description
Utenti ignorati	ignore_users	Specificare gli utenti che si desidera escludere.
Richieste ignorate	ignore_requests	Specificare i modelli di richiesta che si desidera escludere.

Impostazioni di conformità

Le impostazioni di conformità consentono di regolare l'accesso a livello di indice, documento o campo.

Nome	Impostazione del back-end	Description
Registrazione della conformità	enable_compliance	Abilitare o disabilitare la registrazione della conformità.

È possibile specificare le seguenti impostazioni per la registrazione degli eventi di lettura e scrittura.

Nome	Impostazione del back-end	Description
Registrazione della configurazione interna	internal_config	Abilitare o disabilitare la registrazione degli eventi sull'indice `.opendistro_security` .

È possibile specificare le seguenti impostazioni per gli eventi di lettura.

Nome	Impostazione del back-end	Description
Lettura di metadati	read_metadata_only	Includere solo i metadati per gli eventi di lettura. Non includere campi di documento.
Utenti ignorati	read_ignore_users	Non includere determinati utenti per gli eventi di lettura.
Campi osservati	read_watched_fields	Specificare gli indici e i campi da controllare per gli eventi di lettura. L'aggiunta di campi controllati genera un log per accesso ai documenti, che può aumentare notevolmente le dimensioni dei log di verifica. I campi osservati supportano i modelli di indice e i modelli di campo: `{ "index-name-pattern": [ "field-name-pattern" ], "logs*": [ "message" ], "twitter": [ "id", "user*" ] }`

È possibile specificare le seguenti impostazioni per gli eventi di scrittura.

Nome	Impostazione del back-end	Description
Scrittura di metadati	write_metadata_only	Includere solo i metadati per gli eventi di scrittura. Non includere campi di documento.
Differenze dei log	write_log_diffs	Se write_metadata_only è false, includere solo le differenze tra gli eventi di scrittura.
Utenti ignorati	write_ignore_users	Non includere determinati utenti per gli eventi di scrittura.
Osservazione di indici	write_watched_indices	Specificare gli indici o i modelli di indice da controllare per gli eventi di scrittura. L'aggiunta di campi controllati genera un log per accesso ai documenti, che può aumentare notevolmente le dimensioni dei log di verifica.

Esempi di log di verifica

Questa sezione include una configurazione di esempio, una richiesta di ricerca e il log di verifica risultante per tutti gli eventi di lettura e scrittura di un indice.

Fase 1: Configurazione dei log di verifica

Dopo aver abilitato la pubblicazione dei log di controllo in un gruppo CloudWatch Logs, accedi alla pagina di registrazione di audit di OpenSearch Dashboards e scegli Abilita registrazione di controllo.

1. In Impostazioni generali, scegliere Configura e assicurarsi che il livello REST sia abilitato.
2. In Impostazioni di conformità, scegliere Configura.
3. In Scrittura, in Campi osservati, aggiungere `accounts` per tutti gli eventi di scrittura su questo indice.
4. In Lettura, in Campi osservati, aggiungere i campi `ssn` e `id-` dell'indice `accounts`:

```
{
  "accounts-": [
    "ssn",
```

```
    "id-"
  ]
}
```

Fase 2: Esecuzione di eventi di lettura e scrittura

1. Passa a OpenSearch Dashboards, scegli Dev Tools e indicizza un documento di esempio:

```
PUT accounts/_doc/0
{
  "ssn": "123",
  "id-": "456"
}
```

2. Per verificare un evento di lettura, inviare la seguente richiesta:

```
GET accounts/_search
{
  "query": {
    "match_all": {}
  }
}
```

Fase 3: Osservazione dei log

1. Apri la CloudWatch console all'indirizzo https://console.aws.amazon.com/cloudwatch/.
2. Nel pannello di navigazione, selezionare Log groups (Gruppi di log).
3. Scegliere il gruppo di log specificato durante l'abilitazione dei log di verifica. All'interno del gruppo di log, OpenSearch Service crea un flusso di log per ogni nodo del dominio.
4. In Flussi di log, scegliere Cerca tutto.
5. Per gli eventi di lettura e scrittura, consultare i log corrispondenti. È possibile prevedere un ritardo di 5 secondi prima che venga visualizzato il log.

   Log di verifica della scrittura di esempio

```
{
  "audit_compliance_operation": "CREATE",
  "audit_cluster_name": "824471164578:audit-test",
```

```
  "audit_node_name": "be217225a0b77c2bd76147d3ed3ff83c",
  "audit_category": "COMPLIANCE_DOC_WRITE",
  "audit_request_origin": "REST",
  "audit_compliance_doc_version": 1,
  "audit_node_id": "3xNJhm4XS_yTzEgDWcGRjA",
  "@timestamp": "2020-08-23T05:28:02.285+00:00",
  "audit_format_version": 4,
  "audit_request_remote_address": "3.236.145.227",
  "audit_trace_doc_id": "lxnJGXQBqZSlDB91r_uZ",
  "audit_request_effective_user": "admin",
  "audit_trace_shard_id": 8,
  "audit_trace_indices": [
    "accounts"
  ],
  "audit_trace_resolved_indices": [
    "accounts"
  ]
}
```


Log di verifica della lettura di esempio


```
{
  "audit_cluster_name": "824471164578:audit-docs",
  "audit_node_name": "806f6050cb45437e2401b07534a1452f",
  "audit_category": "COMPLIANCE_DOC_READ",
  "audit_request_origin": "REST",
  "audit_node_id": "saSevm9ASte0-pjAtYi2UA",
  "@timestamp": "2020-08-31T17:57:05.015+00:00",
  "audit_format_version": 4,
  "audit_request_remote_address": "54.240.197.228",
  "audit_trace_doc_id": "config:7.7.0",
  "audit_request_effective_user": "admin",
  "audit_trace_shard_id": 0,
  "audit_trace_indices": [
    "accounts"
  ],
  "audit_trace_resolved_indices": [
    "accounts"
  ]
}
```

Per includere il corpo della richiesta, torna alle impostazioni di conformità nei OpenSearch dashboard e disabilita Write metadata. Per escludere eventi da un utente specifico, aggiungere l'utente a Utenti ignorati.

Per una descrizione di ogni campo del log di verifica, consultare Riferimento ai campi dei log di verifica. Per informazioni sulla ricerca e l'analisi dei dati dei log di controllo, consulta Analyzing Log Data with CloudWatch Logs Insights nella Amazon CloudWatch Logs User Guide.

Configurazione dei log di verifica tramite la REST API

Ti consigliamo di utilizzare OpenSearch le dashboard per configurare i log di controllo, ma puoi anche utilizzare l'API REST per il controllo degli accessi a grana fine. In questa sezione è riportata una richiesta di esempio. La documentazione completa sull'API REST è disponibile nella documentazione. OpenSearch

```
PUT _opendistro/_security/api/audit/config
{
  "enabled": true,
  "audit": {
    "enable_rest": true,
    "disabled_rest_categories": [
      "GRANTED_PRIVILEGES",
      "AUTHENTICATED"
    ],
    "enable_transport": true,
    "disabled_transport_categories": [
      "GRANTED_PRIVILEGES",
      "AUTHENTICATED"
    ],
    "resolve_bulk_requests": true,
    "log_request_body": true,
    "resolve_indices": true,
    "exclude_sensitive_headers": true,
    "ignore_users": [
      "kibanaserver"
    ],
    "ignore_requests": [
      "SearchRequest",
      "indices:data/read/*",
      "/_cluster/health"
    ]
  },
```

```
  "compliance": {
    "enabled": true,
    "internal_config": true,
    "external_config": false,
    "read_metadata_only": true,
    "read_watched_fields": {
      "read-index-1": [
        "field-1",
        "field-2"
      ],
      "read-index-2": [
        "field-3"
      ]
    },
    "read_ignore_users": [
      "read-ignore-1"
    ],
    "write_metadata_only": true,
    "write_log_diffs": false,
    "write_watched_indices": [
      "write-index-1",
      "write-index-2",
      "log-*",
      "*"
    ],
    "write_ignore_users": [
      "write-ignore-1"
    ]
  }
}
```

Monitoraggio degli eventi del OpenSearch servizio con Amazon EventBridge

Amazon OpenSearch Service si integra con Amazon EventBridge per notificarti determinati eventi che influiscono sui tuoi domini. Gli eventi AWS relativi ai servizi vengono forniti quasi EventBridge in tempo reale. Gli stessi eventi vengono inviati anche ad Amazon CloudWatch Events, il predecessore di Amazon EventBridge. Puoi compilare regole semplici che indichino quali eventi sono considerati di interesse per te e quali azioni automatizzate intraprendere quando un evento corrisponde a una regola. Le azioni che possono essere attivate automaticamente includono le seguenti:

- Invocare una funzione AWS Lambda
- Richiamo di un Run Command di Amazon EC2
- Inoltro dell'evento a Amazon Kinesis Data Streams
- Attivazione di una macchina a stati AWS Step Functions
- Notifica di un argomento Amazon SNS o di una coda Amazon SQS

Per ulteriori informazioni, consulta la sezione Guida introduttiva ad Amazon EventBridge nella Amazon EventBridge User Guide.

Argomenti

- Eventi di aggiornamento del software di servizio
- Eventi di regolazione automatica
- Eventi sull'integrità del cluster
- Eventi di endpoint VPC
- Eventi di ritiro dei nodi
- Eventi di ritiro dei nodi degradati
- Eventi di errore di dominio
- Tutorial: ascolto degli EventBridge eventi OpenSearch di Amazon Service
- Esercitazione: Invio di avvisi Amazon SNS per gli aggiornamenti software disponibili

Eventi di aggiornamento del software di servizio

OpenSearch Il servizio invia eventi EventBridge quando si verifica uno dei seguenti eventi di aggiornamento del software di servizio.

Aggiornamento del software di servizio disponibile

OpenSearch Il servizio invia questo evento quando è disponibile un aggiornamento del software di servizio.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
```

```
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Available",
    "severity": "Informational",
    "description": "Service software update R20220928 available. Service Software
 Deployment Mechanism:
                   Blue/Green. For more information on deployment configuration,
 please
                   see: https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/managedomains-configuration-changes.html"
  }
}
```

Aggiornamento del software di servizio pianificato

OpenSearch Il servizio invia questo evento quando è stato pianificato un aggiornamento del software di servizio. Per gli aggiornamenti opzionali, ricevi la notifica nella data pianificata e hai la possibilità di riprogrammarli in qualsiasi momento. Per gli aggiornamenti richiesti, ricevi la notifica tre giorni prima della data pianificata e hai la possibilità di riprogrammarla entro la finestra obbligatoria.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Scheduled",
```

```
    "severity": "High",
    "description": "A new service software update [R20200330-p1] has been scheduled at
 [21st May 2023 12:40 GMT].
                    Please see documentation for more information on scheduling
 software updates:
                    https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/service-software.html."
  }
}
```

Aggiornamento del software di servizio riprogrammato

OpenSearch Il servizio invia questo evento quando un aggiornamento opzionale del software di servizio viene riprogrammato. Per ulteriori informazioni, consulta the section called “Aggiornamenti facoltativi e obbligatori”.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Rescheduled",
    "severity": "High",
    "description": "The service software update [R20200330-p1], which was originally
 scheduled for
                    [21st May 2023 12:40 GMT], has been rescheduled to [23rd May 2023
 12:40 GMT].
                    Please see documentation for more information on scheduling
 software updates:
                    https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/service-software.html."
  }
}
```

Aggiornamento del software di servizio avviato

OpenSearch Il servizio invia questo evento quando viene avviato un aggiornamento del software di servizio.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Started",
    "severity": "Informational",
    "description": "Service software update [R20200330-p1] started.
  }
}
```

Aggiornamento del software di servizio completato

OpenSearch Il servizio invia questo evento quando viene completato un aggiornamento del software di servizio.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
```

```
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Completed",
    "severity": "Informational",
    "description": "Service software update [R20200330-p1] completed."
  }
}
```

Aggiornamento del software di servizio annullato

OpenSearch Il servizio invia questo evento quando un aggiornamento del software di servizio viene annullato.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Cancelled",
    "severity": "Informational",
    "description": "The scheduled service software update [R20200330-p1] has been
 cancelled as a
                    newer update is available. Please schedule the latest update."
  }
}
```

Aggiornamento programmato del software di servizio annullato

OpenSearch Il servizio invia questo evento quando un aggiornamento del software di servizio precedentemente pianificato per il dominio è stato annullato.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Cancelled",
    "severity": "Informational",
    "description": "The scheduled service software update [R20200330-p1] has been
 cancelled."
  }
}
```

Aggiornamento del software di servizio non eseguito

OpenSearch Il servizio invia questo evento quando non è in grado di avviare un aggiornamento del software di servizio.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Unexecuted",
    "severity": "Informational",
```

```
    "description": "The scheduled service software update [R20200330-p1] cannot be
 started. Reason: [reason]"
  }
}
```

Aggiornamento del software di servizio non riuscito

OpenSearch Il servizio invia questo evento quando un aggiornamento del software di servizio fallisce.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Failed",
    "severity": "High",
    "description": "Installation of service software update [R20200330-p1] failed.
 [reason].
  }
}
```

Aggiornamento del software di servizio richiesto

OpenSearch Il servizio invia questo evento quando è richiesto un aggiornamento del software di servizio. Per ulteriori informazioni, consulta the section called “Aggiornamenti facoltativi e obbligatori”.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
```

```
  "detail-type": "Amazon OpenSearch Service Software Update Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2016-11-01T13:12:22Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Service Software Update",
    "status": "Required",
    "severity": "High",
    "description": "Service software update [R20200330-p1] available. Update
                    will be automatically installed after [21st May 2023] if no
                    action is taken. Service Software Deployment Mechanism: Blue/Green.
                    For more information on deployment configuration, please see:
                    https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/managedomains-configuration-changes.html"
  }
}
```

Eventi di regolazione automatica

OpenSearch Il servizio invia eventi EventBridge quando si verifica uno dei seguenti eventi Auto-Tune.

Regolazione automatica in sospeso

OpenSearch Il servizio invia questo evento quando Auto-Tune ha identificato i consigli di ottimizzazione per migliorare le prestazioni e la disponibilità del cluster. Questo evento verrà visualizzato solo per i domini con la regolazione automatica disabilitata.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "3acb26c8-397c-4c89-a80a-ce672a864c55",
  "detail-type": "Amazon OpenSearch Service Auto-Tune Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2020-10-30T22:06:31Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
```

```
    "event": "Auto-Tune Event",
    "severity": "Informational",
    "status": "Pending",
    "description": "Auto-Tune recommends the following new settings for your
 domain: { JVM Heap size : 60%}. Enable Auto-Tune to improve cluster stability and
 performance.",
    "scheduleTime": "{iso8601-timestamp}"
  }
}
```

Regolazione automatica avviata

OpenSearch Il servizio invia questo evento quando Auto-Tune inizia ad applicare nuove impostazioni al dominio.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "3acb26c8-397c-4c89-a80a-ce672a864c55",
  "detail-type": "Amazon OpenSearch Service Auto-Tune Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2020-10-30T22:06:31Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Auto-Tune Event",
    "severity": "Informational",
    "status": "Started",
    "scheduleTime": "{iso8601-timestamp}",
    "startTime": "{iso8601-timestamp}",
    "description" : "Auto-Tune is applying the following settings to your domain: { JVM
 Heap size : 60%}."
  }
}
```

Auto-Tune richiede una distribuzione pianificata blue/green

OpenSearch Il servizio invia questo evento quando Auto-Tune ha identificato raccomandazioni di ottimizzazione che richiedono un'implementazione pianificata. blue/green

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "3acb26c8-397c-4c89-a80a-ce672a864c55",
  "detail-type": "Amazon OpenSearch Service Auto-Tune Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2020-10-30T22:06:31Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Auto-Tune Event",
    "severity": "Low",
    "status": "Pending",
    "startTime": "{iso8601-timestamp}",
    "description": "Auto-Tune has identified the following settings for your domain
 that require a blue/green deployment: { JVM Heap size : 60%}.
                   You can schedule the deployment for your preferred time."
  }
}
```

Regolazione automatica annullata

OpenSearch Il servizio invia questo evento quando la pianificazione di Auto-Tune è stata annullata perché non ci sono consigli di ottimizzazione in sospeso.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "3acb26c8-397c-4c89-a80a-ce672a864c55",
  "detail-type": "Amazon OpenSearch Service Auto-Tune Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2020-10-30T22:06:31Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
```

```
    "event": "Auto-Tune Event",
    "severity": "Low",
    "status": "Cancelled",
    "scheduleTime": "{iso8601-timestamp}",
    "description": "Auto-Tune has cancelled the upcoming blue/green deployment."
  }
}
```

Regolazione automatica completata

OpenSearch Il servizio invia questo evento quando Auto-Tune ha completato la blue/green distribuzione e il cluster è operativo con le nuove impostazioni JVM.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "3acb26c8-397c-4c89-a80a-ce672a864c55",
  "detail-type": "Amazon OpenSearch Service Auto-Tune Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2020-10-30T22:06:31Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Auto-Tune Event",
    "severity": "Informational",
    "status": "Completed",
    "completionTime": "{iso8601-timestamp}",
    "description": "Auto-Tune has completed the blue/green deployment and successfully
 applied the following settings: { JVM Heap size : 60%}."
  }
}
```

Regolazione automatica disabilitata e modifiche ripristinate

OpenSearch Il servizio invia questo evento quando Auto-Tune è stato disabilitato e le modifiche applicate sono state ripristinate.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "3acb26c8-397c-4c89-a80a-ce672a864c55",
  "detail-type": "Amazon OpenSearch Service Auto-Tune Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2020-10-30T22:06:31Z",
  "region": "us-east-1",
  "resources": [ "arn:aws:es:us-east-1:123456789012:domain/test-domain" ],
  "detail": {
    "event": "Auto-Tune Event",
    "severity": "Informational",
    "status": "Completed",
    "description": "Auto-Tune is now disabled. All settings have been reverted. Auto-
Tune will continue to evaluate
                    cluster performance and provide recommendations.",
    "completionTime": "{iso8601-timestamp}"
  }
}
```

Regolazione automatica disabilitata e modifiche conservate

OpenSearch Il servizio invia questo evento quando Auto-Tune è stato disabilitato e le modifiche applicate sono state mantenute.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "3acb26c8-397c-4c89-a80a-ce672a864c55",
  "detail-type": "Amazon OpenSearch Service Auto-Tune Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2020-10-30T22:06:31Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Auto-Tune Event",
    "severity": "Informational",
```

```
    "status": "Completed",
    "description": "Auto-Tune is now disabled. The most-recent settings by Auto-Tune
 have been retained.
                    Auto-Tune will continue to evaluate cluster performance and provide
 recommendations.",
    "completionTime": "{iso8601-timestamp}"
  }
}
```

Eventi sull'integrità del cluster

OpenSearch Il servizio invia determinati eventi EventBridge quando lo stato del cluster è compromesso.

Avviato il recupero del cluster rosso

OpenSearch Il servizio invia questo evento dopo che lo stato del cluster è rimasto ininterrottamente in rosso per più di un'ora. Tenta di ripristinare automaticamente uno o più indici rossi da uno snapshot per correggere lo stato del cluster.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service Cluster Status Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2016-11-01T13:12:22Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
   ],
   "detail":{
      "event":"Automatic Snapshot Restore for Red Indices",
      "status":"Started",
      "severity":"High",
      "description":"Your cluster status is red. We have started automatic snapshot
 restore for the red indices.
                     No action is needed from your side. Red indices [red-index-0, red-
index-1]"
```

```
   }
}
```

Recupero del cluster rosso parzialmente completato

OpenSearch Il servizio invia questo evento quando è stato in grado di ripristinare solo un sottoinsieme di indici rossi da un'istantanea durante il tentativo di correggere lo stato di un cluster rosso.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service Cluster Status Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2016-11-01T13:12:22Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
   ],
   "detail":{
      "event":"Automatic Snapshot Restore for Red Indices",
      "status":"Partially Restored",
      "severity":"High",
      "description":"Your cluster status is red. We were able to restore the following
 Red indices from
                   snapshot: [red-index-0]. Indices not restored: [red-index-1].
 Please refer https://docs.aws.amazon.com/opensearch-service/latest/developerguide/
handling-errors.html#handling-errors-red-cluster-status for troubleshooting steps."
   }
}
```

Recupero del cluster rosso non riuscito

OpenSearch Il servizio invia questo evento quando non riesce a ripristinare alcun indice durante il tentativo di correggere lo stato di un cluster rosso.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service Cluster Status Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2016-11-01T13:12:22Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
   ],
   "detail":{
      "event":"Automatic Snapshot Restore for Red Indices",
      "status":"Failed",
      "severity":"High",
      "description":"Your cluster status is red. We were unable to restore the Red
 indices automatically.
                   Indices not restored: [red-index-0, red-index-1]. Please refer
 https://docs.aws.amazon.com/opensearch-service/latest/developerguide/handling-
errors.html#handling-errors-red-cluster-status for troubleshooting steps."
   }
}
```

Partizioni da eliminare

OpenSearch Il servizio invia questo evento quando tenta di correggere automaticamente lo stato del cluster rosso dopo che era rimasto in rosso ininterrottamente per 14 giorni, ma uno o più indici rimangono rossi. Dopo altri 7 giorni (21 giorni totali in cui il colore è continuamente rosso), il OpenSearch Servizio procede all'eliminazione degli shard non assegnati su tutti gli indici rossi.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service Cluster Status Notification",
   "source":"aws.es",
   "account":"123456789012",
```

```
   "time":"2022-04-09T10:36:48Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
   ],
   "detail":{
      "severity":"Medium",
      "description":"Your cluster status is red. Please fix the red indices as soon as
 possible.
                    If not fixed by 2022-04-12 01:51:47+00:00, we will delete all
 unassigned shards,
                    the unit of storage and compute, for these red indices to recover
 your domain and make it green.
                    Please refer to https://docs.aws.amazon.com/opensearch-service/
latest/developerguide/handling-errors.html#handling-errors-red-cluster-status for
 troubleshooting steps.
                    test_data, test_data1",
      "event":"Automatic Snapshot Restore for Red Indices",
      "status":"Shard(s) to be deleted"
   }
}
```

Partizioni eliminate

OpenSearch Il servizio invia questo evento dopo che lo stato del cluster è rimasto in rosso ininterrottamente per 21 giorni. Si procede all'eliminazione delle partizioni non assegnate (archiviazione e calcolo) su tutti gli indici rossi. Per informazioni dettagliate, vedi the section called "Correzione automatico di cluster rossi".

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service Cluster Status Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2022-04-09T10:54:48Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
```

```
   ],
   "detail":{
      "severity":"High",
      "description":"We have deleted unassinged shards, the unit of storage and
 compute, in
                    red indices: index-1, index-2 because these indices were red for
 more than
                    21 days and could not be restored with the automated restore
 process.
                    Please refer to https://docs.aws.amazon.com/opensearch-service/
latest/developerguide/handling-errors.html#handling-errors-red-cluster-status for
 troubleshooting steps.",
      "event":"Automatic Snapshot Restore for Red Indices",
      "status":"Shard(s) deleted"
   }
}
```

Avviso di conteggio di partizioni elevate

OpenSearch Il servizio invia questo evento quando il numero medio di shard negli hot data node ha superato il 90% del limite predefinito consigliato di 1.000. Sebbene le versioni successive di Elasticsearch OpenSearch supportino un limite massimo configurabile di shard per nodo, consigliamo di non avere più di 1.000 shard per nodo. Vedi Scelta del numero di partizioni.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2016-11-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"High Shard Count",
     "status":"Warning",
     "severity":"Low",
     "description":"One or more data nodes have close to 1000 shards. To ensure optimum
 performance and stability of your
```

```
                    cluster, please refer to the best practice guidelines - https://
docs.aws.amazon.com/opensearch-service/latest/developerguide/sizing-domains.html#bp-
sharding."
  }
}
```

Superamento del limite del numero di partizioni

OpenSearch Il servizio invia questo evento quando il numero medio di shard nei nodi di hot data ha superato il limite predefinito consigliato di 1.000. Sebbene le versioni successive di Elasticsearch OpenSearch supportino un limite massimo configurabile per il numero massimo di shard per nodo, consigliamo di non avere più di 1.000 shard per nodo. Vedi Scelta del numero di partizioni.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2016-11-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"High Shard Count",
     "status":"Warning",
     "severity":"Medium",
     "description":"One or more data nodes have more than 1000 shards. To ensure
 optimum performance and stability of your
                    cluster, please refer to the best practice guidelines - https://
docs.aws.amazon.com/opensearch-service/latest/developerguide/sizing-domains.html#bp-
sharding."
  }
}
```

Spazio su disco ridotto

OpenSearch Il servizio invia questo evento quando uno o più nodi del cluster hanno meno del 25% dello spazio di archiviazione disponibile o meno di 25 GB.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"Low Disk Space",
     "status":"Warning",
     "severity":"Medium",
     "description":"One or more data nodes in your cluster has less than 25% of storage
 space or less than 25GB.
                  Your cluster will be blocked for writes at 20% or 20GB. Please refer
 to the documentation for more information - https://docs.aws.amazon.com/opensearch-
service/latest/developerguide/handling-errors.html#troubleshooting-cluster-block"
  }
}
```

Violazione del watermark su disco ridotta

OpenSearch Il servizio invia questo evento quando tutti i nodi del cluster hanno meno del 10% dello spazio di archiviazione disponibile o meno di 10 GB. Quando tutti i nodi superano la soglia di esaurimento del disco, ogni nuovo indice genera un cluster giallo e, quando tutti i nodi scendono al di sotto del limite massimo del disco, viene visualizzato un cluster rosso.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
```

```
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"Low Disk Watermark Breach",
     "status":"Warning",
     "severity":"Medium",
     "description":"Low Disk Watermark threshold is about to be breached. Once the
 threshold is breached, new index creation will be blocked on all
                   nodes to prevent the cluster status from turning red. Please
 increase disk size to suit your storage needs. For more information,
                   see https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/handling-errors.html#troubleshooting-cluster-block".
  }
}
```

Saldo di burst EBS inferiore al 70%

OpenSearch Il servizio invia questo evento quando il saldo del burst di EBS su uno o più nodi di dati scende al di sotto del 70%. L'esaurimento del burst balance di EBS può causare un'indisponibilità diffusa del cluster e una limitazione delle I/O richieste, il che può portare a latenze e timeout elevati nelle richieste di indicizzazione e ricerca. Per la procedura per correggere questo problema, consulta la sezione the section called “Saldo di burst EBS basso”.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"EBS Burst Balance",
     "status":"Warning",
     "severity":"Medium",
     "description":"EBS burst balance on one or more data nodes is below 70%.
                   Follow https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/handling-errors.html#handling-errors-low-ebs-burst
```

```
                    to fix this issue."
  }
}
```

Saldo di burst EBS inferiore al 20%

OpenSearch Il servizio invia questo evento quando il saldo del burst di EBS su uno o più nodi di dati scende al di sotto del 20%. L'esaurimento del burst balance di EBS può causare un'indisponibilità diffusa del cluster e una limitazione delle I/O richieste, il che può portare a latenze e timeout elevati nelle richieste di indicizzazione e ricerca. Per la procedura per correggere questo problema, consulta la sezione the section called “Saldo di burst EBS basso”.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"EBS Burst Balance",
     "status":"Warning",
     "severity":"High",
     "description":"EBS burst balance on one or more data nodes is below 20%.
                    Follow https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/handling-errors.html#handling-errors-low-ebs-burst
                    to fix this issue.
  }
}
```

Limitazione della velocità di trasmissione effettiva del disco

OpenSearch Il servizio invia questo evento quando le richieste di lettura e scrittura verso il dominio vengono limitate a causa delle limitazioni di throughput dei volumi EBS o dell'istanza EC2. Se ricevi questa notifica, valuta la possibilità di aumentare i volumi o le istanze seguendo le best practice consigliate. AWS Se il tipo di volume ègp2, aumenta la dimensione del volume. Se il tipo di volume è

lo stessogp3, fornisci una maggiore velocità effettiva. Puoi anche verificare che la base dell'istanza e il throughput EBS massimo siano maggiori o uguali al throughput di volume assegnato e aumentare di conseguenza.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"Disk Throughput Throttle",
     "status":"Warning",
     "severity":"Medium",
     "description":"Your domain is experiencing throttling due to instance or volume
 throughput limitations.
                   Please consider scaling your domain to suit your throughput needs.
 In July 2023, we improved
                   the accuracy of throughput throttle calculation by replacing ‘Max
 volume throughput’ with
                   ‘Provisioned volume throughput’. Please refer to the documentation
 for more information."
   }
}
```

Frammenti di grandi dimensioni

OpenSearch Il servizio invia questo evento quando uno o più shard nel cluster hanno superato i 50 GiB o i 65 GiB. Per garantire prestazioni e stabilità ottimali del cluster, riduci le dimensioni degli shard.

Per ulteriori informazioni, consulta le migliori pratiche di sharding.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"Large Shard Size",
     "status":"Warning",
     "severity":"Medium",
     "description":"One or more shards are larger than 65GiB. To ensure optimum cluster
 performance and stability, reduce shard sizes.
                    For more information, see https://docs.aws.amazon.com/opensearch-
service/latest/developerguide/monitoring-events.html#monitoring-events-large-shard-
size."
  }
}
```

Elevato utilizzo JVM

OpenSearch Il servizio invia questo evento quando la `JVMMemoryPressure` metrica per il tuo dominio ha superato l'80%. Se supera il 92% per 30 minuti, tutte le operazioni di scrittura sul cluster verranno bloccate. Per garantire una stabilità ottimale del cluster, riduci il traffico verso il cluster o ridimensiona il dominio per fornire memoria sufficiente per il carico di lavoro.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
```

```
     "event":"High JVM Usage",
     "status":"Warning",
     "severity":"High",
     "description":"JVM memory pressure has exceeded 80%. If it exceeds 92% for 30
 minutes, all write operations to your cluster
                    will be blocked. To ensure optimum cluster stability, reduce
 traffic to the cluster or use larger instance types.
                    For more information, see https://docs.aws.amazon.com/opensearch-
service/latest/developerguide/monitoring-events.html#monitoring-events-high-jvm."
  }
}
```

GC insufficiente

OpenSearch Il servizio invia questo evento quando la JVM massima è superiore al 70% e la differenza tra il massimo e il minimo è inferiore al 30%. Ciò può indicare che la JVM non è in grado di recuperare memoria sufficiente durante i cicli di raccolta dei rifiuti per il carico di lavoro. Ciò può portare a risposte sempre più lente e latenze più elevate e, in alcuni casi, persino a cadute dei nodi a causa dei controlli di integrità scaduti. Per garantire una stabilità ottimale del cluster, riduci il traffico verso il cluster o ridimensiona il dominio per fornire memoria sufficiente per il carico di lavoro.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"Insufficient GC",
     "status":"Warning",
     "severity":"Medium",
     "description":"Maximum JVM is above 70% and JVM range is less than 30%. This may
 indicate insufficient garbage collection for your workload.
                    For more information, see https://docs.aws.amazon.com/opensearch-
service/latest/developerguide/monitoring-events.html#monitoring-events-insufficient-
gc."
```

```
  }
}
```

Avviso di routing dell'indice personalizzato

OpenSearch Il servizio invia questo evento quando il dominio è in stato di elaborazione e contiene indici con impostazioni index.routing.allocation personalizzate che possono causare il blocco delle distribuzioni blu-verdi. Verifica che le impostazioni siano applicate correttamente.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"Custom Index Routing Warning",
     "status":"Warning",
     "severity":"Medium",
     "description":"Your domain is in processing state and contains indice(s) with
 custom index.routing.allocation
                   settings which can cause blue-green deployments to get stuck.
 Verify settings are applied properly.
                   For more information, see https://docs.aws.amazon.com/opensearch-
service/latest/developerguide/monitoring-events.html#monitoring-events-index-routing."
  }
}
```

Shard lock non riuscito

OpenSearch Il servizio invia questo evento quando il dominio non è integro a causa di shard non assegnati con. `[ShardLockObtainFailedException]` Per ulteriori informazioni, consulta Come posso risolvere l'eccezione dello shard lock in memoria in Amazon OpenSearch Service?

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2017-12-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"Failed Shard Lock",
     "status":"Warning",
     "severity":"Medium",
     "description":"Your domain is unhealthy due to unassigned shards with
 [ShardLockObtainFailedException]. For more information,
                   see https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/monitoring-events.html#monitoring-events-failed-shard-lock."
}
```

Eventi di endpoint VPC

OpenSearch Il servizio invia determinati eventi agli endpoint EventBridge correlati all'AWS PrivateLink interfaccia.

Creazione dell'endpoint VPC non riuscita

OpenSearch Il servizio invia questo evento quando non è in grado di creare un endpoint VPC richiesto. Questo errore potrebbe verificarsi perché hai raggiunto il limite del numero di endoint VPC consentiti all'interno di una regione. Questo errore verrà visualizzato anche se non esiste una sottorete o un gruppo di sicurezza specificato.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service VPC Endpoint Notification",
   "source":"aws.es",
   "account":"123456789012",
```

```
   "time":"2016-11-01T13:12:22Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
   ],
   "detail":{
      "event":"VPC Endpoint Create Validation",
      "status":"Failed",
      "severity":"High",
      "description":"Unable to create VPC endpoint aos-0d4c74c0342343 for domain
                    arn:aws:es:eu-south-1:123456789012:domain/my-domain due to the
 following validation failures: You've reached the limit on the
                    number of VPC endpoints that you can create in the AWS Region."
   }
}
```

Aggiornamento dell'endpoint VPC non riuscito

OpenSearch Il servizio invia questo evento quando non è in grado di eliminare un endpoint VPC richiesto.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service VPC Endpoint Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2016-11-01T13:12:22Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
   ],
   "detail":{
      "event":"VPC Endpoint Update Validation",
      "status":"Failed",
      "severity":"High",
      "description":"Unable to update VPC endpoint aos-0d4c74c0342343 for domain
                    arn:aws:es:eu-south-1:123456789012:domain/my-domain due to the
 following validation failures: <failure message>."
```

```
   }
}
```

Eliminazione dell'endpoint VPC non riuscita

OpenSearch Il servizio invia questo evento quando non è in grado di eliminare un endpoint VPC richiesto.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service VPC Endpoint Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2016-11-01T13:12:22Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
   ],
   "detail":{
      "event":"VPC Endpoint Delete Validation",
      "status":"Failed",
      "severity":"High",
      "description":"Unable to delete VPC endpoint aos-0d4c74c0342343 for domain
                     arn:aws:es:eu-south-1:123456789012:domain/my-domain due to the
 following validation failures: Specified subnet doesn't exist."
   }
}
```

Eventi di ritiro dei nodi

OpenSearch Il servizio invia gli eventi EventBridge quando si verifica uno dei seguenti eventi di ritiro del nodo.

Ritiro del nodo pianificato

OpenSearch Il servizio invia questo evento quando è stato pianificato il ritiro di un nodo.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2023-04-07T10:07:33Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Node Retirement Notification",
    "status": "Scheduled",
    "severity": "Medium",
    "description": "An automated action to retire and replace a node has been scheduled
 on your domain.
                    The node will be replaced in the next off-peak window. For more
 information, see
                    https://docs.aws.amazon.com/opensearch-service/latest/
developerguide/monitoring-events.html."
  }
}
```

Ritiro del nodo completato

OpenSearch Il servizio invia questo evento quando il ritiro di un nodo è completato.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2023-04-07T10:07:33Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Node Retirement Notification",
```

```
    "status": "Completed",
    "severity": "Medium",
    "description": "The node has been retired and replaced with a new node."
  }
}
```

Ritiro del nodo non riuscito

OpenSearch Il servizio invia questo evento quando il ritiro di un nodo fallisce.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "Amazon OpenSearch Service Notification",
  "source": "aws.es",
  "account": "123456789012",
  "time": "2023-04-07T10:07:33Z",
  "region": "us-east-1",
  "resources": ["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail": {
    "event": "Node Retirement Notification",
    "status": "Failed",
    "severity": "Medium",
    "description": "Node retirement failed. No actions are required from your end. We
 will automatically
                   retry replacing the node."
  }
}
```

Eventi di ritiro dei nodi degradati

OpenSearch Il servizio invia questi eventi quando è necessaria la sostituzione di un nodo a causa di un hardware danneggiato su un nodo.

Notifica di ritiro del nodo danneggiato

OpenSearch Il servizio invia questo evento quando l'azione automatica di ritiro e sostituzione di un nodo danneggiato è stata pianificata per il dominio.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"db233454-aad1-7676-3b15-10a84b052baa",
   "detail-type":"Amazon OpenSearch Service Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2024-01-11T08:16:06Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-node-replacement"
   ],
   "detail":{
      "severity":"Medium",
      "description":"An automated action to retire and replace a node has
 been scheduled on your domain. For more information, please see https://
docs.aws.amazon.com/opensearch-service/latest/developerguide/monitoring-events.html.",
      "event":"Degraded Node Retirement Notification",
      "status":"Scheduled"
   }
}
```

Ritiro del nodo degradato completato

OpenSearch Il servizio invia questo evento quando un nodo danneggiato viene ritirato e sostituito con un nuovo nodo.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"7444215c-90f9-a52d-bcda-e85973a9a762",
   "detail-type":"Amazon OpenSearch Service Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2024-01-11T10:20:30Z",
   "region":"us-east-1",
   "resources":[
```

```
      "arn:aws:es:us-east-1:123456789012:domain/test-node-replacement"
   ],
   "detail":{
      "severity":"Medium",
      "description":"The node has been retired and replaced with a new node.",
      "event":"Degraded Node Retirement Notification",
      "status":"Completed"
   }
}
```

Ritiro del nodo danneggiato non riuscito

OpenSearch Il servizio invia questo evento se il ritiro del nodo danneggiato non riesce.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"c328e9bb-93b9-c0b2-b17a-df527fdf96b6",
   "detail-type":"Amazon OpenSearch Service Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2024-01-11T08:31:38Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-node-replacement"
   ],
   "detail":{
      "severity":"Medium",
      "description":"Node retirement failed. No actions are required from your end. We
 will automatically re-try replacing the node.",
      "event":"Degraded Node Retirement Notification",
      "status":"Failed"
   }
}
```

Eventi di errore di dominio

OpenSearch Il servizio invia gli eventi EventBridge quando si verifica uno dei seguenti errori di dominio.

Errore di convalida dell'aggiornamento del dominio

OpenSearch Il servizio invia questo evento se rileva uno o più errori di convalida durante il tentativo di aggiornare o eseguire una modifica della configurazione su un dominio. Per risolvere questi errori, consulta the section called “Risoluzione degli errori di convalida”.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
   "version":"0",
   "id":"01234567-0123-0123-0123-012345678901",
   "detail-type":"Amazon OpenSearch Service Domain Update Notification",
   "source":"aws.es",
   "account":"123456789012",
   "time":"2016-11-01T13:12:22Z",
   "region":"us-east-1",
   "resources":[
      "arn:aws:es:us-east-1:123456789012:domain/test-domain"
   ],
   "detail":{
      "event":"Domain Update Validation",
      "status":"Failed",
      "severity":"High",
      "description":"Unable to perform updates to your domain due to the following
 validation failures: <failures>
                    Please see the documentation for more information https://
docs.aws.amazon.com/opensearch-service/latest/developerguide/managedomains-
configuration-changes.html#validation"
   }
}
```

Chiave KMS inaccessibile

OpenSearch Il servizio invia questo evento quando non può accedere alla tua chiave. AWS KMS

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
```

```
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Domain Error Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2016-11-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"KMS Key Inaccessible",
     "status":"Error",
     "severity":"High",
     "description":"The KMS key associated with this domain is inaccessible. You are at
 risk of losing access to your domain.
                   For more information, please refer to https://docs.aws.amazon.com/
opensearch-service/latest/developerguide/encryption-at-rest.html#disabled-key."
  }
}
```

Isolamento del dominio

OpenSearch Il servizio invia questo evento quando il dominio diventa isolato e non può ricevere, leggere o scrivere richieste perché non è raggiungibile dalla rete.

Esempio

Di seguito è illustrato un evento di esempio di questo tipo:

```
{
  "version":"0",
  "id":"01234567-0123-0123-0123-012345678901",
  "detail-type":"Amazon OpenSearch Service Notification",
  "source":"aws.es",
  "account":"123456789012",
  "time":"2023-11-01T13:12:22Z",
  "region":"us-east-1",
  "resources":["arn:aws:es:us-east-1:123456789012:domain/test-domain"],
  "detail":{
     "event":"Domain Isolation Notification",
     "status":"Error",
     "severity":"High",
     "description":"Your OpenSearch Service domain has been isolated. An isolated
 domain is unreachable by network and cannot receive, read, or write requests. For more
```

```
 information and assistance, please contact AWS Support at https://docs.aws.amazon.com/
opensearch-service/latest/developerguide/encryption-at-rest.html#disabled-key."
  }
}
```

Tutorial: ascolto degli EventBridge eventi OpenSearch di Amazon Service

In questo tutorial, configuri una semplice AWS Lambda funzione che ascolta gli eventi di Amazon OpenSearch Service e li scrive in un flusso di log di CloudWatch Logs.

Prerequisiti

Questo tutorial presuppone che tu abbia un dominio Service esistente OpenSearch . Se non è stato creato un dominio, attenersi alla procedura descritta in Creazione e gestione dei domini per crearne uno.

Fase 1: Creazione della funzione Lambda

In questa procedura, si crea una semplice funzione Lambda che funga da destinazione per i messaggi degli eventi OpenSearch di servizio.

Per creare una funzione Lambda di destinazione

1. Apri la AWS Lambda console all'indirizzo https://console.aws.amazon.com/lambda/.
2. Scegli Create function (Crea funzione) e Author from scratch (Crea da zero).
3. Per Nome funzione, immettere event-handler.
4. In Runtime, scegli Python 3.8.
5. Scegli Crea funzione.
6. Nella sezione Function code (Codice della funzione), modifica il codice di esempio affinché corrisponda all'esempio seguente:

```
import json

def lambda_handler(event, context):
    if event["source"] != "aws.es":
        raise ValueError("Function only supports input from events with a source
 type of: aws.es")

    print(json.dumps(event))
```

Questa è una semplice funzione Python 3.8 che stampa gli eventi inviati da Service. OpenSearch Se tutto è configurato correttamente, alla fine di questo tutorial, i dettagli dell'evento vengono visualizzati nel flusso di log CloudWatch Logs associato a questa funzione Lambda.

7. Seleziona Implementa.

Fase 2: Registrazione di una regola di evento

In questo passaggio, crei una EventBridge regola che acquisisce gli eventi dai tuoi OpenSearch domini di servizio. Questa regola acquisisce tutti gli eventi all'interno dell'account in cui è definita. Gli stessi messaggi di evento contengono informazioni sull'origine dell'evento, tra cui il dominio da cui ha origine. È possibile utilizzare queste informazioni per filtrare e ordinare gli eventi a livello di programmazione.

Per creare una regola EventBridge

1. Apri la EventBridge console all'indirizzo https://console.aws.amazon.com/events/.
2. Scegli Crea regola.
3. Assegnare alla regola il nome event-rule.
4. Scegli Next (Successivo).
5. Per lo schema dell'evento, seleziona AWS services, Amazon OpenSearch Service e All Events. Questo modello si applica a tutti i domini di OpenSearch servizio e a tutti gli eventi OpenSearch del servizio. In alternativa, è possibile creare una regola più specifica per filtrare alcuni risultati.
6. Premere Next (Successivo).
7. Per la destinazione, scegli Lambda function (Funzione Lambda). Nel menu a discesa della funzione, scegli event-handler.
8. Premere Next (Successivo).
9. Ignorare i tag e premere di nuovo Next (Successivo).
10. Rivedi la configurazione e scegli Create rule (Crea regola).

Fase 3: Test della configurazione

La prossima volta che ricevi una notifica nella sezione Notifiche della console di OpenSearch servizio, se tutto è configurato correttamente, la funzione Lambda viene attivata e scrive i dati dell'evento in un flusso di log di CloudWatch Logs per la funzione.

Per testare la configurazione

1. Apri la console all' CloudWatch indirizzo. https://console.aws.amazon.com/cloudwatch/
2. Nel pannello di navigazione, scegli Logs e seleziona il gruppo di log per la tua funzione Lambda (ad esempio,aws/lambda/event/-handler).
3. Seleziona un flusso di log per visualizzare i dati di evento.

Esercitazione: Invio di avvisi Amazon SNS per gli aggiornamenti software disponibili

In questo tutorial, configuri una regola di EventBridge evento Amazon che acquisisce le notifiche per gli aggiornamenti del software di servizio disponibili in Amazon OpenSearch Service e ti invia una notifica e-mail tramite Amazon Simple Notification Service (Amazon SNS).

Prerequisiti

Questo tutorial presuppone che tu disponga di un dominio di servizio esistente OpenSearch . Se non è stato creato un dominio, attenersi alla procedura descritta in Creazione e gestione dei domini per crearne uno.

Fase 1: Creazione e sottoscrizione a un argomento Amazon SNS

Configurare un argomento Amazon SNS che funga da destinazione evento per la nuova regola di evento.

Per creare una destinazione Amazon SNS

1. Apri la console Amazon SNS nella versione v3/home. https://console.aws.amazon.com/sns/
2. Scegliere Argomenti e Crea argomento.
3. Per il tipo di processo, scegliere Standard e assegnare al processo il nome software-update.
4. Scegli Create topic (Crea argomento).
5. Una volta creato l'argomento, scegliere Crea una sottoscrizione.
6. Per Protocollo, scegli E-mail. In Endpoint, inserire un indirizzo e-mail a cui si ha accesso correntemente, quindi scegliere Crea sottoscrizione.
7. Controllare l'account e-mail e attendere di ricevere una e-mail di conferma della sottoscrizione. Una volta ricevuta, seleziona Confirm subscription (Conferma sottoscrizione).

Fase 2: Registrazione di una regola di evento

Quindi, registrare una regola di evento che acquisisca solo gli eventi di aggiornamento del software di servizio.

Per creare una regola di evento

1. Apri la console all' EventBridge indirizzo. https://console.aws.amazon.com/events/
2. Scegli Crea regola.
3. Assegnare alla regola il nome softwareupdate-rule.
4. Scegli Next (Successivo).
5. Per il modello di evento, seleziona AWS servizi, Amazon OpenSearch Service (cluster gestito) e Amazon OpenSearch Service Software Update Notification. Questo schema corrisponde a qualsiasi evento di aggiornamento del software di OpenSearch servizio fornito da Service. Per ulteriori informazioni sui modelli di eventi, consulta i modelli di EventBridge eventi di Amazon nella Amazon EventBridge User Guide.
6. Facoltativamente, è possibile solo in base alle gravità specifiche. Per le gravità di ogni evento, consulta the section called “Eventi di aggiornamento del software di servizio”.
7. Scegli Next (Successivo).
8. Per destinazione, scegli SNS topic (Argomento SNS) e seleziona software-update.
9. Scegli Next (Successivo).
10. Ignora i tag e scegli Next (Successivo).
11. Rivedi la configurazione della regola e scegli Create rule (Crea regola).

La prossima volta che ricevi una notifica dal OpenSearch Servizio su un aggiornamento del software di servizio disponibile, se tutto è configurato correttamente, Amazon SNS dovrebbe inviarti un'e-mail di avviso sull'aggiornamento.

Monitoraggio delle chiamate API di Amazon OpenSearch Service con AWS CloudTrail

Amazon OpenSearch Service si integra con AWS CloudTrail un servizio che fornisce un registro delle azioni intraprese da un utente, un ruolo o un AWS servizio nel OpenSearch Servizio. CloudTrail acquisisce tutte le chiamate API di configurazione per OpenSearch Service come eventi.

> **Note**
>
> CloudTrail acquisisce solo le chiamate all'API di configurazione, ad esempio `eCreateDomain`. `GetUpgradeStatus` CloudTrail non acquisisce le chiamate verso OpenSearch APIs, ad esempio `_search` and`_bulk`. Per queste chiamate, consultare the section called “Monitoraggio dei log di verifica”.

Le chiamate acquisite includono chiamate dalla console OpenSearch di servizio o da un AWS SDK. AWS CLI Se crei un trail, puoi abilitare la distribuzione continua di CloudTrail eventi a un bucket Amazon S3, inclusi gli eventi per Service. OpenSearch Se non configuri un percorso, puoi comunque visualizzare gli eventi più recenti sulla CloudTrail console nella cronologia degli eventi. Utilizzando le informazioni raccolte da CloudTrail, puoi determinare la richiesta che è stata effettuata al OpenSearch Servizio, l'indirizzo IP da cui è stata effettuata la richiesta, chi ha effettuato la richiesta, quando è stata effettuata e dettagli aggiuntivi.

Per ulteriori informazioni CloudTrail, consulta la Guida AWS CloudTrail per l'utente.

Informazioni sul OpenSearch servizio Amazon in CloudTrail

CloudTrail è abilitato sul tuo account al Account AWS momento della creazione dell'account. Quando si verifica un'attività nel OpenSearch Servizio, tale attività viene registrata in un CloudTrail evento insieme ad altri eventi di AWS servizio nella Cronologia degli eventi. È possibile visualizzare, cercare e scaricare gli eventi recenti nell'account Account AWS . Per ulteriori informazioni, vedere Visualizzazione degli eventi con la cronologia degli CloudTrail eventi.

Per una registrazione continua degli eventi nel tuo Account AWS account, inclusi gli eventi di OpenSearch Service for Service, crea un percorso. Un trail consente di CloudTrail inviare file di log a un bucket Amazon S3. Per impostazione predefinita, quando si crea un percorso nella console, questo sarà valido in tutte le Regioni AWS. Il trail registra gli eventi di tutte le regioni della AWS partizione e consegna i file di log al bucket Amazon S3 specificato. Inoltre, puoi configurare altri AWS servizi per analizzare ulteriormente e agire in base ai dati sugli eventi raccolti nei log. CloudTrail Per ulteriori informazioni, consultare gli argomenti seguenti:

- Creare un percorso per il tuo Account AWS
- AWS integrazioni di servizi con Logs CloudTrail
- Configurazione delle notifiche Amazon SNS per CloudTrail

- Ricezione di file di CloudTrail registro da più regioni e ricezione di file di CloudTrail registro da più account

Tutte le azioni dell'API di configurazione del OpenSearch servizio vengono registrate CloudTrail e documentate nell'Amazon OpenSearch Service API Reference.

Ogni evento o voce di log contiene informazioni sull'utente che ha generato la richiesta. Le informazioni di identità consentono di determinare quanto segue:

- Se la richiesta è stata effettuata con credenziali utente root o AWS Identity and Access Management (IAM)
- Se la richiesta è stata effettuata con le credenziali di sicurezza temporanee per un ruolo o un utente federato.
- Se la richiesta è stata effettuata da un altro AWS servizio

Per ulteriori informazioni, consulta Elemento CloudTrail userIdentity.

Informazioni sulle voci dei file OpenSearch di registro di Amazon Service

Un trail è una configurazione che consente la distribuzione di eventi come file di log in un bucket Amazon S3 specificato dall'utente. CloudTrail i file di registro contengono una o più voci di registro. Un evento rappresenta una singola richiesta proveniente da qualsiasi fonte e include informazioni sull'azione richiesta, la data e l'ora dell'azione, i parametri della richiesta e così via. CloudTrail i file di registro non sono una traccia ordinata dello stack delle chiamate API pubbliche, quindi non vengono visualizzati in un ordine specifico.

L'esempio seguente mostra una voce di CloudTrail registro che dimostra l'`CreateDomain`operazione:


```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "AIDACKCEVSQ6C2EXAMPLE",
    "arn": "arn:aws:iam::123456789012:user/test-user",
    "accountId": "123456789012",
    "accessKeyId": "access-key",
    "userName": "test-user",
    "sessionContext": {
```

```
      "attributes": {
        "mfaAuthenticated": "false",
        "creationDate": "2018-08-21T21:59:11Z"
      }
    },
    "invokedBy": "signin.amazonaws.com"
  },
  "eventTime": "2018-08-21T22:00:05Z",
  "eventSource": "es.amazonaws.com",
  "eventName": "CreateDomain",
  "awsRegion": "us-west-1",
  "sourceIPAddress": "123.123.123.123",
  "userAgent": "signin.amazonaws.com",
  "requestParameters": {
    "engineVersion": "OpenSearch_1.0",
    "clusterConfig": {
      "instanceType": "m4.large.search",
      "instanceCount": 1
    },
    "snapshotOptions": {
      "automatedSnapshotStartHour": 0
    },
    "domainName": "test-domain",
    "encryptionAtRestOptions": {},
    "eBSOptions": {
      "eBSEnabled": true,
      "volumeSize": 10,
      "volumeType": "gp2"
    },
    "accessPolicies": "{\"Version\":\"2012-10-17\",\"Statement\":[{\"Effect\":\"Allow
\",\"Principal\":{\"AWS\":[\"123456789012\"]},\"Action\":[\"es:*\"],\"Resource\":
\"arn:aws:es:us-west-1:123456789012:domain/test-domain/*\"}]}",
    "advancedOptions": {
      "rest.action.multi.allow_explicit_index": "true"
    }
  },
  "responseElements": {
    "domainStatus": {
      "created": true,
      "clusterConfig": {
        "zoneAwarenessEnabled": false,
        "instanceType": "m4.large.search",
        "dedicatedMasterEnabled": false,
        "instanceCount": 1
```

```
      },
      "cognitoOptions": {
        "enabled": false
      },
      "encryptionAtRestOptions": {
        "enabled": false
      },
      "advancedOptions": {
        "rest.action.multi.allow_explicit_index": "true"
      },
      "upgradeProcessing": false,
      "snapshotOptions": {
        "automatedSnapshotStartHour": 0
      },
      "eBSOptions": {
        "eBSEnabled": true,
        "volumeSize": 10,
        "volumeType": "gp2"
      },
      "engineVersion": "OpenSearch_1.0",
      "processing": true,
      "aRN": "arn:aws:es:us-west-1:123456789012:domain/test-domain",
      "domainId": "123456789012/test-domain",
      "deleted": false,
      "domainName": "test-domain",
      "accessPolicies": "{\"Version\":\"2012-10-17\",\"Statement\":[{\"Effect\":\"Allow
\",\"Principal\":{\"AWS\":\"arn:aws:iam::123456789012:root\"},\"Action\":\"es:*\",
\"Resource\":\"arn:aws:es:us-west-1:123456789012:domain/test-domain/*\"}]}"
    }
  },
  "requestID": "12345678-1234-1234-1234-987654321098",
  "eventID": "87654321-4321-4321-4321-987654321098",
  "eventType": "AwsApiCall",
  "recipientAccountId": "123456789012"
}
```

Monitoraggio operativo unificato con Cluster Insights

Amazon OpenSearch Service ora include Cluster Insights, una soluzione di monitoraggio che offre una visibilità operativa completa dei cluster tramite un'unica dashboard. Ciò elimina la complessità di dover analizzare e correlare vari log e metriche per identificare potenziali rischi per la disponibilità o le prestazioni del cluster. La soluzione automatizza il consolidamento dei dati operativi critici tra nodi, indici e shard, trasformando la risoluzione dei problemi complessi in un processo semplificato. È possibile rilevare problemi come frammenti di grandi dimensioni e filigrane su disco insufficienti, visualizzare metriche dettagliate a livello di nodo, indice e shard e accedere alle migliori pratiche di sicurezza e resilienza.

> **Note**
>
> Cluster Insights è disponibile tramite l'interfaccia utente del OpenSearch servizio senza costi aggiuntivi per tutti gli utenti che utilizzano OpenSearch la versione 2.17 o successiva. Inoltre, i domini con OpenSearch le versioni 2.17 e 2.19 devono disporre dell'ultimo aggiornamento del software di servizio.

Vantaggi

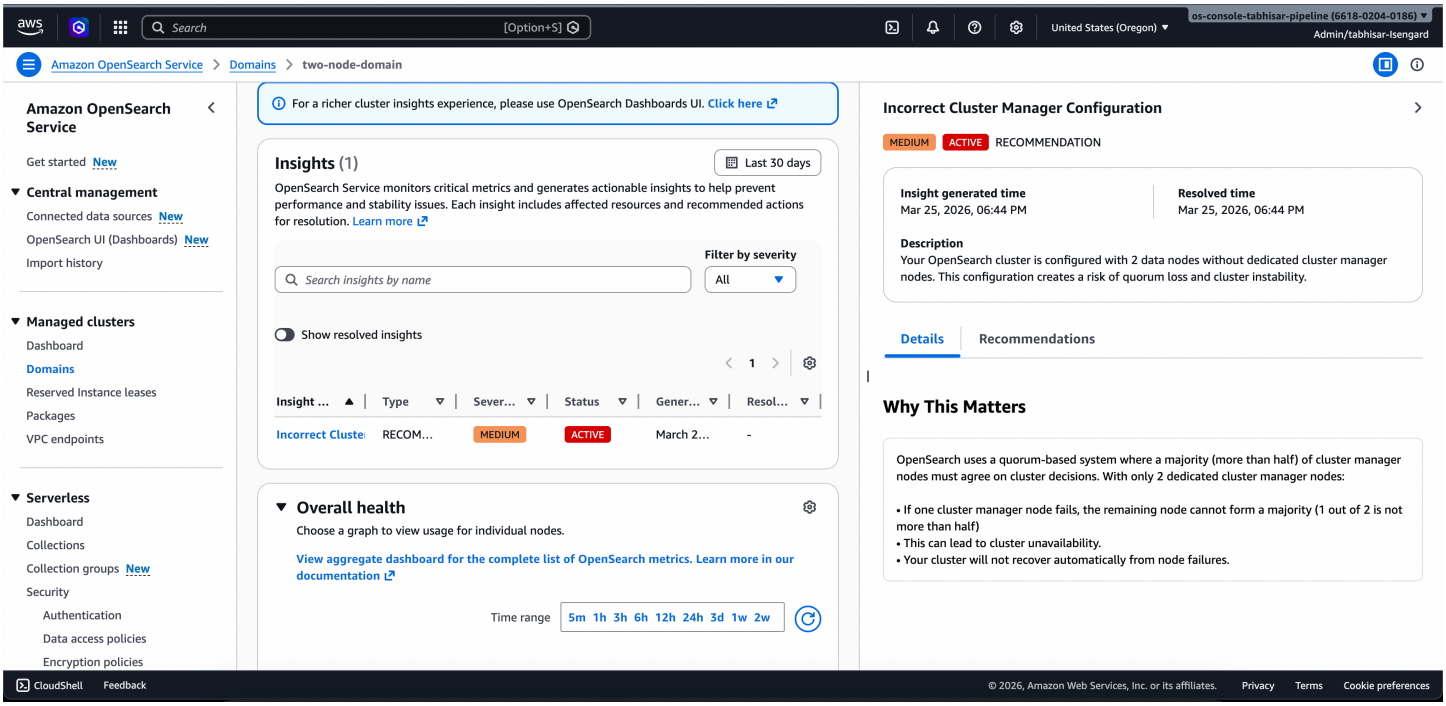
- Monitoraggio proattivo: monitora lo stato del cluster in modo proattivo con metriche dettagliate delle prestazioni su tutti i componenti, dai singoli nodi e indici agli shard e alle query di ricerca.
- Visibilità unificata: consolida i dati di monitoraggio in un'unica dashboard
- Consigli pratici: ottieni step-by-step indicazioni per la risoluzione dei problemi
- Copertura completa: monitora la sicurezza, la stabilità e la resilienza tra i cluster OpenSearch
- Ottimizzazione delle query: identifica le query che richiedono molte risorse e ottimizza le prestazioni

Con Cluster Insights, è possibile mantenere prestazioni ottimali dei cluster, ridurre il sovraccarico operativo e garantire best practice coerenti tra i cluster OpenSearch

Accedi a Cluster Insights tramite Console

Rivedi i consigli su prestazioni e resilienza e apporta le modifiche di configurazione necessarie, il tutto all'interno della stessa console. Nella Console, nella scheda Cluster Health, puoi accedere a Cluster Insights che elenca tutti gli Insights attivi. Fai clic su qualsiasi Insight per visualizzare i consigli.

Screen-1: Cluster Insights nella scheda Cluster Health



Accedi a Cluster Insights e alle metriche dettagliate tramite l'interfaccia utente OpenSearch

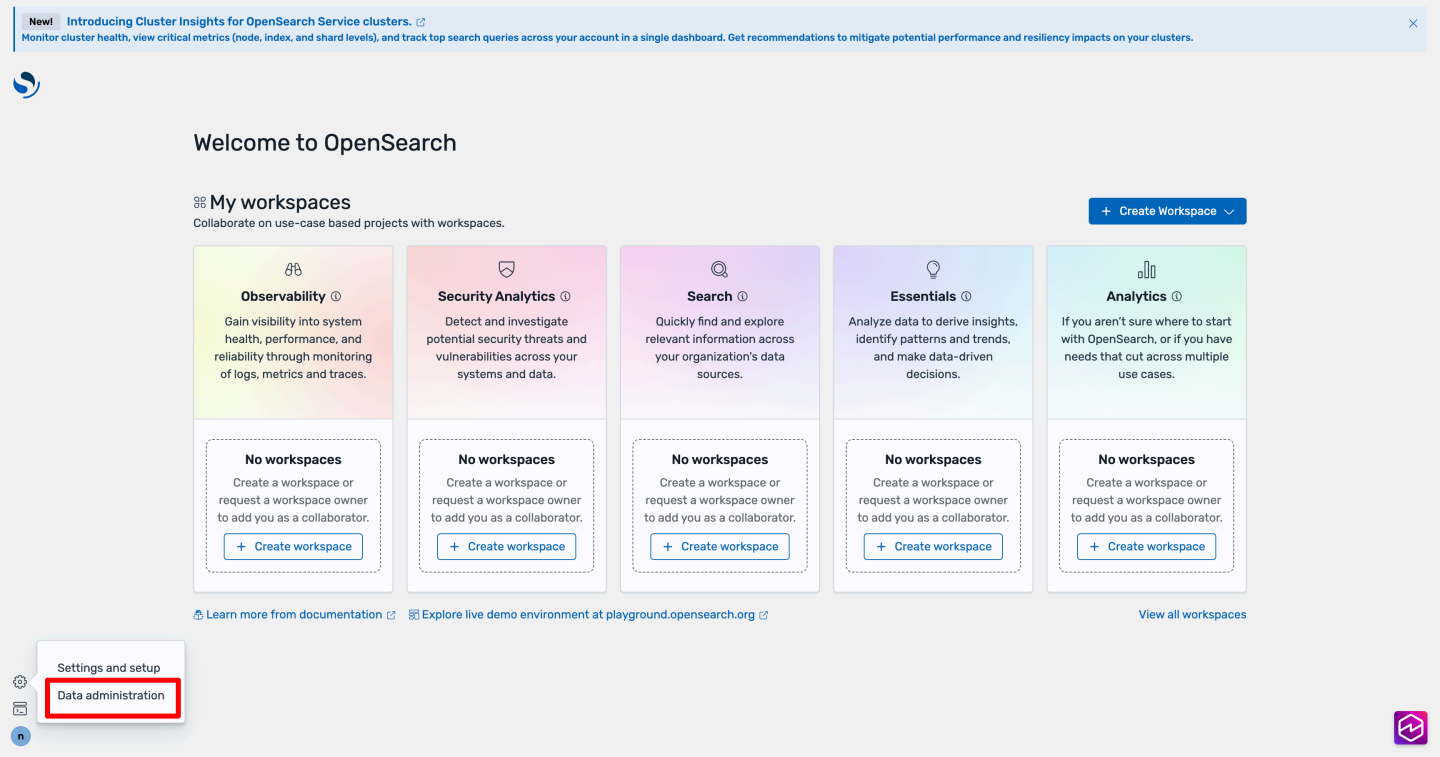
Puoi visualizzare le informazioni dettagliate per uno specifico cluster OpenSearch di servizi tramite l'OpenSearch interfaccia utente (dashboard). Nell' OpenSearch interfaccia utente, un'applicazione è semplicemente un costrutto organizzativo come una cartella. Ogni applicazione può connettersi e visualizzare informazioni dettagliate per più cluster OpenSearch di servizi. L'accesso a Cluster Insights richiede un ruolo amministrativo nell'applicazione OpenSearch UI.

> **Note**
>
> L'accesso a Cluster Insights richiede un ruolo amministrativo nell'applicazione OpenSearch UI.

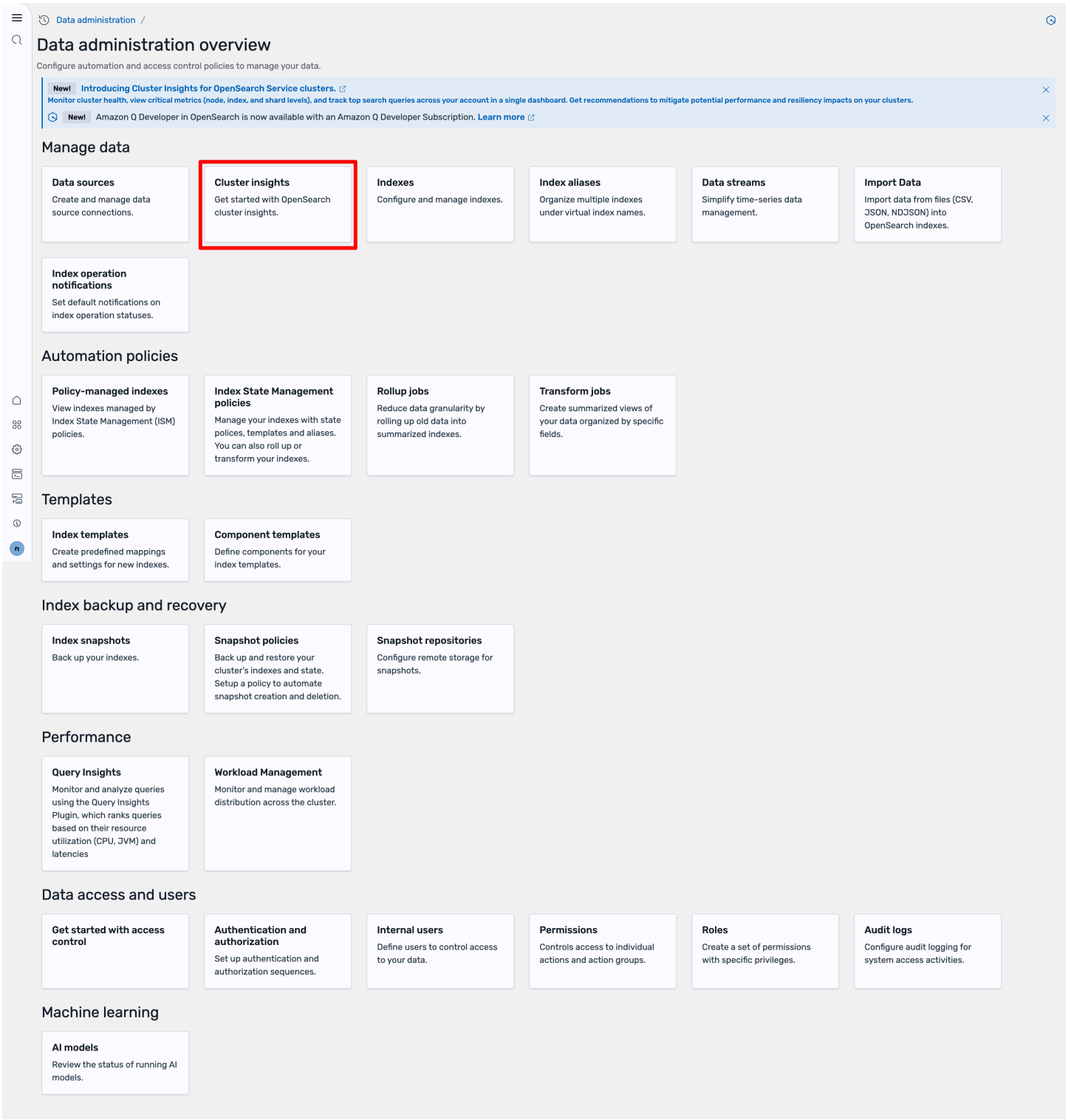
Crea e configura un'applicazione per visualizzare Cluster Insights

1. Apri la console OpenSearch di servizio all'indirizzo https://console.aws.amazon.com/aos/home
2. Scegli OpenSearch UI (Dashboards) dalla barra di navigazione a sinistra
3. Completa i seguenti passaggi per creare e configurare un'applicazione:
    1. Creare un'applicazione OpenSearch di servizio
    2. Associa fonti di dati
4. Dopo aver completato i due passaggi precedenti, puoi visualizzare Cluster Insights nella dashboard OpenSearch dell'interfaccia utente nella sezione Impostazioni > Amministratore dati > Cluster Insights. L'icona Impostazioni si trova nella parte inferiore sinistra della schermata dell' OpenSearch interfaccia utente.

Screen-2: Accedi a Data Administrator dall'interfaccia utente OpenSearch



Screen-3: Cluster Insights nella sezione Gestisci dati



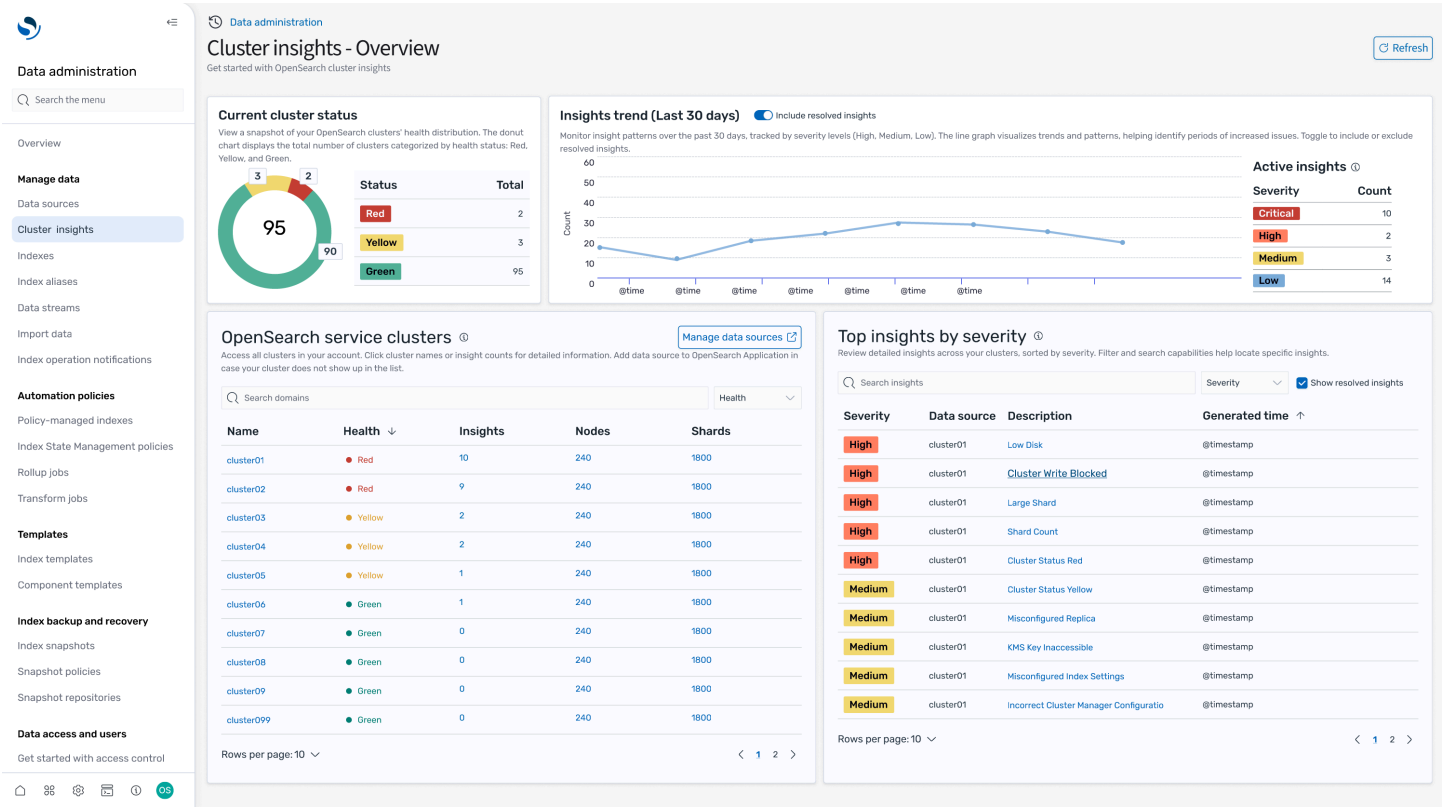
Comprendere Cluster Insights

Questa sezione descrive i vari approfondimenti disponibili in Cluster Insights.

Dashboard di panoramica

La pagina Panoramica di Cluster Insights, come illustrato nella schermata seguente, offre una panoramica di alto livello dello stato del cluster a livello di applicazione e comprende le seguenti sezioni:

Screen-4: pagina iniziale di Cluster Insights nell'applicazione UI. OpenSearch



Stato attuale del cluster

Un grafico ad anello mostra lo stato di salute del cluster:

- Verde: tutti gli shard e le repliche primari sono allocati ai nodi
- Giallo: tutti gli shard primari vengono allocati, ma alcune repliche no
- Rosso: almeno uno shard primario non è allocato a nessun nodo

Andamento degli approfondimenti

Il grafico delle tendenze tiene traccia dei modelli di problemi degli ultimi 30 giorni, aiutandoti a identificare i problemi emergenti e a monitorare i progressi nella risoluzione.

Informazioni aperte attuali

Un conteggio organizzato in base alla gravità degli approfondimenti aperti negli ultimi 30 giorni.

OpenSearch Cluster di servizi

Questa sezione elenca tutti i OpenSearch cluster con le statistiche chiave, tra cui il numero di nodi, il conteggio degli shard e le query attive.

Le migliori informazioni in base alla gravità

Puoi esaminare gli approfondimenti in tutti i domini della tua applicazione. Questa sezione dà priorità ai problemi che richiedono attenzione immediata (critici e di gravità elevata). Ogni approfondimento include una descrizione e consigli specifici, che possono aiutarti a concentrarti prima sulle questioni critiche.

Dettagli degli approfondimenti

Ogni approfondimento nella sezione Principali approfondimenti per gravità è interattivo e fornisce un'analisi dettagliata. Ad esempio, quando scegli l'analisi Large Shard Size:

1. Potete vedere quanti shard superano la soglia e quali indici sono interessati.
2. Una mappa delle risorse identifica ogni frammento sovradimensionato con il relativo indice, ID e dimensione corrente.
3. La scheda dei consigli fornisce step-by-step indicazioni per la correzione.
4. La scheda Cronologia mostra una cronologia delle azioni di riparazione delle risorse.

Dettagli del cluster

Quando si seleziona un cluster specifico nella sezione Cluster di OpenSearch servizi, OpenSearch visualizza le informazioni dettagliate relative a quel cluster nelle seguenti schede: Stato del cluster, visualizzazione Nodi, visualizzazione Indice, Visualizzazione Shard e Visualizzazione Query. La scheda Cluster Health mostra le seguenti informazioni:

Panoramica di

Le informazioni chiave includono lo stato del cluster, il conteggio degli shard, il conteggio dei nodi, il conteggio degli indici e le statistiche dei documenti.

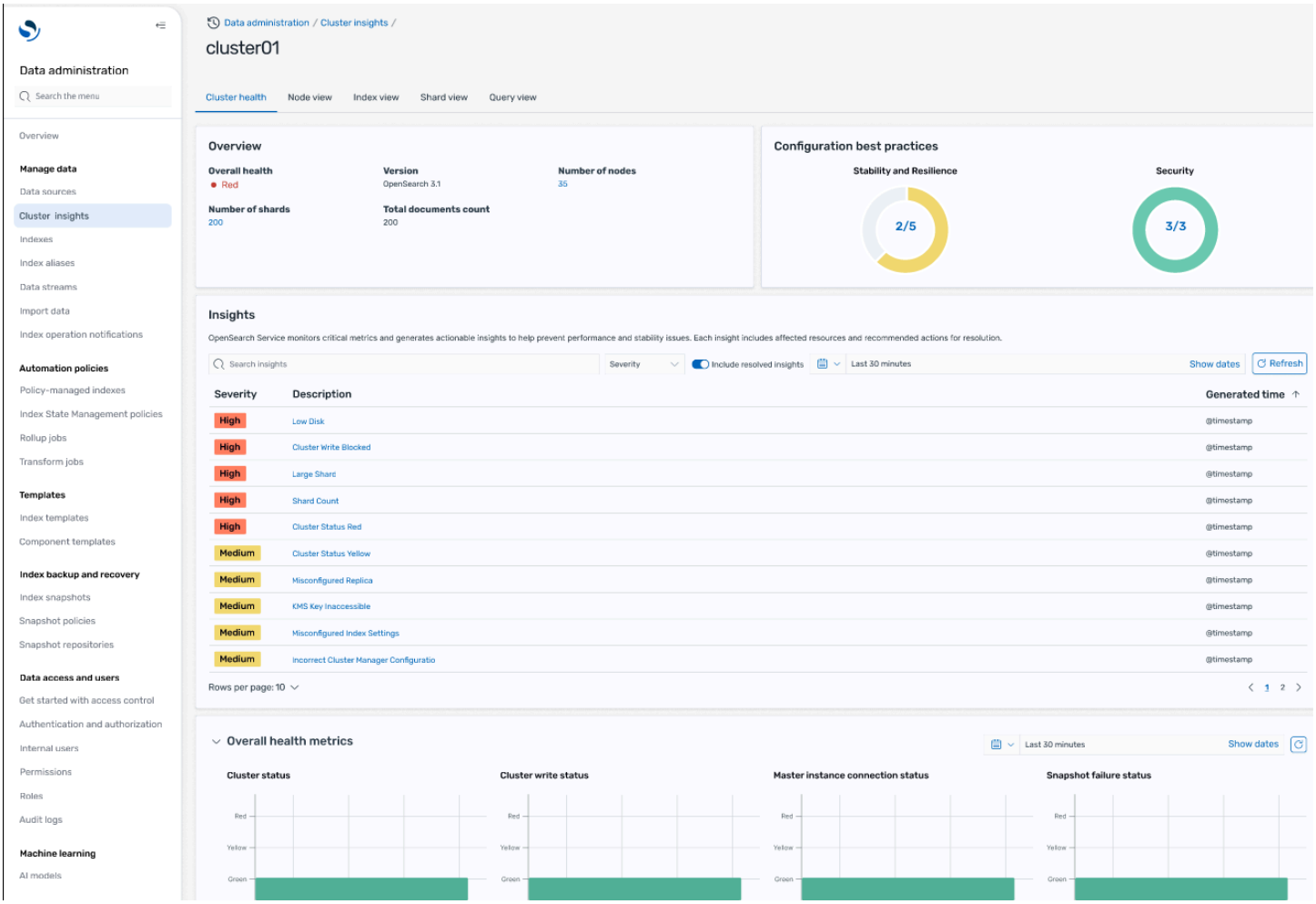
Best practice per la configurazione

I grafici Donut mostrano la conformità con le impostazioni consigliate per la resilienza e la sicurezza.

Approfondimenti

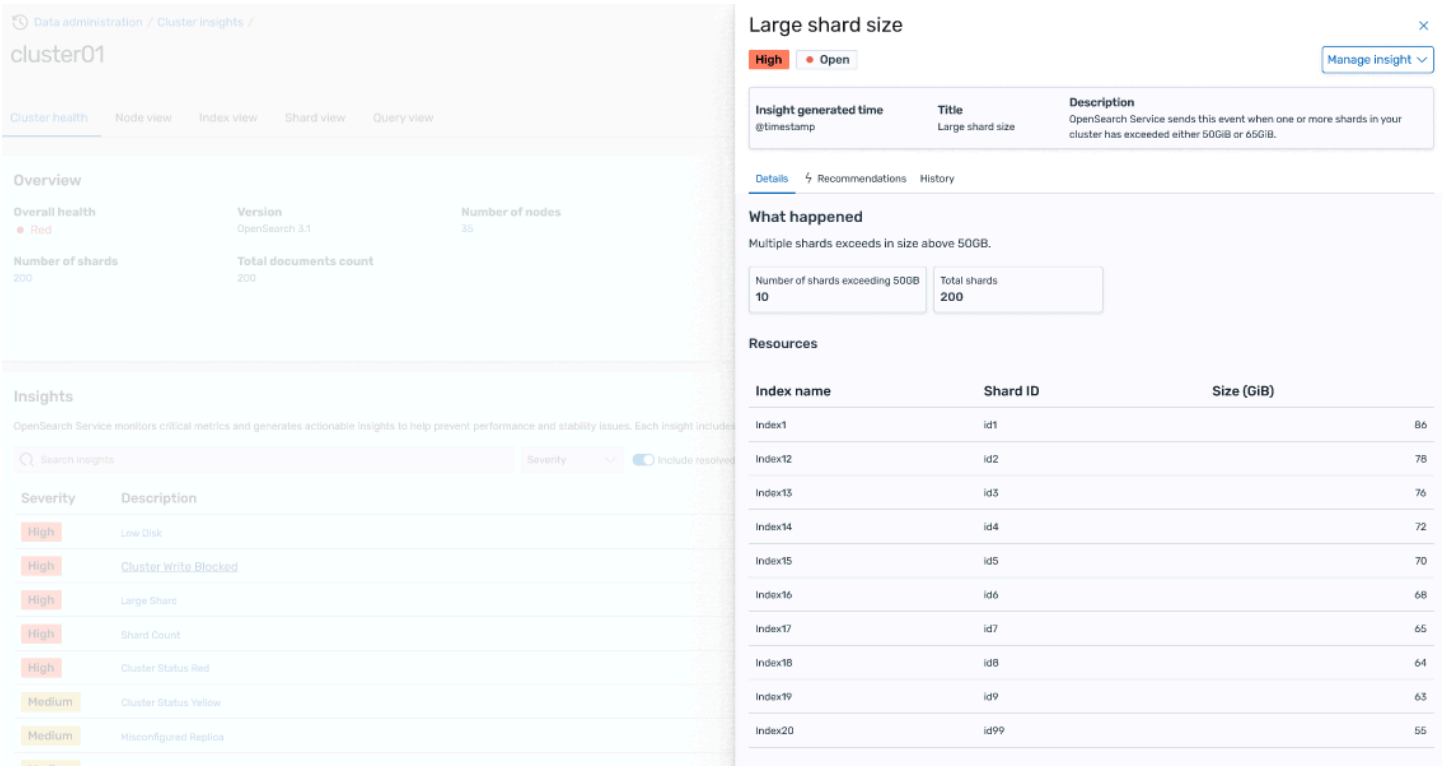
Una tabella elenca gli approfondimenti recenti generati per il cluster, con le stesse indicazioni dettagliate sulla suddivisione e la correzione disponibili nella pagina di panoramica.

Screen-5: La panoramica di Cluster Health fornisce metriche chiave, best practice e approfondimenti



Quando fai clic su qualsiasi approfondimento, puoi visualizzare i dettagli, le risorse interessate e i consigli. Inoltre, puoi anche visualizzare la cronologia delle risorse fisse.

Screen-6: dettagli di Insight. Fornisce dettagli, consigli e cronologia storica.



Sezione Metriche

I grafici interattivi in questa sezione mostrano le seguenti metriche del cluster:

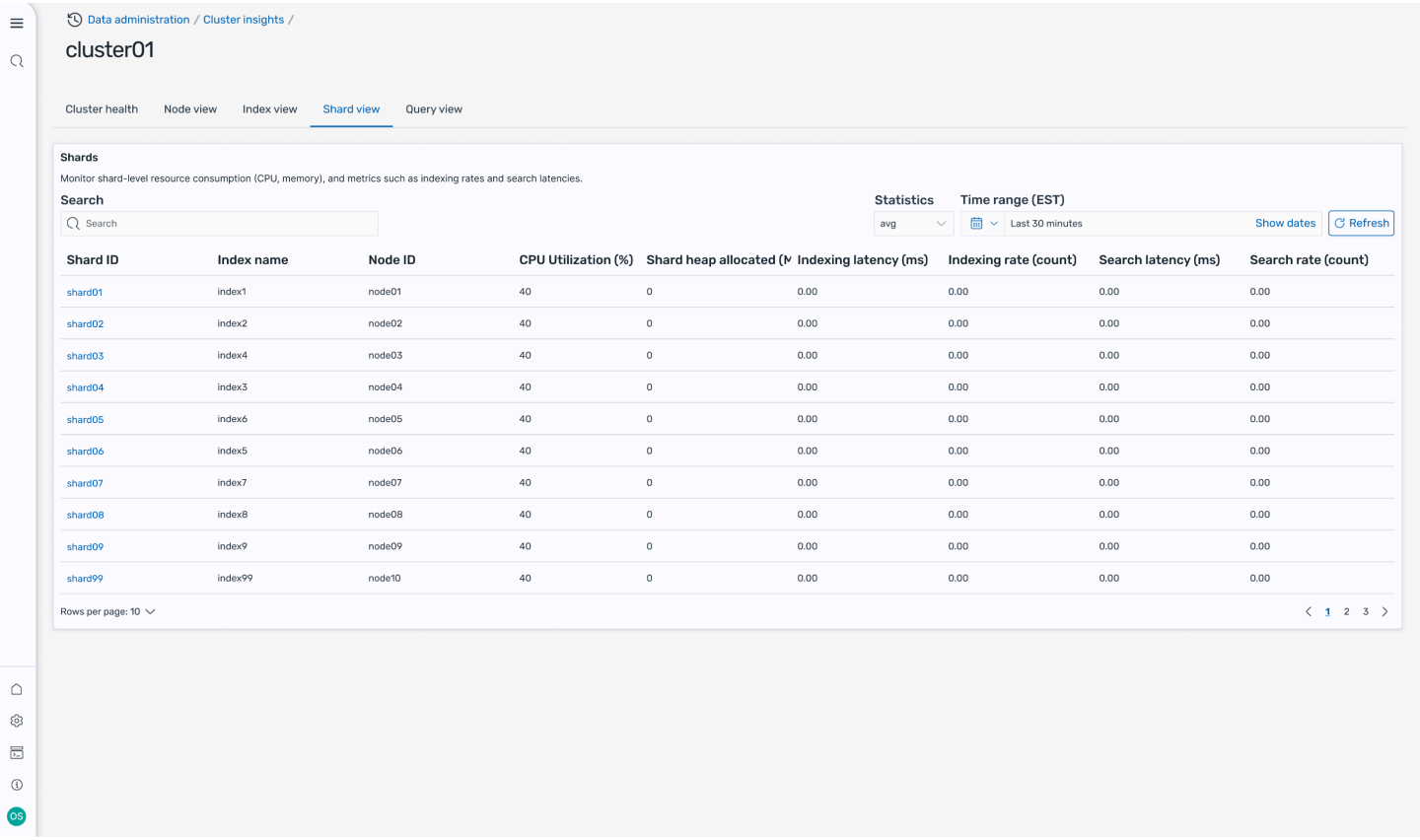
- Metriche generali sullo stato del cluster, ad esempio Cluster Status, Write status e documenti ricercabili
- KPIs (Indicatori chiave di prestazione) come tassi e latenze di indicizzazione e ricerca
- Metriche di utilizzo delle risorse come JVM e utilizzo della CPU

Visualizzazioni di nodi, indici e frammenti

Le viste Node, Index e Shard utilizzano le OpenSearch statistiche per fornire una visibilità dettagliata sulle operazioni del cluster. Puoi visualizzare:

- Metriche in tempo reale come l'utilizzo della CPU e la pressione della memoria JVM
- Ricerca e indicizzazione dei dati sulle prestazioni
- Hotspot di risorse tra i componenti del cluster
- Diagnostica granulare a livello di nodo
- Top shard heap allocato

Screen-7: metriche a livello di nodo, indice e shard



Visualizzazione delle interrogazioni

> **Note**
>
> La funzionalità Query View è supportata per OpenSearch le versioni 2.19 o successive.

La pagina Query View consente di monitorare le query che richiedono molte risorse con:

Configurazione degli accessi

La visualizzazione delle prime N query richiede autorizzazioni di controllo degli accessi granulari. Assicurati quanto segue:

- Il controllo granulare degli accessi è abilitato sul tuo dominio.
- Il tuo ruolo IAM (o utente interno) è mappato su un OpenSearch ruolo con le autorizzazioni del cluster richieste per l'analisi delle query.

- Per un accesso amministrativo completo, associa l'ARN del tuo ruolo IAM come ruolo di backend a entrambi i ruoli all_access e security_manager. Puoi farlo nelle OpenSearch dashboard in Sicurezza > Ruoli > seleziona il ruolo > Utenti mappati > Gestisci la mappatura o utilizzando l'API di sicurezza (PUT _plugins/_ _access). security/api/rolesmapping/all

Senza una corretta mappatura dei ruoli, gli utenti potrebbero ricevere 403 errori proibiti quando tentano di accedere ai dati di Query Insights. Per i dettagli, consulta Controllo granulare degli accessi.

Dashboard live

Visualizza le statistiche di esecuzione, l'utilizzo della CPU e della memoria e l'avanzamento del completamento per ogni query.

Query top N

Una tabella classificata mostra le query più significative con dettagli tra cui:

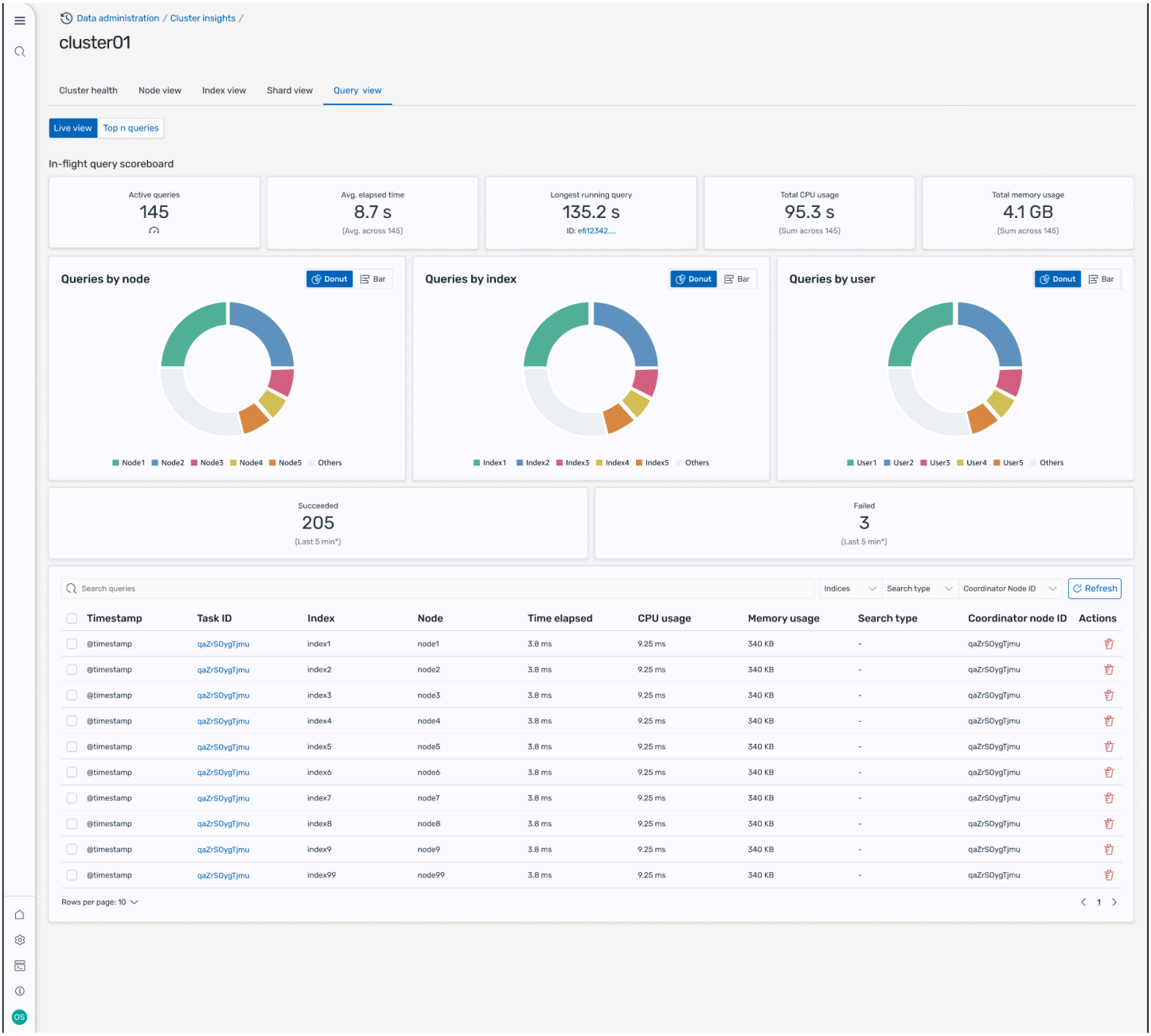
- Numero di interrogazioni
- Latenza, utilizzo della CPU e della memoria
- Tipo di ricerca e nodo coordinatore
- Indici di destinazione e numero di frammenti

Dettagli della query

Fate doppio clic su una query per vedere:

- Pagamento esatto dell'interrogazione e fasi di esecuzione
- Suddivisione della latenza per ogni fase (espansione, interrogazione, recupero)
- Consigli per l'ottimizzazione

Screen-8: visualizzazione dal vivo in volo. Puoi anche visualizzare le query Top-N



Accedi a Insights tramite EventBridge eventi Amazon

Puoi monitorare gli approfondimenti tramite Amazon EventBridge Events. Per ulteriori dettagli, consulta le notifiche.

Catalogo Insights

Cluster Insights offre le seguenti informazioni utili sui cluster interessati

Approfondimenti	Description
Disco insufficiente	Quando la capacità del cluster FreeStorage scende al di sotto del 25% della capacità totale o di 25 GB, a seconda di quale soglia viene raggiunta per prima, viene generata un'analisi di elevata gravità.
Scrittura del cluster bloccata	Una panoramica sulla gravità critica viene generata quando la capacità del cluster FreeStorage scende al di sotto del 20% della capacità totale o di 20 GB, a seconda della soglia raggiunta per prima.
Shard di grandi dimensioni	Un insight con bassa severità viene generato quando la dimension e dello shard è compresa tra 50 GB e 100 GB. La gravità passa a Media quando la dimensione dello shard è compresa tra 100 GB e 200 GB e aumenta ulteriormente a Alta quando la dimensione dello shard supera i 200 GB.
Numero di frammenti	Un insight di elevata severità viene generato quando il cluster supera i 25 shard per GB di heap configurato o quando il numero di shard per nodo raggiunge o supera 900. Per la OpenSearch versione 2.17 e successive, queste informazioni verranno generate in base alla configurazione del gestore del cluster, come dettagliato nella presente documentazione.
Stato del cluster (rosso)	Una visione della gravità critica viene generata quando la sostituzi one di un nodo danneggiato sul OpenSearch cluster determina lo stato del cluster rosso. Ciò richiede un'azione immediata per ripristinare lo stato verde del cluster.
Stato del cluster (giallo)	Una visione di livello medio viene generata quando la sostituzione di un nodo danneggiato sul OpenSearch cluster determina lo stato giallo del cluster.
Replica configurata in modo errato	Una panoramica di livello medio viene generata quando nel cluster non sono configurate repliche o quando il numero di repliche supera il numero totale di nodi di dati.

Approfondimenti	Description
Chiave KMS inaccessibile	Una visione della gravità critica viene generata quando il OpenSearch cluster perde l'accesso alla chiave KMS utilizzata per la crittografia a riposo. Ciò impedisce al cluster di accedere ai dati crittografati e influisce sulle operazioni del cluster, richiedendo un'azione immediata per ripristinare l'accesso.
Impostazioni dell'indice configurate in modo errato	Quando si imposta l'indice «index.routing.allocation», viene generata una visione di gravità media. require/include/exclude. **<**attribute**>**"esiste su qualsiasi indice del OpenSearch cluster. Questa impostazione può impedire il successo delle distribuzioni blu-verdi durante gli aggiornamenti del cluster.
Configurazione errata di Cluster Manager	Una visione della gravità critica viene generata quando il OpenSearch cluster ha 2 nodi di dati senza nodi di gestione del cluster dedicati o 2 nodi di gestione del cluster dedicati. Queste configurazioni creano un rischio di perdita del quorum e instabilità del cluster.
Strategia di sharding non ottimale	(a) Viene generata un'analisi di media gravità che indica che uno o più indici hanno un numero di frammenti inferiore all'80% dei nodi di dati. Ciò può comportare una distribuzione non ottimale del carico di lavoro nel cluster. (b) Viene generata un'analisi di media gravità che indica che uno o più indici hanno frammenti con il 20% in più di dati rispetto alla dimensione media dei frammenti per nodo. Ciò può comportare una distribuzione non ottimale del carico di lavoro nel cluster.

Approfondimenti	Description
Sovraccarico del cluster	(a) Un'analisi dettagliata del sovraccarico del cluster indica che oltre il 70% dei nodi di dati del cluster registra un utilizzo elevato della CPU, della memoria, dell'I/O del disco, della velocità effettiva del disco o dell'utilizzo del disco. Ciò indica che il cluster potrebbe aver raggiunto i limiti di capacità. (b) Un'analisi della gravità critica indica che oltre il 70% dei nodi di dati sta registrando un elevato utilizzo delle risorse in termini di CPU, memoria, I/O del disco, velocità effettiva del disco o utilizzo del disco e si osservano limitazioni o rifiuti delle richieste. Ciò richiede un'azione immediata per scalare il cluster.

Sicurezza in Amazon OpenSearch Service

La sicurezza del cloud AWS è la massima priorità. In qualità di AWS cliente, puoi beneficiare di un data center e di un'architettura di rete progettati per soddisfare i requisiti delle organizzazioni più sensibili alla sicurezza.

La sicurezza è una responsabilità condivisa tra AWS te e te. Il modello di responsabilità condivisa descrive questo come sicurezza del cloud e sicurezza nel cloud:

- Sicurezza del cloud: AWS è responsabile della protezione dell'infrastruttura che gestisce AWS i servizi nel AWS cloud. AWS ti fornisce anche servizi che puoi utilizzare in modo sicuro. I revisori di terze parti testano e verificano regolarmente l'efficacia della sicurezza come parte dei programmi di conformitàAWS. Per maggiori informazioni sui programmi di conformità che si applicano ad Amazon OpenSearch Service, consulta AWS Services in Scope by Compliance Program.
- Sicurezza nel cloud: la tua responsabilità è determinata dal AWS servizio che utilizzi. L'utente è anche responsabile di altri fattori, tra cui la riservatezza dei dati, i requisiti della propria azienda e le leggi e normative vigenti.

Questa documentazione aiuta a capire come applicare il modello di responsabilità condivisa quando si utilizza il OpenSearch Servizio. I seguenti argomenti mostrano come configurare il OpenSearch servizio per soddisfare gli obiettivi di sicurezza e conformità. Scopri anche come utilizzare altri AWS servizi che ti aiutano a monitorare e proteggere le risorse OpenSearch del Servizio.

Argomenti

- Protezione dei dati in Amazon OpenSearch Service
- Identity and Access Management in Amazon OpenSearch Service
- Prevenzione del problema "confused deputy" tra servizi
- Controllo granulare degli accessi in Amazon Service OpenSearch
- Convalida della conformità per Amazon Service OpenSearch
- Resilienza in Amazon Service OpenSearch
- Autenticazione e autorizzazione JWT per Amazon Service OpenSearch
- Sicurezza dell'infrastruttura in Amazon OpenSearch Service
- Autenticazione SAML per dashboard OpenSearch
- IAM Identity Center Trusted Identity Propagation Support per OpenSearch

- Configurazione dell'autenticazione Amazon Cognito per dashboard OpenSearch
- Utilizzo di ruoli collegati ai servizi per Amazon Service OpenSearch

Protezione dei dati in Amazon OpenSearch Service

Il modello di responsabilità AWS di si applica alla protezione dei dati in Amazon OpenSearch Service. Come descritto in questo modello, AWS è responsabile della protezione dell'infrastruttura globale che gestisce tutti i Cloud AWS. L'utente è responsabile del controllo dei contenuti ospitati su questa infrastruttura. L'utente è inoltre responsabile della configurazione della protezione e delle attività di gestione per i Servizi AWS utilizzati. Per maggiori informazioni sulla privacy dei dati, consulta le Domande frequenti sulla privacy dei dati. Per informazioni sulla protezione dei dati in Europa, consulta il post del blog relativo al AWS Modello di responsabilità condivisa e GDPR nel AWS Blog sulla sicurezza.

Ai fini della protezione dei dati, consigliamo di proteggere Account AWS le credenziali e configurare i singoli utenti con AWS IAM Identity Center or AWS Identity and Access Management (IAM). In tal modo, a ogni utente verranno assegnate solo le autorizzazioni necessarie per svolgere i suoi compiti. Suggeriamo, inoltre, di proteggere i dati nei seguenti modi:

- Utilizza l'autenticazione a più fattori (MFA) con ogni account.
- SSL/TLS Da utilizzare per comunicare con AWS le risorse. È richiesto TLS 1.2 ed è consigliato TLS 1.3.
- Configura l'API e la registrazione delle attività degli utenti con AWS CloudTrail. Per informazioni sull'utilizzo dei CloudTrail percorsi per acquisire AWS le attività, consulta Lavorare con i CloudTrail percorsi nella Guida per l'AWS CloudTrail utente.
- Utilizza soluzioni di AWS crittografia, insieme a tutti i controlli di sicurezza predefiniti all'interno Servizi AWS.
- Utilizza i servizi di sicurezza gestiti avanzati, come Amazon Macie, che aiutano a individuare e proteggere i dati sensibili archiviati in Amazon S3.
- Se hai bisogno di moduli crittografici convalidati FIPS 140-3 per accedere AWS tramite un'interfaccia a riga di comando o un'API, usa un endpoint FIPS. Per ulteriori informazioni sugli endpoint FIPS disponibili, consulta il Federal Information Processing Standard (FIPS) 140-3.

Ti consigliamo di non inserire mai informazioni riservate o sensibili, ad esempio gli indirizzi e-mail dei clienti, nei tag o nei campi di testo in formato libero, ad esempio nel campo Nome. Ciò include

quando si lavora con OpenSearch Service o altro Servizi AWS utilizzando la console, l'API o. AWS CLI AWS SDKs I dati inseriti nei tag o nei campi di testo in formato libero utilizzati per i nomi possono essere utilizzati per i la fatturazione o i log di diagnostica. Quando si fornisce un URL a un server esterno, suggeriamo vivamente di non includere informazioni sulle credenziali nell'URL per convalidare la richiesta al server.

Crittografia dei dati inattivi per Amazon OpenSearch Service

OpenSearch I domini di servizio offrono la crittografia dei dati inattivi, una funzionalità di sicurezza che aiuta a prevenire l'accesso non autorizzato ai tuoi dati. La funzionalità utilizza AWS Key Management Service (AWS KMS) per archiviare e gestire le chiavi di crittografia e l'algoritmo Advanced Encryption Standard con chiavi a 256 bit (AES-256) per eseguire la crittografia. Se abilitata, la funzionalità crittografa gli elementi seguenti di un dominio:

- Tutti gli indici (compresi quelli archiviati) UltraWarm
- OpenSearch registri
- File di swap
- Tutti gli altri dati presenti nella directory dell'applicazione
- Snapshot automatici

Gli elementi seguenti non vengono crittografati quando abiliti la crittografia dei dati a riposo, ma puoi eseguire operazioni aggiuntive per proteggerli:

- Istantanee manuali: al momento non è possibile utilizzare AWS KMS le chiavi per crittografare le istantanee manuali. Tuttavia, è possibile usare la crittografia lato server con chiavi gestite da S3 o chiavi KMS per crittografare il bucket che si utilizza come repository degli snapshot. Per istruzioni, consultare the section called “Registrazione di un repository di snapshot manuali”.
- Slow log e log degli errori: se pubblichi log e desideri crittografarli, puoi crittografare il relativo gruppo di log CloudWatch Logs utilizzando la stessa chiave del dominio del servizio. AWS KMS OpenSearch Per ulteriori informazioni, consulta Encrypt log data in CloudWatch Logs using AWS Key Management Service nella Amazon CloudWatch Logs User Guide.

> **Note**
>
> Non puoi abilitare la crittografia a riposo su un dominio esistente se UltraWarm sul dominio è abilitata la memorizzazione a freddo. È necessario innanzitutto UltraWarm disabilitare

la memorizzazione a freddo, abilitare la crittografia a riposo e quindi riattivare UltraWarm la conservazione a freddo. Se si desidera conservare gli indici in una conservazione a caldo UltraWarm o a freddo, è necessario spostarli nella memorizzazione a caldo prima di UltraWarm disattivarli o archiviarli a freddo.

OpenSearch Il servizio supporta solo chiavi KMS con crittografia simmetrica, non asimmetriche. Per informazioni su come creare chiavi simmetriche, consulta Creare una chiave KMS nella Guida per gli sviluppatori.AWS Key Management Service

Indipendentemente dal fatto che la crittografia a riposo sia abilitata, tutti i domini crittografano automaticamente i pacchetti personalizzati utilizzando AES-256 e chiavi gestite dal servizio. OpenSearch

Permissions

Per utilizzare la console di OpenSearch servizio per configurare la crittografia dei dati inattivi, è necessario disporre delle autorizzazioni di lettura per AWS KMS, ad esempio la seguente politica basata sull'identità:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:List*",
        "kms:Describe*"
      ],
      "Resource": "*"
    }
  ]
}
```

Se desideri utilizzare una chiave diversa da quella di AWS proprietà, devi disporre anche delle autorizzazioni per creare concessioni per la chiave. Queste autorizzazioni in genere hanno la forma di una policy basata su risorse specificata quando crei la chiave.

Se desideri mantenere la tua chiave esclusiva per OpenSearch Service, puoi aggiungere la ViaService condizione kms: a quella policy chiave:

```
"Condition": {
  "StringEquals": {
    "kms:ViaService": "es.us-west-1.amazonaws.com"
  },
  "Bool": {
    "kms:GrantIsForAWSResource": "true"
  }
}
```

Per ulteriori informazioni, consulta Policy delle chiavi in AWS KMS nella Guida per gli sviluppatori di AWS Key Management Service .

Abilitazione della crittografia dei dati a riposo

La crittografia dei dati archiviati su nuovi domini richiede Elasticsearch OpenSearch 5.1 o versione successiva. Per abilitarla su domini esistenti è necessario Elasticsearch 6.7 OpenSearch o versione successiva.

Come abilitare la crittografia dei dati a riposo (console)

1. Apri il dominio nella AWS console, quindi scegli Azioni e Modifica configurazione di sicurezza.
2. In Encryption (Crittografia), seleziona Enable encryption of data at rest (Abilita la crittografia dei dati a riposo).
3. Scegli una AWS KMS chiave da usare, quindi scegli Salva modifiche.

Puoi abilitare la crittografia anche tramite l'API di configurazione. La seguente richiesta abilita la crittografia dei dati a riposo su un dominio esistente:

```
{
   "ClusterConfig":{
      "EncryptionAtRestOptions":{
         "Enabled": true,
         "KmsKeyId":"arn:aws:kms:us-east-1:123456789012:alias/my-key"
      }
   }
}
```

Chiave KMS disabilitata o eliminata

Se disabiliti o elimini la chiave che hai usato per crittografare un dominio, il dominio diventa inaccessibile. OpenSearch Il servizio ti invia una notifica che ti informa che non può accedere alla chiave KMS. Riabilita immediatamente la chiave per accedere al dominio.

Il team OpenSearch di assistenza non può aiutarti a recuperare i dati se la chiave viene eliminata. AWS KMS elimina le chiavi solo dopo un periodo di attesa di almeno sette giorni. Se la tua chiave è in attesa di cancellazione, annulla la cancellazione o prendi uno Snapshot manuale del dominio per evitare la perdita di dati.

Disabilitazione della crittografia dei dati a riposo

Dopo aver configurato un dominio per crittografare i dati a riposo, non puoi disabilitare l'impostazione. Puoi invece acquisire una snapshot manuale del dominio esistente, creare un altro dominio, migrare i dati ed eliminare il dominio precedente.

Monitoraggio dei domini che crittografano dati a riposo

I domini che crittografano i dati a riposo hanno due parametri aggiuntivi: `KMSKeyError` e `KMSKeyInaccessible`. Questi parametri vengono visualizzati solo se il dominio rileva un problema con la chiave di crittografia. Per una descrizione completa di questi parametri, consulta the section called “Parametri cluster”. Puoi visualizzarli utilizzando la console OpenSearch di servizio o la CloudWatch console Amazon.

> **Tip**
>
> Ogni metrica rappresenta un problema significativo per un dominio, quindi ti consigliamo di creare CloudWatch allarmi per entrambi. Per ulteriori informazioni, consulta the section called “Allarmi consigliati CloudWatch ”.

Altre considerazioni

- La rotazione automatica delle chiavi preserva le proprietà delle AWS KMS chiavi, quindi la rotazione non ha alcun effetto sulla capacità di accedere ai dati. OpenSearch I domini OpenSearch di Encrypted Service non supportano la rotazione manuale delle chiavi, che comporta la creazione di una nuova chiave e l'aggiornamento di eventuali riferimenti alla vecchia chiave. Per ulteriori informazioni, consulta Rotate AWS KMS keys nella AWS Key Management Service Developer Guide.

- Alcuni tipi di istanza non supportano la crittografia dei dati a riposo. Per dettagli, consulta the section called “Tipi di istanze supportati”.
- I domini che crittografano i dati a riposo usano un nome di repository diverso per i propri snapshot automatici. Per ulteriori informazioni, consultare the section called “Ripristino di snapshot”.
- Anche se consigliamo vivamente di abilitare la crittografia dei dati a riposo, questa potrebbe aggiungere un sovraccarico aggiuntivo della CPU e alcuni millisecondi di latenza. Tuttavia, la maggior parte dei casi d'uso non è sensibile a queste differenze e l'entità dell'impatto dipende dalla configurazione del cluster, dei client e del profilo di utilizzo.

Node-to-node crittografia per Amazon OpenSearch Service

Node-to-node la crittografia fornisce un ulteriore livello di sicurezza oltre alle funzionalità predefinite di Amazon OpenSearch Service.

Ogni dominio OpenSearch di servizio, indipendentemente dal fatto che utilizzi l'accesso VPC, risiede all'interno del proprio VPC dedicato. Questa architettura impedisce ai potenziali aggressori di intercettare il traffico tra i nodi e protegge il cluster. OpenSearch Per impostazione predefinita, tuttavia, il traffico all'interno del VPC non è crittografato. Node-to-nodela crittografia abilita la crittografia TLS 1.2 per tutte le comunicazioni all'interno del VPC.

Se invii dati al OpenSearch Servizio tramite HTTPS, la node-to-node crittografia aiuta a garantire che i dati rimangano crittografati durante la OpenSearch distribuzione (e la ridistribuzione) in tutto il cluster. Se i dati arrivano non crittografati tramite HTTP, OpenSearch Service li crittografa dopo aver raggiunto il cluster. Puoi richiedere che tutto il traffico verso il dominio arrivi tramite HTTPS utilizzando la console o l'API AWS CLI di configurazione.

Node-to-node la crittografia è necessaria se si abilita il controllo granulare degli accessi.

Attivazione della crittografia node-to-node

Node-to-node la crittografia su nuovi domini richiede qualsiasi versione di OpenSearch Elasticsearch 6.0 o successiva. L'abilitazione della node-to-node crittografia su domini esistenti richiede qualsiasi versione di Elasticsearch 6.7 o OpenSearch successiva. Scegli il dominio esistente nella console AWS , seleziona Operazioni, quindi Modifica configurazione di sicurezza.

In alternativa, puoi utilizzare l'API o di configurazione. AWS CLI Per ulteriori informazioni, consulta il riferimento alle API AWS CLI Command Reference e OpenSearch Service.

Disabilitazione della crittografia node-to-node

Dopo aver configurato un dominio per utilizzare node-to-node la crittografia, non puoi disabilitare l'impostazione. Puoi invece acquisire una snapshot manuale del dominio crittografato, creare un altro dominio, migrare i dati ed eliminare il dominio precedente.

Identity and Access Management in Amazon OpenSearch Service

Amazon OpenSearch Service offre diversi modi per controllare l'accesso ai tuoi domini. In questa sezione sono descritti i diversi tipi di policy, il modo in cui interagiscono tra loro ed è riportato come creare le policy personalizzate.

> **Important**
>
> Il supporto VPC introduce alcune considerazioni aggiuntive sul controllo degli accessi ai OpenSearch servizi. Per ulteriori informazioni, consulta the section called “Informazioni sulle policy d'accesso nei domini VPC”.

Tipi di policy

OpenSearch Il servizio supporta tre tipi di politiche di accesso:

- the section called “Policy basate sulle risorse”
- the section called “Policy basate sull’identità”
- the section called “Policy basate su IP”

Policy basate sulle risorse

Quando si crea un dominio, viene aggiunta una policy basata su risorse, talvolta denominata policy di accesso al dominio. Queste policy specificano le operazioni che un principale può eseguire sulle risorse secondarie del dominio (con l'eccezione della ricerca tra cluster). Le sottorisorse includono OpenSearch indici e. APIs L'elemento della policy JSON Principal in IAM specifica gli account, gli utenti o i ruoli a cui è consentito l'accesso. L'elemento della policy Resource JSON specifica a quali sottorisorse possono accedere questi principali.

Ad esempio, la seguente policy basata sulle risorse concede un accesso completo `test-user` (`es:*`) alle risorse secondarie in `test-domain`:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "arn:aws:iam::123456789012:user/test-user"
        ]
      },
      "Action": [
        "es:*"
      ],
      "Resource": "arn:aws:es:us-west-1:987654321098:domain/test-domain/*"
    }
  ]
}
```

Due considerazioni importanti si applicano a questa policy:

- Questi privilegi si applicano solo a questo dominio. A meno che non vengano create policy simili su altri domini, `test-user` può accedere solo a `test-domain`.
- I caratteri `/*` finali nell'elemento `Resource` sono significativi e indicano che le policy basate sulle risorse si applicano solo alle risorse secondarie del dominio e non al dominio stesso. Nelle policy basate sulle risorse, l'operazione `es:*` equivale a `es:ESHttp*`.

  Ad esempio, `test-user` può effettuare richieste su un indice (`GET https://search-test-domain.us-west-1.es.amazonaws.com/test-index`), ma non è in grado di aggiornare la configurazione del dominio (`POST https://es.us-west-1.amazonaws.com/2021-01-01/opensearch/domain/test-domain/config`). Nota la differenza tra i due endpoint. L'accesso all'API di configurazione richiede una politica basata sull'identità.

È possibile specificare un nome di indice parziale aggiungendo un carattere jolly. Questo esempio identifica gli indici che iniziano con `commerce`:

```
arn:aws:es:us-west-1:987654321098:domain/test-domain/commerce*
```

In questo caso, l'uso del carattere jolly significa che `test-user` può fare richieste agli indici nel dominio `test-domain` che hanno nomi che iniziano con `commerce`.

Per limitare ulteriormente `test-user`, è possibile applicare la seguente policy:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "arn:aws:iam::123456789012:user/test-user"
        ]
      },
      "Action": [
        "es:ESHttpGet"
      ],
      "Resource": "arn:aws:es:us-west-1:987654321098:domain/test-domain/commerce-
data/_search"
    }
  ]
}
```

Ora `test-user` è in grado di eseguire un'unica operazione: effettua una ricerca sull'indice `commerce-data`. Tutti gli altri indici all'interno del dominio sono inaccessibili e senza le autorizzazioni necessarie per utilizzare le operazioni `es:ESHttpPut` o `es:ESHttpPost`, `test-user` non è in grado di aggiungere o modificare i documenti.

Quindi, è possibile decidere di configurare un ruolo per gli utenti avanzati. Questa politica consente `power-user-role` l'accesso ai metodi HTTP GET e PUT per tutti gli URIs elementi dell'indice:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
```

```
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "arn:aws:iam::123456789012:role/power-user-role"
        ]
      },
      "Action": [
        "es:ESHttpGet",
        "es:ESHttpPut"
      ],
      "Resource": "arn:aws:es:us-west-1:987654321098:domain/test-domain/commerce-
data/*"
    }
  ]
}
```

Se il dominio si trova in un VPC o utilizza un controllo granulare degli accessi, è possibile usare una policy di accesso al dominio aperto. In caso contrario, la policy di accesso al dominio deve contenere alcune limitazioni, sia per principal che per indirizzo IP.

Per ulteriori informazioni su tutte le azioni disponibili, consultare the section called “Riferimenti agli elementi della policy”. Per un controllo molto più dettagliato sui dati, utilizzare una policy di accesso al dominio aperto con il controllo granulare degli accessi.

Policy basate sull’identità

A differenza delle policy basate sulle risorse, che fanno parte di ogni dominio di OpenSearch servizio, le policy basate sull'identità vengono associate a utenti o ruoli utilizzando il servizio (IAM). AWS Identity and Access Management Proprio come le policy basate sulle risorse, le policy basate su identità specificano chi può accedere a un servizio, quali azioni può eseguire e, ove applicabile, le risorse su cui può eseguire tali operazioni.

Mentre le policy basate su identità tendono a essere più generiche. Spesso regolamentano solo le operazioni delle API di configurazione che possono essere eseguite da un utente. Dopo aver implementato queste politiche, è possibile utilizzare le politiche basate sulle risorse (o il controllo granulare degli accessi) in Service per offrire agli utenti l'accesso agli indici e. OpenSearch OpenSearch APIs

> **Note**
>
> Gli utenti con la `AmazonOpenSearchServiceReadOnlyAccess` policy AWS gestita non possono visualizzare lo stato di integrità del cluster sulla console. Per consentire loro di visualizzare lo stato di integrità del cluster (e altri OpenSearch dati), aggiungi l'`es:ESHttpGet`azione a una politica di accesso e allegala ai loro account o ruoli.

Poiché le policy basate su identità si collegano a utenti o ruoli (principali), il formato JSON non specifica un principale. La policy seguente consente di concedere l'accesso alle azioni che iniziano con `Describe` e `List`. Questa combinazione di azioni fornisce accesso in sola lettura alle configurazioni di dominio, ma non ai dati archiviati nel dominio:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Action": [
        "es:Describe*",
        "es:List*"
      ],
      "Effect": "Allow",
      "Resource": "*"
    }
  ]
}
```

Un amministratore potrebbe avere pieno accesso al OpenSearch servizio e a tutti i dati archiviati su tutti i domini:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
```

```
      "Action": [
        "es:*"
      ],
      "Effect": "Allow",
      "Resource": "*"
    }
  ]
}
```

Le policy basate sull'identità consentono di utilizzare i tag per controllare l'accesso all'API di configurazione. La policy riportata di seguito, ad esempio, consente ai principal collegati di visualizzare e aggiornare la configurazione di un dominio se il dominio dispone del tag `team:devops`:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Action": [
      "es:UpdateDomainConfig",
      "es:DescribeDomain",
      "es:DescribeDomainConfig"
    ],
    "Effect": "Allow",
    "Resource": "*",
    "Condition": {
      "ForAnyValue:StringEquals": {
        "aws:ResourceTag/team": [
          "devops"
        ]
      }
    }
  }]
}
```

Puoi anche utilizzare i tag per controllare l'accesso all' OpenSearch API. Le politiche basate su tag per l' OpenSearch API si applicano solo ai metodi HTTP. Ad esempio, la seguente politica

consente ai principali collegati di inviare richieste GET e PUT all' OpenSearch API se il dominio ha il `environment:production` tag:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Action": [
      "es:ESHttpGet",
      "es:ESHttpPut"
    ],
    "Effect": "Allow",
    "Resource": "*",
    "Condition": {
      "ForAnyValue:StringEquals": {
        "aws:ResourceTag/environment": [
          "production"
        ]
      }
    }
  }]
}
```

Per un controllo più granulare dell' OpenSearch API, prendi in considerazione l'utilizzo di un controllo granulare degli accessi.

> **Note**
>
> Dopo averne aggiunto uno o più OpenSearch APIs a qualsiasi politica basata su tag, devi eseguire un'unica operazione sui tag (ad esempio aggiungere, rimuovere o modificare un tag) affinché le modifiche abbiano effetto su un dominio. È necessario utilizzare il software di servizio R20211203 o versione successiva per includere le operazioni OpenSearch API nelle politiche basate su tag.

OpenSearch Il servizio supporta le chiavi `RequestTag` di condizione `TagKeys` globali per l'API di configurazione, non l'API. OpenSearch Queste condizioni si applicano solo alle chiamate API che includono tag all'interno della richiesta, ad esempio `CreateDomain`, `AddTags` e `RemoveTags`. La

policy seguente consente ai principal collegati di creare domini, ma solo se includono il tag `team:it` nella richiesta:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": {
    "Effect": "Allow",
    "Action": [
      "es:CreateDomain",
      "es:AddTags"
    ],
    "Resource": "*",
    "Condition": {
      "StringEquals": {
        "aws:RequestTag/team": [
          "it"
        ]
      }
    }
  }
}
```

Per ulteriori informazioni sull'utilizzo dei tag per il controllo degli accessi e sulle differenze tra le politiche basate sulle risorse e quelle basate sull'identità, consulta la sezione Definizione delle autorizzazioni in base agli attributi con autorizzazione ABAC nella IAM User Guide.

Policy basate su IP

Le policy basate su IP limitano l'accesso a un dominio a uno o più indirizzi IP o blocchi CIDR. Tecnicamente, le policy basate su IP non sono un tipo distinto di policy. Sono invece solo politiche basate sulle risorse che specificano un principale anonimo e includono una condizione speciale. Per informazioni, consulta IAM JSON Policy elements: Condition nella IAM User Guide.

L'attrattiva principale delle policy basate su IP è che consentono richieste non firmate a un dominio di OpenSearch servizio, il che consente di utilizzare client come curl e OpenSearch Dashboards o accedere al dominio tramite un server proxy. Per ulteriori informazioni, consultare the section called “Utilizzo di un proxy per accedere al servizio da dashboard OpenSearch”.

> **Note**
>
> Se è stato abilitato l'accesso VPC al dominio, non è possibile configurare una policy basata su IP. Puoi invece utilizzare `security groups` per controllare quali indirizzi IP possono accedere al dominio. Per ulteriori informazioni, consulta i seguenti argomenti:
>
> - the section called “Informazioni sulle policy d'accesso nei domini VPC”
> - Controlla il traffico verso AWS le tue risorse utilizzando i gruppi di sicurezza nella Amazon VPC User Guide

La seguente policy concede a tutte le richieste HTTP che provengono dall'intervallo di IP specificato l'accesso a `test-domain`:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
      "Action": [
        "es:ESHttp*"
      ],
      "Condition": {
        "IpAddress": {
          "aws:SourceIp": [
            "192.0.2.0/24"
          ]
        }
      },
      "Resource": "arn:aws:es:us-west-1:987654321098:domain/test-domain/*"
    }
  ]
}
```

Se il dominio dispone di un endpoint pubblico e non utilizza il controllo granulare degli accessi, è consigliabile combinare le entità IAM e gli indirizzi IP. Questa policy concede l'accesso HTTP `test-user` solo se la richiesta proviene dall'intervallo IP specificato:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "AWS": [
        "arn:aws:iam::987654321098:user/test-user"
      ]
    },
    "Action": [
      "es:ESHttp*"
    ],
    "Condition": {
      "IpAddress": {
        "aws:SourceIp": [
          "192.0.2.0/24"
        ]
      }
    },
    "Resource": "arn:aws:es:us-west-1:987654321098:domain/test-domain/*"
  }]
}
```

Effettuazione e firma di richieste OpenSearch di assistenza

Anche se configuri una politica di accesso completamente aperta basata su risorse, tutte le richieste all'API di configurazione del OpenSearch servizio devono essere firmate. Se le tue policy specificano ruoli o utenti IAM, OpenSearch APIs anche le richieste devono essere firmate utilizzando AWS Signature Version 4. Il metodo della firma differisce in base alle API:

- Per effettuare chiamate all'API OpenSearch di configurazione del servizio, ti consigliamo di utilizzare una delle AWS SDKs. Semplifica SDKs notevolmente il processo e può farti risparmiare una notevole quantità di tempo rispetto alla creazione e alla firma delle tue richieste. Gli endpoint dell'API di configurazione utilizzano il seguente formato:

```
es.region.amazonaws.com/2021-01-01/
```

Ad esempio, la seguente richiesta consente di apportare una modifica di configurazione al dominio `movies`, ma l'utente deve firmarla manualmente (scelta non consigliata):

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/movies/config
{
  "ClusterConfig": {
    "InstanceType": "c5.xlarge.search"
  }
}
```

Se utilizzi uno di questi SDKs, come Boto 3, l'SDK gestisce automaticamente la firma della richiesta:

```
import boto3

client = boto3.client(es)
response = client.update_domain_config(
  DomainName='movies',
  ClusterConfig={
    'InstanceType': 'c5.xlarge.search'
  }
)
```

Per un esempio di codice Java, consulta the section called “Usando il AWS SDKs”.

- Per effettuare chiamate verso OpenSearch APIs, è necessario firmare le proprie richieste. Quindi OpenSearch APIs usa il seguente formato:

```
domain-id.region.es.amazonaws.com
```

Ad esempio, la seguente richiesta esegue una ricerca nell'indice `movies` per thor:

```
GET https://my-domain.us-east-1.es.amazonaws.com/movies/_search?q=thor
```

> **Note**
>
> Il servizio ignora i parametri passati URLs per le richieste HTTP POST firmate con Signature Version 4.

Quando le policy entrano in collisione

Sorgono problemi quando le policy dissentono o non effettuano alcuna menzione esplicita di un utente. Il funzionamento di IAM nella IAM User Guide fornisce un riepilogo conciso della logica di valutazione delle policy:

- Come impostazione predefinita, tutte le richieste vengono negate.
- Un permesso esplicito sostituisce questa impostazione di default.
- Un rifiuto esplicito sovrascrive tutti i consensi.

Ad esempio, se una policy basata sulle risorse ti concede l'accesso a una sottorisorsa di dominio (un OpenSearch indice o un'API), ma una policy basata sull'identità ti nega l'accesso, ti viene negato l'accesso. Se una policy basata sull'identità consente l'accesso e una policy basata sulle risorse non specifica se si dispone o meno dell'accesso, è consentito l'accesso. Consultare la seguente tabella di policy che si sovrappongono per un riepilogo dei risultati per le risorse secondarie del dominio.

	Consentito nella policy basata sulle risorse	Rifiutato nella policy basata sulle risorse	Non consentito né rifiutato nella policy basata sulle risorse
Allowed in identity-based policy	Consenso	Rifiuta	Consenso
Denied in identity-based policy	Rifiuta	Rifiuta	Rifiuta
Neither allowed nor denied in identity-based policy	Consenso	Rifiuta	Rifiuta

Riferimenti agli elementi della policy

OpenSearch Il servizio supporta la maggior parte degli elementi delle policy nello IAM Policy Elements Reference, ad eccezione di. `NotPrincipal` La tabella riportata di seguito mostra gli elementi più comuni.

Elemento della policy JSON	Riepilogo
`Version`	La versione corrente del linguaggio della policy è `2012-10-17` . Tutte le policy d'accesso devono specificare questo valore.
`Effect`	L'elemento specifica se l'istruzione consente o nega l'accesso alle azioni specificate. I valori validi sono `Allow` e `Deny`.
`Principal`	Questo elemento specifica il ruolo Account AWS o l'utente IAM a cui è consentito o negato l'accesso a una risorsa e può assumere diverse forme: • AWS account: `"Principal":{"AWS": ["123456789012"]}` o `"Principal":{"AWS": ["arn:aws:iam::123456789012:root"]}` • Utenti IAM: `"Principal":{"AWS": ["arn:aws:iam::123456789012:user/test-user"]}` • Ruoli IAM: `"Principal":{"AWS": ["arn:aws:iam::123456789012:role/test-role"]}` **Important** La specifica del carattere jolly * consente l'accesso anonimo al dominio, che non è consigliabile a meno che non si aggiunga una condizione basata su IP, non si usi il supporto VPC o non si abiliti il controllo granulare degli accessi. Inoltre, esamina attentamente le seguenti politiche per confermare che non garantiscano un ampio accesso:

Elemento della policy JSON	Riepilogo
	• Policy basate sull'identità collegate ai principali AWS associati (ad esempio, ruoli IAM) • Politiche basate sulle risorse allegate alle AWS risorse associate (ad esempio, chiavi KMS) AWS Key Management Service

Elemento della policy JSON	Riepilogo
`Action`	OpenSearch Il servizio utilizza `ESHttp*` azioni per i metodi HTTP. OpenSearch Il resto delle operazioni si applica all'API di configurazione. Alcune azioni `es:` supportano le autorizzazioni a livello di risorsa. Ad esempio, è possibile assegnare a un utente autorizzazioni per eliminare un determinato dominio senza garantirgli le autorizzazioni per eliminare qualsiasi dominio. Altre azioni si applicano solo al servizio. Ad esempio, `es:ListDomainNames` non ha senso nel contesto di un singolo dominio e quindi richiede un carattere jolly. Per un elenco di tutte le azioni disponibili e se si applicano alle sottoriso rse del dominio (`test-domain/*`), alla configurazione del dominio (`test-domain`) o solo al servizio (`*`), consulta Azioni, risorse e chiavi di condizione per Amazon OpenSearch Service nel Service Authorization Reference Le policy basate sulle risorse differiscono dalle autorizzazioni a livello di risorsa. Le policy basate sulle risorse sono policy JSON complete che si collegano ai domini. Le autorizzazioni a livello di risorsa consentono di limitare le azioni a particolari domini o risorse secondarie. In pratica, si possono considerare le autorizzazioni a livello di risorsa una parte opzionale di una policy basata sulle risorse o sull'identità. Mentre le autorizzazioni a livello di risorsa per `es:CreateDomain` potrebbero sembrare poco intuitive (dopo tutto, perché offrire a un utente le autorizzazioni per creare un dominio già esistente?) l'uso di un carattere jolly consente di implementare uno schema di denominaz ione semplice per i domini, ad esempio `"Resource": "arn:aws:es:us-west-1:987654321098:domain/my-team-name-*"`. Naturalmente, è ugualmente possibile includere azioni insieme a elementi di risorse meno restrittivi, come i seguenti:

Elemento della policy JSON	Riepilogo
	JSON `{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Action": [ "es:ESHttpGet", "es:DescribeDomain" ], "Resource": "*" } ] }` Per ulteriori informazioni sull'accoppiamento di azioni e risorse, fare riferimento all'elemento `Resource` in questa tabella.

Elemento della policy JSON	Riepilogo
`Condition`	OpenSearch Il servizio supporta la maggior parte delle condizioni descritte nelle chiavi di contesto delle condizioni AWS globali nella IAM User Guide. Le eccezioni più importanti includono la `aws:PrincipalTag` chiave, che OpenSearch Service non supporta. Durante la configurazione di una policy basata su IP, è necessario specificare gli indirizzi IP o i blocchi CIDR come condizione, come ad esempio: `"Condition": { "IpAddress": { "aws:SourceIp": [ "192.0.2.0/32" ] } }` Come indicato inthe section called “Policy basate sull’identità”, i tasti `aws:ResourceTag` `aws:RequestTag` , e `aws:TagKeys` condition si applicano all'API di configurazione e al. OpenSearch APIs

Elemento della policy JSON	Riepilogo

`Resource`	OpenSearch Il servizio utilizza `Resource` gli elementi in tre modi fondamentali:
	
	• Per azioni che si applicano al OpenSearch Servizio stesso`es:ListDomainNames` , come o per consentire l'accesso completo, usa la seguente sintassi:
	
	`"Resource": "*"`
	
	• Per le azioni che implicano una configurazione del dominio, ad esempio `es:DescribeDomain` , è possibile utilizzare la sintassi seguente:
	
	`"Resource": "arn:aws:es:` *`region`*`:`*`aws-account-id`*`:domain/`*`domain-name`*` "`
	
	• Per le azioni che si applicano alle risorse secondarie di un dominio, ad esempio `es:ESHttpGet` , è possibile utilizzare la sintassi seguente:
	
	`"Resource": "arn:aws:es:` *`region`*`:`*`aws-account-id`*`:domain/`*`domain-name`*` /*"`
	
	Non è necessario utilizzare un jolly. OpenSearch Il servizio consente di definire una politica di accesso diversa per ogni OpenSearch indice o API. Ad esempio, è possibile limitare le autorizzazioni di un utente per l'indice `test-index` :
	
	`"Resource": "arn:aws:es:` *`region`*`:`*`aws-account-id`*`:domain/`*`domain-name`*` /test-index"`
	
	Invece dell'accesso completo a `test-index` , è preferibile limitare la policy all'API di ricerca:
	
	`"Resource": "arn:aws:es:` *`region`*`:`*`aws-account-id`*`:domain/`*`domain-name`*` /test-index/_search"`

Elemento della policy JSON	Riepilogo
	È possibile anche controllare l'accesso ai singoli documenti: `"Resource": "arn:aws:es:  region:aws-account-id:domain/domain-name /test-index/test-type/1"` In sostanza, se OpenSearch esprime la sottorisorsa come URI, è possibile controllarne l'accesso utilizzando una politica di accesso. Per un controllo ancora maggiore sulle risorse a cui un utente può accedere, vedere the section called “Controllo granulare degli accessi”. Per ulteriori informazioni su quali azioni supportano le autorizzazioni a livello di risorsa, fare riferimento all'elemento `Action` in questa tabella.

Opzioni avanzate e considerazioni sulle API

OpenSearch Il servizio ha diverse opzioni avanzate, una delle quali ha implicazioni sul controllo degli accessi:. `rest.action.multi.allow_explicit_index` Con l'impostazione predefinita true, consente agli utenti di ignorare le autorizzazioni a livello di risorsa secondaria in determinate circostanze.

Ad esempio, considerare la seguente policy basata sulle risorse:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "arn:aws:iam::123456789012:user/test-user"
        ]
      },
```

```
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": [
        "arn:aws:es:us-west-1:987654321098:domain/test-domain/test-index/*",
        "arn:aws:es:us-west-1:987654321098:domain/test-domain/_bulk"
      ]
    },
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "arn:aws:iam::123456789012:user/test-user"
        ]
      },
      "Action": [
        "es:ESHttpGet"
      ],
      "Resource": "arn:aws:es:us-west-1:987654321098:domain/test-domain/
restricted-index/*"
    }
  ]
}
```

Questa politica garantisce l'accesso `test-user` completo `test-index` e alla OpenSearch massa delle API. Consente inoltre `GET` le richieste a `restricted-index`.

La seguente richiesta di indicizzazione, come è prevedibile, ha esito negativo a causa di un errore delle autorizzazioni:

```
PUT https://search-test-domain.us-west-1.es.amazonaws.com/restricted-index/movie/1
{
  "title": "Your Name",
  "director": "Makoto Shinkai",
  "year": "2016"
}
```

A differenza dell'API dell'indice, l'API bulk consente la creazione, l'aggiornamento e l'eliminazione di molti documenti in una sola chiamata. Tuttavia spesso si specificano queste operazioni nel corpo della richiesta, anziché nell'URL della richiesta. Poiché OpenSearch Service utilizza URLs per controllare l'accesso alle sottorisorse del dominio, `test-user` può, in effetti, utilizzare l'API in blocco

per apportare modifiche a. `restricted-index` Anche se l'utente non dispone di autorizzazioni POST per l'indice, la seguente richiesta ha esito positivo:

```
POST https://search-test-domain.us-west-1.es.amazonaws.com/_bulk
{ "index" : { "_index": "restricted-index", "_type" : "movie", "_id" : "1" } }
{ "title": "Your Name", "director": "Makoto Shinkai", "year": "2016" }
```

In questo caso, la policy d'accesso non riesce a soddisfare i suoi intenti. Per impedire agli utenti di aggirare questi tipi di restrizioni, è possibile modificare `rest.action.multi.allow_explicit_index` in false. Se questo valore è falso, tutte le chiamate a bulk, mget e msearch APIs che specificano i nomi degli indici nel corpo della richiesta smettono di funzionare. In altre parole, le chiamate a `_bulk` non funzionano, ma le chiamate a `test-index/_bulk` sì. Questo secondo endpoint contiene un nome dell'indice, perciò non è necessario specificarne uno nel corpo nella richiesta.

OpenSearch Le dashboard si basano molto su mget e msearch, quindi è improbabile che funzionino correttamente dopo questa modifica. Per una correzione parziale, puoi lasciare impostato `rest.action.multi.allow_explicit_index` come true e negare a determinati utenti l'accesso a uno o più di questi. APIs

Per informazioni su come modificare questa impostazione, consultare the section called “Impostazioni avanzate del cluster”.

Analogamente, la seguente policy basata sulle risorse contiene due problemi di lieve entità:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::123456789012:user/test-user"
      },
      "Action": "es:ESHttp*",
      "Resource": "arn:aws:es:us-west-1:987654321098:domain/test-domain/*"
    },
    {
      "Effect": "Deny",
```

```
      "Principal": {
        "AWS": "arn:aws:iam::123456789012:user/test-user"
      },
      "Action": "es:ESHttp*",
      "Resource": "arn:aws:es:us-west-1:987654321098:domain/test-domain/
restricted-index/*"
    }
  ]
}
```

- Nonostante il rifiuto esplicito, `test-user` può comunque effettuare chiamate come `GET https://search-test-domain.us-west-1.es.amazonaws.com/_all/_search` e `GET https://search-test-domain.us-west-1.es.amazonaws.com/*/_search` per accedere ai documenti in `restricted-index`.
- Poiché l'elemento `Resource` fa riferimento a `restricted-index/*`, `test-user` non dispone delle autorizzazioni per accedere direttamente ai documenti dell'indice. L'utente, tuttavia, dispone delle autorizzazioni necessarie per eliminare l'intero indice. Per prevenire l'accesso e l'eliminazione, la policy deve specificare `restricted-index*`.

Anziché combinare permessi ampi e negazioni strette, la soluzione più sicura è seguire il principio del privilegio minimo e concedere solo le autorizzazioni necessarie per eseguire un'operazione. Per ulteriori informazioni sul controllo dell'accesso a singoli indici o OpenSearch operazioni, vedere. the section called “Controllo granulare degli accessi”

> **Important**
>
> Specificando il carattere jolly * si abilita l'accesso anonimo al dominio. Non è consigliabile utilizzare il carattere jolly. Inoltre, esamina attentamente le seguenti politiche per verificare che non garantiscano un accesso ampio:
>
> - Politiche basate sull'identità collegate ai AWS principali associati (ad esempio, ruoli IAM)
> - Politiche basate sulle risorse collegate alle AWS risorse associate (ad esempio, chiavi KMS) AWS Key Management Service

Configurazione delle policy di accesso

- Per istruzioni sulla creazione o la modifica di politiche basate su risorse e IP in Service, vedere. OpenSearch the section called “Configurazione delle policy di accesso”
- Per istruzioni su come creare o modificare le policy basate sull'identità in IAM, consulta Definizione delle autorizzazioni IAM personalizzate con le policy gestite dal cliente nella Guida per l'utente IAM.

Altre policy di esempio

Sebbene questo capitolo includa molti esempi di policy, il controllo degli AWS accessi è un argomento complesso che può essere compreso meglio attraverso esempi. Per ulteriori informazioni, consultare Esempi di policy IAM basate su identità nella Guida per l'utente di IAM.

Riferimento alle autorizzazioni dell'API Amazon OpenSearch Service

Quando configuri il controllo degli accessi, scrivi politiche di autorizzazione che puoi allegare a un'identità IAM (politiche basate sull'identità). Per ulteriori informazioni, consultare gli argomenti seguenti nella Referenza sull'autorizzazione del servizio:

- Azioni, risorse e chiavi di condizione per Service. OpenSearch
- Azioni, risorse e chiavi di condizione per OpenSearch Ingestion.

Questo riferimento contiene informazioni su quali operazioni API possono essere utilizzate in una policy IAM. Include anche la AWS risorsa per la quale è possibile concedere le autorizzazioni e le chiavi di condizione che è possibile includere per un controllo granulare degli accessi.

Le operazioni, il valore della risorsa e le condizioni vengono specificati rispettivamente nei campi `Action`, `Resource` e `Condition` della policy. Per specificare un'azione per OpenSearch Service, utilizzate il `es:` prefisso seguito dal nome dell'operazione API (ad esempio,). `es:CreateDomain` Per specificare un'azione per OpenSearch Ingestion, utilizzate il `osis:` prefisso seguito dall'operazione API (ad esempio,). `osis:CreatePipeline`

AWS politiche gestite per Amazon OpenSearch Service

Una politica AWS gestita è una politica autonoma creata e amministrata da. AWS AWS le politiche gestite sono progettate per fornire autorizzazioni per molti casi d'uso comuni, in modo da poter iniziare ad assegnare autorizzazioni a utenti, gruppi e ruoli.

Tieni presente che le policy AWS gestite potrebbero non concedere le autorizzazioni con il privilegio minimo per i tuoi casi d'uso specifici, poiché sono disponibili per tutti i clienti. AWS Si consiglia pertanto di ridurre ulteriormente le autorizzazioni definendo policy gestite dal cliente specifiche per i propri casi d'uso.

Non è possibile modificare le autorizzazioni definite nelle politiche gestite. AWS Se AWS aggiorna le autorizzazioni definite in una politica AWS gestita, l'aggiornamento ha effetto su tutte le identità principali (utenti, gruppi e ruoli) a cui è associata la politica. AWS è più probabile che aggiorni una policy AWS gestita quando ne Servizio AWS viene lanciata una nuova o quando diventano disponibili nuove operazioni API per i servizi esistenti.

Per ulteriori informazioni, consultare Policy gestite da AWS nella Guida per l'utente di IAM.

AmazonOpenSearchDirectQueryGlueCreateAccess

Concede ad Amazon OpenSearch Service Direct Query Service l'accesso a `CreateDatabase`,`CreatePartition`,`CreateTable`, e `BatchCreatePartition` AWS Glue API.

Puoi trovare la AmazonOpenSearchDirectQueryGlueCreateAccesspolicy nella console IAM.

AmazonOpenSearchServiceFullAccess

Garantisce l'accesso completo alle operazioni e alle risorse dell'API di configurazione del OpenSearch servizio per un Account AWS.

Puoi trovare la AmazonOpenSearchServiceFullAccesspolicy nella console IAM.

AmazonOpenSearchServiceReadOnlyAccess

Concede l'accesso in sola lettura a tutte le risorse del OpenSearch servizio per un. Account AWS

Puoi trovare la AmazonOpenSearchServiceReadOnlyAccesspolicy nella console IAM.

AmazonOpenSearchServiceRolePolicy

Non è possibile collegare `AmazonOpenSearchServiceRolePolicy`alle entità IAM. Questa policy è associata a un ruolo collegato al servizio che consente a OpenSearch Service di accedere alle risorse dell'account. Per ulteriori informazioni, consulta the section called “Permissions”.

Puoi trovare la AmazonOpenSearchServiceRolePolicypolicy nella console IAM.

AmazonOpenSearchServiceCognitoAccess

Fornisci le autorizzazioni minime di Amazon Cognito necessarie per abilitare l'autenticazione Cognito.

Puoi trovare la AmazonOpenSearchServiceCognitoAccesspolicy nella console IAM.

AmazonOpenSearchIngestionServiceRolePolicy

Non è possibile collegare `AmazonOpenSearchIngestionServiceRolePolicy`alle entità IAM. Questa policy è associata a un ruolo collegato al servizio che consente a OpenSearch Ingestion di abilitare l'accesso VPC per le pipeline di ingestione, creare tag e pubblicare metriche relative all'importazione sul tuo account. CloudWatch Per ulteriori informazioni, consulta the section called "Uso di ruoli collegati ai servizi".

AmazonOpenSearchIngestionServiceRolePolicyPuoi trovare la policy nella console IAM.

OpenSearchIngestionSelfManagedVpcePolicy

Non è possibile collegare `OpenSearchIngestionSelfManagedVpcePolicy`alle entità IAM. Questa policy è associata a un ruolo collegato al servizio che consente a OpenSearch Ingestion di abilitare l'accesso VPC autogestito per le pipeline di ingestione, creare tag e pubblicare metriche relative all'importazione sul tuo account. CloudWatch Per ulteriori informazioni, consulta the section called "Uso di ruoli collegati ai servizi".

OpenSearchIngestionSelfManagedVpcePolicyPuoi trovare la policy nella console IAM.

AmazonOpenSearchIngestionFullAccess

Garantisce l'accesso completo alle operazioni e alle risorse dell'API OpenSearch Ingestion per un. Account AWS

Puoi trovare la AmazonOpenSearchIngestionFullAccesspolicy nella console IAM.

AmazonOpenSearchIngestionReadOnlyAccess

Concede l'accesso in sola lettura a tutte le risorse di OpenSearch Ingestion per un. Account AWS

Puoi trovare la AmazonOpenSearchIngestionReadOnlyAccesspolicy nella console IAM.

AmazonOpenSearchServerlessServiceRolePolicy

Fornisce le Amazon CloudWatch autorizzazioni minime necessarie per inviare dati metrici OpenSearch Serverless a. CloudWatch

Puoi trovare la AmazonOpenSearchServerlessServiceRolePolicypolicy nella console IAM.

OpenSearch Aggiornamenti del servizio alle politiche AWS gestite

Visualizza i dettagli sugli aggiornamenti alle politiche AWS gestite per OpenSearch Service da quando questo servizio ha iniziato a tenere traccia delle modifiche.

Modifica	Descrizione	Data
È stato aggiornato il `AmazonOpenSearchIngestionServiceRolePolicy`	L'aggiornamento concede a OpenSearch Ingestion l'autorizzazione a modificare gli endpoint VPC creati da per condividere le pipeline OpenSearch tra loro. VPCs Per la policy JSON, consultare Console IAM.	28 agosto 2025
Aggiornato il `AmazonOpenSearchServiceRolePolicy`	È stata aggiunta la seguente dichiarazione alla politica. Quando Amazon OpenSearch Service assume il ruolo `AWSServiceRoleForAmazonOpenSearchService` collegato ai servizi, questa nuova dichiarazione nella politica consente di OpenSearch aggiornare l'ambito di accesso di qualsiasi AWS IAM Identity Center applicazione gestita solo da. OpenSearch `{ "Effect": "Allow",`	31 marzo 2025

Modifica	Descrizione	Data
	`"Action": "sso:PutApplicationAccessScope", "Resource": "arn:aws:sso::*:application/*/*", "Condition": { "StringEquals": { "aws:ResourceOrgID": "${aws:PrincipalOrgID}" } } }`	
Aggiornato `AmazonOpenSearchServerlessServiceRolePolicy`	È stato aggiunto il Sid `AllowAOSSCloudwatchMetrics` alla politica. `AmazonOpenSearchServerlessServiceRolePolicy` Un Sid è un ID di dichiarazione che funge da identificatore opzionale per l'informativa sulla politica.	12 luglio 2024

Modifica	Descrizione	Data
Aggiunto `OpenSearchIngestionSelfManagedVpcePolicy`	Una nuova policy che consente a OpenSearch Ingestion di abilitare l'accesso VPC autogestito per le pipeline di ingestione, creare tag e pubblicare metriche relative all'importazione sul tuo account. CloudWatch Per la policy JSON, consultare Console IAM.	12 giugno 2024
Aggiunto `AmazonOpenSearchDirectQueryGlueCreateAccess`	Concede ad Amazon OpenSearch Service Direct Query Service l'accesso a `CreateDatabase` `CreatePartition` ,`CreateTable` , e `BatchCreatePartition` AWS Glue API.	6 maggio 2024
Aggiornati `AmazonOpenSearchServiceRolePolicy` e `AmazonElasticsearchServiceRolePolicy`	Sono state aggiunte le autorizzazioni necessarie al ruolo collegato al servizio per assegnare e annullare l'assegnazione degli indirizzi . IPv6 Anche la policy obsoleta di Elasticsearch è stata aggiornata per garantire la compatibilità con le versioni precedenti.	18 ottobre 2023

Modifica	Descrizione	Data
Aggiunto `AmazonOpenSearchIngestionServiceRolePolicy`	Una nuova policy che consente a OpenSearch Ingestion di abilitare l'accesso VPC per le pipeline di ingestione, creare tag e pubblicare metriche relative all'importazione sul tuo account. CloudWatch Per la policy JSON, consultare Console IAM.	26 aprile 2023
Aggiunto `AmazonOpenSearchIngestionFullAccess`	Una nuova policy che garantisce l'accesso completo alle operazioni e alle risorse dell'API OpenSearch Ingestion per un. Account AWS Per la policy JSON, consultare Console IAM.	26 aprile 2023
Aggiunto `AmazonOpenSearchIngestionReadOnlyAccess`	Una nuova politica che garantisce l'accesso in sola lettura a tutte le risorse di OpenSearch Ingestion per un. Account AWS Per la policy JSON, consultare Console IAM.	26 aprile 2023

Modifica	Descrizione	Data
Aggiunto `AmazonOpenSearchServerlessServiceRolePolicy`	Una nuova policy che fornisce le autorizzazioni minime necessarie per inviare dati metrici OpenSearch Serverless a Amazon CloudWatch Per la policy JSON, consultare Console IAM.	29 novembre 2022
Aggiornati `AmazonOpenSearchServiceRolePolicy` e `AmazonElasticsearchServiceRolePolicy`	Sono state aggiunte le autorizzazioni necessarie per il ruolo collegato al servizio per creare endpoint VPC gestiti dal servizio OpenSearch . Alcune azioni possono essere eseguite solo quando la richiesta contiene il tag `OpenSearchManaged=true` . Anche la policy obsoleta di Elasticsearch è stata aggiornata per garantire la compatibilità con le versioni precedenti.	7 novembre 2022

Modifica	Descrizione	Data
Aggiornati `AmazonOpenSearchServiceRolePolicy` e `AmazonElasticsearchServiceRolePolicy`	È stato aggiunto il supporto per l'`PutMetricData` azione, necessaria per pubblicare i parametri OpenSearch del cluster su Amazon CloudWatch. Anche la policy obsoleta di Elasticsearch è stata aggiornata per garantire la compatibilità con le versioni precedenti. Per la policy JSON, consultare Console IAM.	12 settembre 2022
Aggiornati `AmazonOpenSearchServiceRolePolicy` e `AmazonElasticsearchServiceRolePolicy`	Aggiunto il supporto per il tipo di risorsa `acm`. La policy fornisce l'autorizzazione minima AWS Certificate Manager (ACM) di sola lettura necessaria al ruolo collegato al servizio per verificare e convalidare le risorse ACM al fine di creare e aggiornare domini personalizzati abilitati agli endpoint. Anche la policy obsoleta di Elasticsearch è stata aggiornata per garantire la compatibilità con le versioni precedenti.	28 luglio 2022

Modifica	Descrizione	Data
Aggiornati `AmazonOpenSearchServiceCognitoAccess` e `AmazonESCognitoAccess`	È stato aggiunto il supporto per l'`UpdateUserPoolClient` azione, necessaria per impostare la configurazione del pool di utenti di Cognito durante l'aggiornamento da Elasticsearch a. OpenSearch Autorizzazioni corrette per l'operazione `SetIdentityPoolRoles` in modo da consentire l'accesso a tutte le risorse. Anche la policy obsoleta di Elasticsearch è stata aggiornata per garantire la compatibilità con le versioni precedenti.	20 dicembre 2021
Aggiornato `AmazonOpenSearchServiceRolePolicy`	Aggiunto il supporto per il tipo di risorsa `security-group` . La policy fornisce le autorizzazioni minime di Amazon EC2 ed Elastic Load Balancing necessarie per il ruolo collegato ai servizi per abilitare l'accesso VPC.	9 settembre 2021

Modifica	Descrizione	Data
• Aggiunto `AmazonOpenSearchServiceFullAccess` • Obsoleta `AmazonESFullAccess`	Questa nuova policy ha lo scopo di sostituire la vecchia policy. Entrambe le policy forniscono l'accesso completo all'API di configurazione del OpenSearch servizio e a tutti i metodi HTTP per. OpenSearch APIs Il controllo granulare degli accessi e le policy basate sulle risorse possono comunque limitare l'accesso.	7 settembre 2021
• Aggiunto `AmazonOpenSearchServiceReadOnlyAccess` • Obsoleta `AmazonESReadOnlyAccess`	Questa nuova policy ha lo scopo di sostituire la vecchia policy. Entrambe le policy forniscono l'accesso in sola lettura all'API di configurazione del OpenSearch servizio (`es:Describe* es:List*`, `es:Get*`) e nessun accesso ai metodi HTTP per. OpenSearch APIs	7 settembre 2021
• Aggiunto `AmazonOpenSearchServiceCognitoAccess` • Obsoleta `AmazonESCognitoAccess`	Questa nuova policy ha lo scopo di sostituire la vecchia policy. Entrambe le policy forniscono le autorizzazioni minime di Amazon Cognito necessarie per abilitare l'autenticazione Cognito.	7 settembre 2021

Modifica	Descrizione	Data
• Aggiunto AmazonOpenSearchServiceRolePolicy • Obsoleta `AmazonElasticsearchServiceRolePolicy`	Questa nuova policy ha lo scopo di sostituire la vecchia policy. Entrambe le policy forniscono le autorizzazioni minime di Amazon EC2 ed Elastic Load Balancing necessarie per il ruolo collegato ai servizi per abilitare l'accesso VPC.	7 settembre 2021
Monitoraggio delle modifiche iniziato	Amazon OpenSearch Service ora tiene traccia delle modifiche alle politiche AWS gestite.	7 settembre 2021

Prevenzione del problema "confused deputy" tra servizi

Il problema confused deputy è un problema di sicurezza in cui un'entità che non dispone dell'autorizzazione per eseguire un'azione può costringere un'entità maggiormente privilegiata a eseguire l'azione. Inoltre AWS, l'impersonificazione tra servizi può portare al confuso problema del vicesceriffo. La rappresentazione tra servizi può verificarsi quando un servizio (il servizio chiamante) effettua una chiamata a un altro servizio (il servizio chiamato). Il servizio chiamante può essere manipolato per utilizzare le proprie autorizzazioni e agire sulle risorse di un altro cliente, a cui normalmente non avrebbe accesso. Per evitare ciò, AWS fornisce strumenti per poterti a proteggere i tuoi dati per tutti i servizi con entità di servizio a cui è stato concesso l'accesso alle risorse dell'account.

Ti consigliamo di utilizzare le chiavi di contesto `aws:SourceArn` e `aws:SourceAccount` global condition nelle politiche delle risorse per limitare le autorizzazioni che Amazon OpenSearch Service concede a un altro servizio alla risorsa. Se il valore `aws:SourceArn` non contiene l'ID account, ad esempio un ARN di un bucket Amazon S3, è necessario utilizzare entrambe le chiavi di contesto delle condizioni globali per limitare le autorizzazioni. Se si utilizzano entrambe le chiavi di contesto delle condizioni globali e il valore `aws:SourceArn` contiene l'ID account, il valore `aws:SourceAccount` e l'account nel valore `aws:SourceArn` deve utilizzare lo stesso ID

account nella stessa dichiarazione di policy. Utilizzare `aws:SourceArn` se si desidera consentire l'associazione di una sola risorsa all'accesso tra servizi. Utilizzare `aws:SourceAccount` se si desidera consentire l'associazione di qualsiasi risorsa in tale account all'uso tra servizi.

Il valore di `aws:SourceArn` deve essere l'ARN del dominio del OpenSearch servizio.

Il modo più efficace per proteggersi dal problema "confused deputy" è quello di usare la chiave di contesto della condizione globale `aws:SourceArn` con l'ARN completo della risorsa. Se non si conosce l'ARN completo della risorsa o si scelgono più risorse, è necessario utilizzare la chiave di contesto della condizione globale `aws:SourceArn` con caratteri jolly (*) per le parti sconosciute dell'ARN. Ad esempio, `arn:aws:es:*:123456789012:*`.

L'esempio seguente mostra come utilizzare le chiavi di contesto `aws:SourceArn` e `aws:SourceAccount` global condition in OpenSearch Service per evitare il confuso problema del vice.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":{
      "Sid":"ConfusedDeputyPreventionExamplePolicy",
      "Effect":"Allow",
      "Principal":{
         "Service":"es.amazonaws.com"
      },
      "Action":"sts:AssumeRole",
      "Condition":{
         "StringEquals":{
            "aws:SourceAccount":"111122223333"
         },
         "ArnLike":{
            "aws:SourceArn":"arn:aws:es:us-east-1:111122223333:domain/my-domain"
         }
      }
   }
}
```

Controllo granulare degli accessi in Amazon Service OpenSearch

Il controllo granulare degli accessi offre modi aggiuntivi per controllare l'accesso ai tuoi dati su Amazon Service. OpenSearch Ad esempio, a seconda di chi effettua la richiesta, è possibile che una ricerca restituisca risultati da un solo indice. Potresti voler nascondere determinati campi nei tuoi documenti o escludere del tutto determinati documenti.

Il controllo granulare degli accessi offre i seguenti vantaggi:

- Controllo degli accessi basato sui ruoli
- Sicurezza a livello di indice, documento e campo
- OpenSearch Dashboard multi-tenancy
- Autenticazione di base HTTP per e dashboard OpenSearch OpenSearch

Argomenti

- Il quadro generale: controllo granulare degli accessi e sicurezza dei servizi OpenSearch
- Concetti chiave
- Informazioni sull'utente principale
- Abilitazione del controllo granulare degli accessi
- Accesso alle OpenSearch dashboard come utente principale
- Gestione delle autorizzazioni
- Configurazioni consigliate
- Limitazioni
- Modifica dell'utente principale
- Utenti principali aggiuntivi
- Snapshot manuali
- Integrazioni
- Differenze della REST API
- Tutorial: configurazione di un dominio con un utente master IAM e autenticazione Amazon Cognito
- Tutorial: Configurare un dominio con il database utente interno e l'autenticazione di base HTTP

Il quadro generale: controllo granulare degli accessi e sicurezza dei servizi OpenSearch

La sicurezza di Amazon OpenSearch Service ha tre livelli principali:

Rete

Il primo livello di sicurezza è la rete, che determina se le richieste raggiungono un dominio OpenSearch di servizio. Se sceglisi sceglieAccesso pubblico quando creisi crea un dominio, le richieste provenienti da qualunquequalsiasi client connesso a Internet possono raggiungere l'endpoint del dominio. Se si sceglie Accesso VPC, i client devono connettersi al VPC (e i gruppi di sicurezza associati devono consentirlo) affinché una richiesta raggiunga l'endpoint. Per ulteriori informazioni, consultare the section called “Supporto per VPC”.

Policy di accesso al dominio

Il secondo livello di protezione è la policy di accesso al dominio. Dopo che una richiesta raggiunge un endpoint di dominio, la policy di accesso basato sulle risorse consente o nega la richiesta di accesso a un determinato URI. Il criterio di accesso accetta o rifiuta le richieste al «bordo» del dominio, prima che raggiungano OpenSearch stesso.

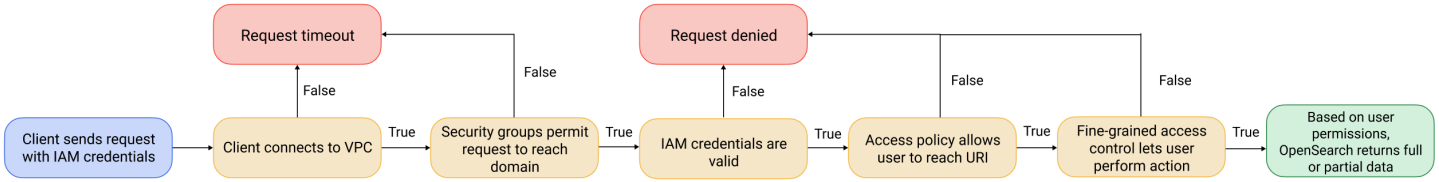
Controllo granulare degli accessi

Il terzo e ultimo livello di sicurezza è il controllo granulare degli accessi. Dopo che una policy di accesso basata sulle risorse consente a una richiesta di raggiungere un endpoint di dominio, il controllo granulare degli accessi valuta le credenziali utente e autentica l'utente o nega la richiesta. Se il controllo granulare degli accessi autentica l'utente, recupera tutti i ruoli mappati a tale utente e utilizza il set completo di autorizzazioni per determinare come gestire la richiesta.

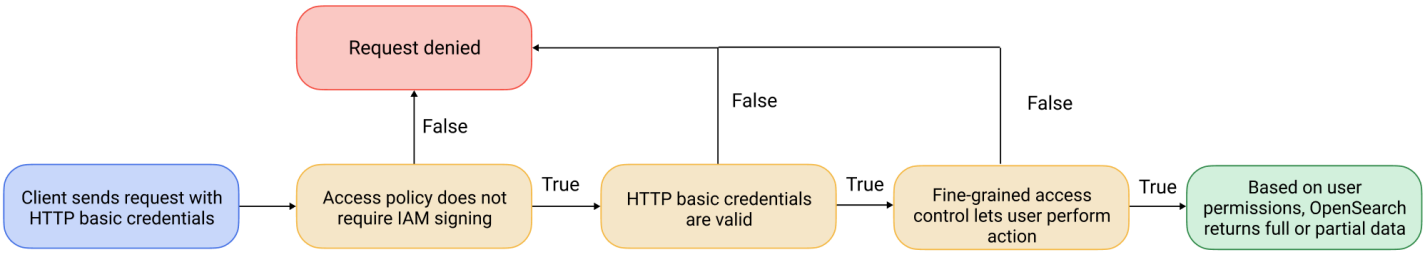
> **Note**
>
> Se una politica di accesso basata sulle risorse contiene ruoli o utenti IAM, i client devono inviare richieste firmate utilizzando AWS Signature Version 4. Pertanto, le policy di accesso possono entrare in conflitto con il controllo granulare degli accessi, soprattutto se si utilizza il database utente interno e l'autenticazione di base HTTP. Non puoi firmare una richiesta con nome utente e password e credenziali IAM. In generale, se si abilita il controllo granulare

degli accessi, si consiglia di utilizzare una policy di accesso al dominio che non richiede richieste firmate.

Questo primo diagramma illustra una configurazione comune: un dominio di accesso VPC con controllo granulare degli accessi abilitato, una policy di accesso basata su IAM e un utente principale IAM.



Il seguente diagramma illustra un'altra configurazione comune: un dominio di accesso pubblico con controllo granulare degli accessi abilitato, una policy di accesso che non utilizza i principal IAM e un utente master nel database utente interno.



Esempio

Considera una richiesta `GET` a `movies/_search?q=thor`. L'utente dispone delle autorizzazioni per cercare l'indice `movies`? In tal caso, l'utente dispone delle autorizzazioni per visualizzare tutti i documenti al suo interno? La risposta dovrebbe omettere o anonimizzare dei campi? Per l'utente master, la risposta potrebbe essere simile a questa:

```
{
  "hits": {
    "total": 7,
    "max_score": 8.772789,
    "hits": [{
        "_index": "movies",
        "_type": "_doc",
        "_id": "tt0800369",
        "_score": 8.772789,
        "_source": {
```

```
          "directors": [
            "Kenneth Branagh",
            "Joss Whedon"
          ],
          "release_date": "2011-04-21T00:00:00Z",
          "genres": [
            "Action",
            "Adventure",
            "Fantasy"
          ],
          "plot": "The powerful but arrogant god Thor is cast out of Asgard to
 live amongst humans in Midgard (Earth), where he soon becomes one of their finest
 defenders.",
          "title": "Thor",
          "actors": [
            "Chris Hemsworth",
            "Anthony Hopkins",
            "Natalie Portman"
          ],
          "year": 2011
        }
      },
      ...
    ]
  }
}
```

Se un utente con autorizzazioni più limitate emette esattamente la stessa richiesta, la risposta potrebbe essere simile a questa:

```
{
  "hits": {
    "total": 2,
    "max_score": 8.772789,
    "hits": [{
        "_index": "movies",
        "_type": "_doc",
        "_id": "tt0800369",
        "_score": 8.772789,
        "_source": {
          "year": 2011,
          "release_date":
 "3812a72c6dd23eef3c750c2d99e205cbd260389461e19d610406847397ecb357",
```

```
          "plot": "The powerful but arrogant god Thor is cast out of Asgard to
 live amongst humans in Midgard (Earth), where he soon becomes one of their finest
 defenders.",
          "title": "Thor"
        }
      },
      ...
    ]
  }
}
```

La risposta ha meno hit e meno campi per ogni hit. Inoltre, il campo `release_date` è anonimizzato. Se un utente senza autorizzazioni effettua la stessa richiesta, il cluster restituisce un errore:

```
{
  "error": {
    "root_cause": [{
      "type": "security_exception",
      "reason": "no permissions for [indices:data/read/search] and User [name=limited-
user, roles=[], requestedTenant=null]"
    }],
    "type": "security_exception",
    "reason": "no permissions for [indices:data/read/search] and User [name=limited-
user, roles=[], requestedTenant=null]"
  },
  "status": 403
}
```

Se un utente fornisce credenziali non valide, il cluster restituisce un'eccezione `Unauthorized`.

Concetti chiave

Per iniziare a utilizzare un controllo granulare degli accessi, considera i seguenti concetti:

- Ruoli: il modo principale di utilizzare il controllo granulare degli accessi. In questo caso, i ruoli sono distinti dai ruoli IAM. I ruoli contengono qualsiasi combinazione di autorizzazioni: a livello di cluster, a livello di indice, a livello di documento e a livello di campo.
- Mappatura: dopo aver configurato un ruolo, lo si associa a uno o più utenti. Ad esempio, è possibile mappare tre ruoli a un singolo utente: un ruolo che fornisce l'accesso a Dashboards, uno che fornisce l'accesso in sola lettura a `index1` e uno che fornisce l'accesso in scrittura a `index2`. In alternativa, è possibile includere tutte queste autorizzazioni in un singolo ruolo.

- Utenti: persone o applicazioni che effettuano richieste al OpenSearch cluster. Gli utenti dispongono di credenziali, chiavi di accesso IAM o nome utente e password, che specificano quando effettuano richieste.

Informazioni sull'utente principale

L'utente principale in OpenSearch Service è una combinazione di nome utente e password o un principale IAM che dispone delle autorizzazioni complete per il OpenSearch cluster sottostante. Un utente è considerato un utente principale se ha tutti gli accessi al OpenSearch cluster e la possibilità di creare utenti interni, ruoli e mappature dei ruoli all'interno delle dashboard. OpenSearch

Un utente master creato nella console di OpenSearch servizio o tramite la CLI viene automaticamente mappato su due ruoli predefiniti:

- `all_access`— Fornisce l'accesso completo a tutte le operazioni a livello di cluster, l'autorizzazione alla scrittura su tutti gli indici del cluster e l'autorizzazione alla scrittura a tutti i tenant.
- `security_manager`— Fornisce l'accesso al plug-in di sicurezza e la gestione di utenti e autorizzazioni.

Con questi due ruoli, l'utente ottiene l'accesso alla scheda Sicurezza nelle OpenSearch dashboard, dove può gestire utenti e autorizzazioni. Se si crea un altro utente interno e lo si associa solo al **`all_access`** ruolo, l'utente non ha accesso alla scheda Sicurezza. È possibile creare utenti principali aggiuntivi mappandoli esplicitamente a entrambi `all_access` i `security_manager` ruoli. Per istruzioni, consulta the section called “Utenti principali aggiuntivi”.

Quando crei un utente master per il tuo dominio, puoi specificare un principale IAM esistente o creare un utente master all'interno del database utenti interno. Considera quanto segue quando decidi quale usare:

- Principal IAM: se scegli un principale IAM per il tuo utente principale, tutte le richieste al cluster devono essere firmate utilizzando AWS Signature Version 4.

  OpenSearch Il servizio non prende in considerazione nessuna delle autorizzazioni del responsabile IAM. L'utente o il ruolo IAM serve esclusivamente per l'autenticazione. Le politiche relative a quell'utente o ruolo non influiscono sull'autorizzazione dell'utente principale. L'autorizzazione viene gestita tramite le varie autorizzazioni del plug-in OpenSearch Security.

Ad esempio, puoi assegnare zero autorizzazioni IAM a un principale IAM e, purché la macchina o la persona possa autenticarsi per quell'utente o ruolo, ha il potere dell'utente principale in Service. OpenSearch

Ti consigliamo IAM se desideri utilizzare gli stessi utenti su più cluster, se desideri utilizzare Amazon Cognito per accedere alle dashboard o se OpenSearch disponi di client che supportano la firma Signature versione 4.

- Database utenti interno: se crei un master nel database utenti interno (con una combinazione di nome utente e password), puoi utilizzare l'autenticazione di base HTTP (oltre alle credenziali IAM) per effettuare richieste al cluster. La maggior parte dei client supporta l'autenticazione di base, incluso curl, che supporta anche la versione 4 di AWS Signature con l'opzione --aws-sigv4. Il database utenti interno è memorizzato in un OpenSearch indice, quindi non è possibile condividerlo con altri cluster.

  È consigliabile utilizzare il database utente interno se non è necessario riutilizzare gli utenti in più cluster, se si desidera utilizzare l'autenticazione di base HTTP per accedere a Dashboards (anziché ad Amazon Cognito) o se si dispone di client che supportano solo l'autenticazione di base. Il database utenti interno è il modo più semplice per iniziare a usare OpenSearch Service.

Abilitazione del controllo granulare degli accessi

Abilita il controllo granulare degli accessi utilizzando la console o l'API di AWS CLI configurazione. Per le fasi, consulta Creazione e gestione dei domini.

Il controllo granulare degli accessi richiede OpenSearch Elasticsearch 6.7 o versione successiva. Richiede inoltre HTTPS per tutto il traffico verso il dominio, la crittografia dei dati inattivi e la crittografia. node-to-node A seconda di come configurate le funzionalità avanzate del controllo granulare degli accessi, l'ulteriore elaborazione delle richieste potrebbe richiedere risorse di calcolo e memoria su singoli nodi di dati. Dopo aver abilitato il controllo granulare degli accessi, non è più possibile disabilitarlo.

Abilitazione del controllo granulare degli accessi su domini esistenti

Puoi abilitare un controllo granulare degli accessi sui domini esistenti in esecuzione o su Elasticsearch 6.7 o versione successiva. OpenSearch

Per abilitare il controllo granulare degli accessi su un dominio esistente (console)

1. Seleziona il dominio e scegli Operazioni quindi Modifica configurazione di sicurezza.
2. True per abilitare il controllo granulare degli accessi.
3. Scegli come creare l'utente master:
   - Se si desidera utilizzare IAM per la gestione degli utenti, scegliere Imposta ARN IAM come utente principale e specificare l'ARN per un ruolo IAM.
   - Se desideri utilizzare il database utenti interno, scegli Crea utente principale e specifica un nome utente e una password.
4. (Opzionale) Seleziona Enable migration period for open/IP-based access policy (Abilita il periodo di migrazione per le policy di accesso open/basati su IP). Questa impostazione consente un periodo di transizione di 30 giorni durante il quale gli utenti esistenti possono continuare ad accedere al dominio senza interruzioni e Policy di accesso basate su IP aperte e esistenti continueranno a lavorare con il tuo dominio. Durante questo periodo di migrazione, consigliamo agli amministratori creare i ruoli necessari e mapparli agli utenti per il dominio. Se si utilizzano policy basate su identità anziché un criterio di accesso aperto o basato su IP, è possibile disabilitare questa impostazione.

   È inoltre necessario aggiornare i client per lavorare con un controllo granulare degli accessi durante il periodo di migrazione. Ad esempio, se mappi i ruoli IAM con un controllo di accesso granulare, devi aggiornare i tuoi client per iniziare a firmare le richieste con AWS Signature Version 4. Se si configura l'autenticazione di base HTTP con un controllo granulare degli accessi, è necessario aggiornare i client per fornire le credenziali di autenticazione di base appropriate nelle richieste.

   Durante il periodo di migrazione, gli utenti che accedono all'endpoint OpenSearch Dashboards per il dominio accederanno direttamente alla pagina Discover anziché alla pagina di accesso. Gli amministratori e gli utenti master possono scegliere Login per accedere con le credenziali di amministratore e configurare i mapping dei ruoli.

   > **Important**
   >
   > OpenSearch Il servizio disabilita automaticamente il periodo di migrazione dopo 30 giorni. Si consiglia di terminarlo non appena crei i ruoli necessari e li mappi agli utenti. Al termine del periodo di migrazione, non è possibile riattivarlo.

5. Scegliere Save changes (Salva modifiche).

Il cambiamento innesca una implementazione blu/verde durante la quale l'integrità del cluster diventa rossa, ma tutte le operazioni del cluster rimangono inalterate.

Per abilitare il controllo granulare degli accessi su un dominio esistente (CLI)

Imposta `AnonymousAuthEnabled` su `true` per abilitare il periodo di migrazione con un controllo granulare degli accessi:

```
aws opensearch update-domain-config --domain-name test-domain --region us-east-1 \
      --advanced-security-options '{ "Enabled": true,
 "InternalUserDatabaseEnabled":true, "MasterUserOptions": {"MasterUserName":"master-
username","MasterUserPassword":"master-password"},"AnonymousAuthEnabled": true}'
```

Informazioni sul ruolo default

Un controllo granulare degli accessi richiede una mappatura dei ruoli. Se il tuo dominio utilizza politiche di accesso basate sull'identità, OpenSearch Service associa automaticamente gli utenti a un nuovo ruolo chiamato default_role per aiutarti a migrare correttamente gli utenti esistenti. Questa mappatura temporanea garantisce che gli utenti possano ancora inviare correttamente le richieste GET e PUT firmate IAM fino a quando non crei i tuoi mapping dei ruoli personalizzati.

Il ruolo non aggiunge vulnerabilità o difetti di sicurezza al dominio del servizio. OpenSearch
Ti consigliamo di eliminare il ruolo predefinito non appena configuri i tuoi ruoli e di mapparli di conseguenza.

Scenari di migrazione

Nella tabella seguente viene descritto il comportamento di ciascun metodo di autenticazione prima e dopo aver abilitato il controllo granulare degli accessi su un dominio esistente e i passaggi che gli amministratori devono adottare per mappare correttamente i propri utenti ai ruoli:

Metodo di autentica zione	Prima dell'abil itazione del controllo dettagliato degli accessi	Dopo l'abilitazione del controllo dettagliato degli accessi	Attività amministrative
Policy basate sull'identità	Tutti gli utenti IAM che soddisfan o la policy IAM possono accedere al dominio.	Non è necessario abilitare il periodo di migrazione. OpenSearch Il servizio mappa automaticamente tutti gli utenti che soddisfano la policy IAM sul default_r ole in modo che possano continuar e ad accedere al dominio.	1. Crea mappature di ruolo personali zzate sul dominio. 2. Elimina il default_role.
Policy basate su IP	Tutti gli utenti degli indirizzi IP o dei blocchi CIDR consentit i possono accedere al dominio.	Durante il periodo di migrazione di 30 giorni, tutti gli utenti degli indirizzi IP o dei blocchi CIDR consentiti possono continuar e ad accedere al dominio.	1. Crea mappature di ruolo personali zzate sul dominio. 2. Aggiorna i tuoi client per fornire credenziali di autenticazione di base o credenziali IAM, a seconda della configurazione della mappatura dei ruoli. 3. Disabilitare il periodo di migrazione. Gli utenti degli indirizzi IP consentit i o dei blocchi CIDR che inviano richieste senza autenticazione di base o credenziali IAM perderanno l'accesso al dominio.

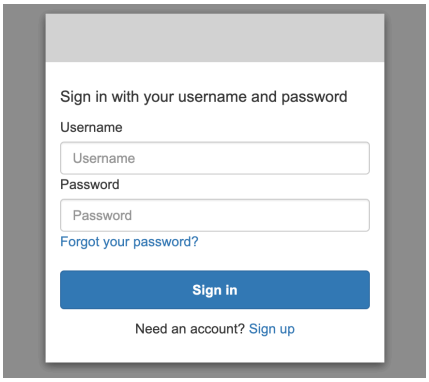
Metodo di autentica zione	Prima dell'abil itazione del controllo dettagliato degli accessi	Dopo l'abilitazione del controllo dettagliato degli accessi	Attività amministrative
Policy di accesso aperto	Tutti gli utenti su Internet possono accedere al dominio.	Durante il periodo di migrazione di 30 giorni, tutti gli utenti su Internet possono continuar e ad accedere al dominio.	1. Crea mappature dei ruoli sul dominio. 2. Aggiorna i tuoi client per fornire credenziali di autenticazione di base o credenziali IAM, a seconda della configurazione della mappatura dei ruoli. 3. Disabilitare il periodo di migrazione. Gli utenti che inviano richieste senza autenticazione di base o credenzia li IAM perderanno l'accesso al dominio.

Accesso alle OpenSearch dashboard come utente principale

Il controllo granulare degli accessi ha un plug-in OpenSearch Dashboards che semplifica le attività di gestione. È possibile utilizzare Dashboards per gestire utenti, ruoli, mappature, gruppi di operazioni e tenant. La pagina di accesso a OpenSearch Dashboards e il metodo di autenticazione sottostante differiscono, tuttavia, a seconda di come gestisci gli utenti e configuri il dominio.

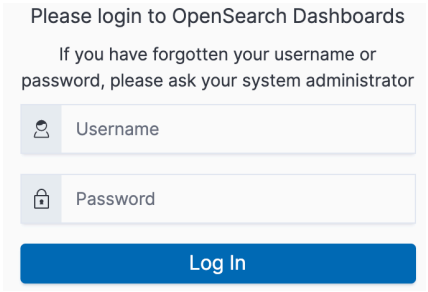
- Se si desidera utilizzare IAM per la gestione degli utenti, utilizzare the section called “Autenticazione Amazon Cognito per dashboard OpenSearch ” per accedere a Dashboards. Altrimenti, Dashboards mostra una pagina di accesso non funzionale. Per informazioni, consultare the section called “Limitazioni”.

  Con l'autenticazione di Amazon Cognito, uno dei ruoli assunti dal pool di identità deve corrispondere al ruolo IAM specificato per l'utente principale. Per ulteriori informazioni su questa configurazione, consulta the section called “(Facoltativo) Configurazione dell'accesso granulare” e the section called “Tutorial: controllo granulare degli accessi con autenticazione Cognito”.



- Se scegli di utilizzare il database utenti interno, puoi accedere a Dashboards con il nome utente e la password principali. È necessario accedere a Dashboards tramite HTTPS. L'autenticazione Amazon Cognito e SAML per Dashboards sostituiscono entrambe questa schermata di accesso.

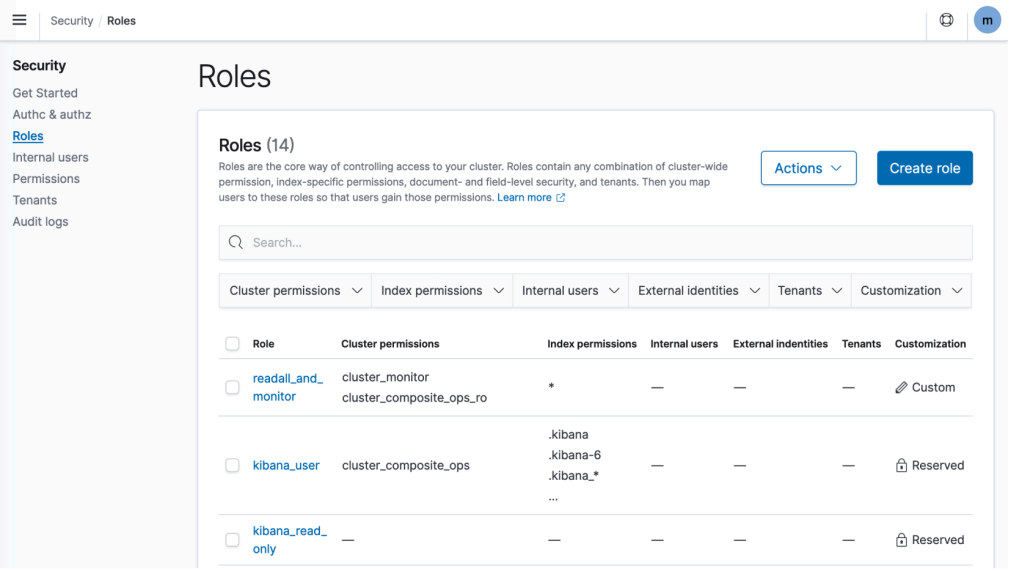
  Per ulteriori informazioni su questa configurazione, consulta the section called “Tutorial: Database utente interno e autenticazione di base”.



- Se si decide di utilizzare l'autenticazione SAML, è possibile accedere utilizzando le credenziali di un provider di identità esterno. Per ulteriori informazioni, consultare the section called “Autenticazione SAML per dashboard OpenSearch”.

Gestione delle autorizzazioni

Come indicato in the section called “Concetti chiave”, è possibile gestire le autorizzazioni del controllo granulare degli accessi a utilizzando ruoli, utenti e mappature. In questa sezione viene descritto come creare e applicare tali risorse. Per eseguire queste operazioni si consiglia di accedere a Dashboards come utente principale.



> **Note**
>
> Le autorizzazioni che scegli di concedere agli utenti variano ampiamente in base al caso d'uso. Non possiamo coprire in maniera fattibile tutti gli scenari contenuti in questa documentazione. Mentre decidi quali autorizzazioni concedere ai tuoi utenti, assicurati di fare riferimento alle autorizzazioni per OpenSearch cluster e indice menzionate nelle sezioni seguenti e segui sempre il principio del privilegio minimo.

Creazione di ruoli

Puoi creare nuovi ruoli per il controllo granulare degli accessi utilizzando le OpenSearch dashboard o l'operazione nell'API REST. `_plugins/_security` Per ulteriori informazioni, consulta Creazione di ruoli.

Il controllo granulare degli accessi include anche numerosi ruoli predefiniti. Client come OpenSearch Dashboards e Logstash inviano un'ampia varietà di richieste OpenSearch, il che può rendere difficile la creazione manuale di ruoli con il set minimo di autorizzazioni. Ad esempio, il ruolo `opensearch_dashboards_user` include le autorizzazioni necessarie a un utente per utilizzare modelli di indice, visualizzazioni, dashboard e tenant. Si consiglia di associarlo a qualsiasi ruolo utente o back-end che accede a Dashboards, insieme a ruoli aggiuntivi che consentono l'accesso ad altri indici.

Amazon OpenSearch Service non offre i seguenti OpenSearch ruoli:

- `observability_full_access`
- `observability_read_access`
- `reports_read_access`
- `reports_full_access`

Amazon OpenSearch Service offre diversi ruoli che non sono disponibili con OpenSearch:

- `ultrawarm_manager`
- `ml_full_access`
- `cold_manager`
- `notifications_full_access`
- `notifications_read_access`

Sicurezza a livello di cluster

Le autorizzazioni a livello di cluster includono la possibilità di eseguire richieste generiche quali `_mget`, `_msearch`, e `_bulk`, monitorare l'integrità, acquisire snapshot e altro ancora. Gestire queste autorizzazioni utilizzando la sezione Autorizzazioni cluster durante la creazione di un ruolo. Per l'elenco completo delle autorizzazioni a livello di cluster, consulta Autorizzazioni cluster.

Invece delle autorizzazioni individuali, spesso puoi ottenere la posizione di sicurezza desiderata utilizzando una combinazione di gruppi di azioni predefiniti. Per un elenco dei gruppi di operazioni a livello di cluster, consultare Livello di cluster.

Sicurezza a livello di indice

Le autorizzazioni a livello di indice includono la possibilità di creare nuovi indici, indici di ricerca, leggere e scrivere documenti, eliminare documenti, gestire alias e altro ancora. Gestire queste autorizzazioni utilizzando la sezione Autorizzazioni indice durante la creazione di un ruolo. Per l'elenco completo delle autorizzazioni a livello di indice, consulta Autorizzazioni indice.

Invece delle autorizzazioni individuali, spesso puoi ottenere la posizione di sicurezza desiderata utilizzando una combinazione di gruppi di azioni predefiniti. Per un elenco dei gruppi di operazioni a livello di indice, consultare Livello di indice.

Sicurezza a livello di documento

La sicurezza a livello di documento consente di limitare i documenti in un indice che un utente può visualizzare. Quando crei un ruolo, specifica uno schema di indice e una OpenSearch query. Tutti gli utenti che si associano a tale ruolo possono visualizzare solo i documenti corrispondenti alla query. La sicurezza a livello di documento influisce sul numero di visite ricevute durante la ricerca.

Per ulteriori informazioni, consultare Sicurezza a livello di documento.

Sicurezza a livello di campo

La sicurezza a livello di campo consente di controllare quali campi di documento possono essere visualizzati dall'utente. Quando si crea un ruolo, aggiungere un elenco di campi da includere o escludere. Se si includono campi, gli utenti per cui si esegue il mapping a tale ruolo possono visualizzare solo i campi. Se si escludono i campi, possono visualizzare tutti i campi tranne quelli esclusi. La sicurezza a livello di campo influisce sul numero di campi inclusi negli hit durante la ricerca.

Per ulteriori informazioni, consultare Sicurezza a livello di campo.

Mascheramento del campo

Il mascheramento dei campi è un'alternativa alla sicurezza a livello di campo che consente di anonimizzare i dati in un campo anziché rimuoverli del tutto. Quando si crea un ruolo, aggiungere un elenco di campi da mascherare. Il mascheramento dei campi influisce sulla possibilità di visualizzare il contenuto di un campo durante la ricerca.

> **Tip**
>
> Se applichi il mascheramento standard a un campo, OpenSearch Service utilizza un hash sicuro e casuale che può causare risultati di aggregazione imprecisi. Per eseguire aggregazioni su campi mascherati, utilizzare invece il mascheramento basato su modello.

Creazione di utenti

Se hai abilitato il database utenti interno, puoi creare utenti utilizzando le OpenSearch dashboard o l'`_plugins/_security`operazione nell'API REST. Per ulteriori informazioni, consultare Creazione di utenti.

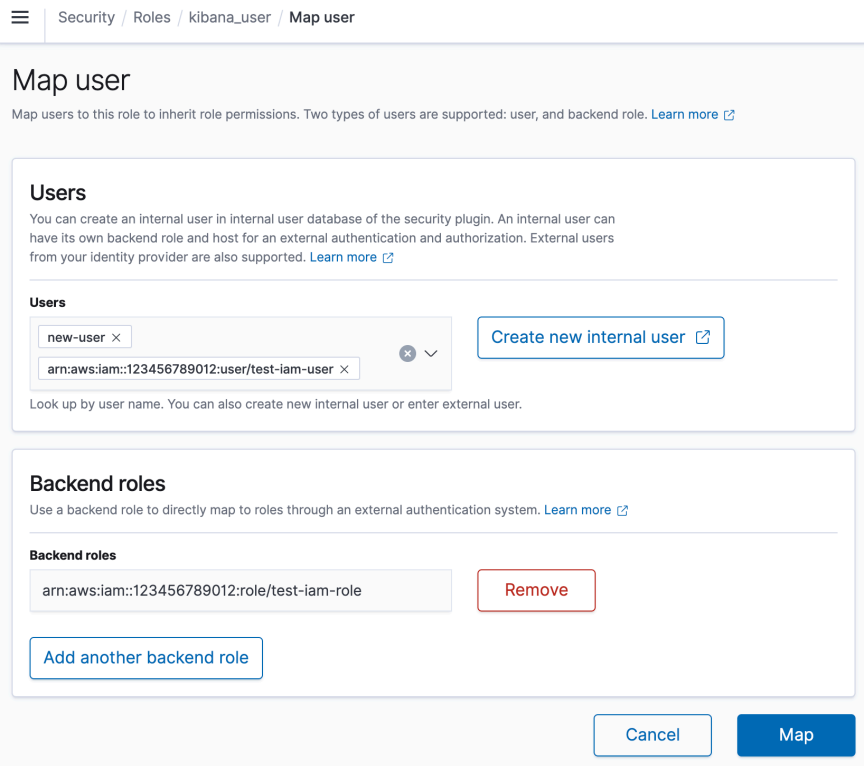
Se è stato scelto IAM per l'utente principale, ignorare questa parte di Dashboards. Creare invece ruoli IAM. Per ulteriori informazioni, consultare la Guida per l'utente IAM.

Mappatura dei ruoli agli utenti

La mappatura dei ruoli è l'aspetto più critico del controllo granulare degli accessi. Il controllo granulare degli accessi dispone di alcuni ruoli predefiniti che consentono di iniziare, ma a meno che non si esegua la mappatura dei ruoli agli utenti, ogni richiesta al cluster termina con un errore di autorizzazioni.

I ruoli di backend possono aiutare a semplificare il processo di mappatura dei ruoli. Invece di mappare lo stesso ruolo a 100 singoli utenti, puoi mappare il ruolo a un singolo ruolo di backend condiviso da tutti i 100 utenti. I ruoli di back-end possono essere ruoli IAM o stringhe arbitrarie.

- Specificare gli utenti ARNs, gli utenti e le stringhe utente di Amazon Cognito nella sezione Utenti. Le stringhe utente di Cognito assumono la forma di `Cognito/`*`user-pool-id`*`/`*`username`*.
- Specificate i ruoli di backend e il ruolo IAM ARNs nella sezione Ruoli di backend.



Puoi mappare i ruoli agli utenti utilizzando OpenSearch le dashboard o l'`_plugins/_security`operazione nell'API REST. Per ulteriori informazioni, consultare Mappa utenti ai ruoli.

Creazione di gruppi di operazioni

I gruppi di azioni sono insiemi di autorizzazioni che è possibile riutilizzare in diverse risorse. È possibile creare nuovi gruppi di azioni utilizzando OpenSearch le dashboard o l'`_plugins/_security`operazione nell'API REST, sebbene i gruppi di azioni predefiniti siano sufficienti per la maggior parte dei casi d'uso. Per ulteriori informazioni sui gruppi di operazioni predefiniti, consultare Gruppi di operazioni predefiniti.

OpenSearch Dashboard multi-tenancy

I tenant sono spazi per salvare modelli di indici, visualizzazioni, pannelli di controllo e altri oggetti Dashboards. La funzionalità multi-tenancy di Dashboards ti consente di condividere in sicurezza il tuo lavoro con altri utenti di Dashboards (o di mantenerlo privato) e di configurare dinamicamente i tenant. È possibile controllare quali ruoli hanno accesso a un tenant e se tali ruoli hanno accesso in lettura o in scrittura. Il tenant globale è l'impostazione predefinita. Per ulteriori informazioni, consulta Dashboards multi-tenancy. OpenSearch

Per visualizzare il tenant corrente o modificare i tenant

1. Vai alle OpenSearch dashboard e accedi.
2. Seleziona l'icona utente in alto a destra e sceglie Cambia tenant.
3. Verificare il tenant prima di creare visualizzazioni o dashboard. Se si desidera condividere il proprio lavoro con tutti gli altri utenti di Dashboards, scegliere Globale. Per condividere il lavoro con un sottoinsieme di utenti di Dashboards, scegliere un tenant condiviso diverso. Altrimenti, scegli Privato.

> **Note**
>
> OpenSearch Dashboards mantiene un indice separato per ogni tenant e crea un modello di indice chiamato. `tenant_template` Non eliminate o modificate l'`tenant_template`indice, poiché se la mappatura dell'indice dei tenant non è configurata correttamente, ciò potrebbe causare il malfunzionamento dei OpenSearch dashboard.

Configurazioni consigliate

A causa del modo in cui il controllo granulare degli accessi interagisce con altre funzionalità di sicurezza, consigliamo diverse configurazioni di controllo granulare degli accessi che funzionano bene per la maggior parte dei casi d'uso.

Description	Utente principale	Policy di accesso al dominio
Utilizza le credenziali IAM per le chiamate a e utilizza l'autenticazione SAML OpenSearch APIs per accedere alle dashboard. Gestire i ruoli del controllo granulare degli accessi utilizzando Dashboards o la REST API.	Utente o ruolo IAM	JSON `{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Principal": { "AWS": "*" }, "Action": "es:ESHttp*", "Resource": "arn:aws:es: us-east-1 :111122223333 :domain/domain-name /*" } ] }`
Utilizza le credenziali IAM o l'autenticazione di base per le chiamate a. OpenSearch APIs Gestire i ruoli del controllo granulare degli accessi utilizzando Dashboards o la REST API. Questa configurazione offre molta flessibilità, soprattutto se disponi di	Nome utente e password	JSON `{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Principal": { "AWS": "*" }, "Action": "es:ESHttp*",`

Description	Utente principale	Policy di accesso al dominio

OpenSearch client che supportano solo l'autenticazione di base.
 	
Se si dispone di un provider di identità esistente, utilizzare Autenticazione SAML per accedere a Dashboards. In caso contrario, gestire gli utenti di Dashboards nel database interno degli utenti.
		<pre>"Resource": "arn:aws:es: us-east-1 :111122223333 :domain/domain-name /*"
		}
		]
		}</pre>

Usa le credenziali IAM per le chiamate verso e usa Amazon Cognito per accedere alle dashboard . OpenSearch APIs Gestire i ruoli del controllo granulare degli accessi utilizzando Dashboards o la REST API.	Utente o ruolo IAM	JSON
		
		<pre>{
		"Version":"2012-10-17",
		"Statement": [
		{
		"Effect": "Allow",
		"Principal": {
		"AWS": "*"
		},
		"Action": "es:ESHttp*",
		"Resource": "arn:aws:es: us-east-1 :111122223333 :domain/domain-name /*"
		}
		]
		}</pre>

Description	Utente principale	Policy di accesso al dominio

Utilizza le credenziali IAM per le chiamate e blocca la OpenSearch APIs maggior parte degli accessi alle dashboard. Gestire i ruoli del controllo granulare degli accessi utilizzando la REST API.	Utente o ruolo IAM	JSON
		{
		"Version":"2012-10-17",
		"Statement": [
		{
		"Effect": "Allow",
		"Principal": {
		"AWS": "*"
		},
		"Action": "es:ESHttp*",
		"Resource": "arn:aws:es: *us-east-1* :*111122223333* :domain/*domain-name* /*"
		},
		{
		"Effect": "Deny",
		"Principal": {
		"AWS": "*"
		},
		"Action": "es:ESHttp*",
		"Resource": "arn:aws:es: *us-east-1* :*111122223333* :domain/*domain-name* /_dashboards*"
		}
		]
		}

Limitazioni

Il controllo granulare degli accessi presenta diverse limitazioni importanti:

- L'aspetto `hosts` delle mappature dei ruoli, che associa i ruoli a nomi host o indirizzi IP, non funziona se il dominio è all'interno di un VPC. È comunque possibile mappare i ruoli agli utenti e ai ruoli di back-end.

- Se sceglisi sceglie IAM per l'utente principale e non abilitisi abilita l'autenticazione Amazon Cognito o SAML, Dashboards visualizza una pagina di accesso non funzionante.funzionale.
- Se si sceglie IAM per l'utente master, è comunque possibile creare utenti nel database utente interno. Poiché l'autenticazione di base HTTP non è abilitata in questa configurazione, tutte le richieste firmate con tali credenziali utente vengono rifiutate.
- Se si utilizza SQL per eseguire una query su un indice a cui non si ha accesso, viene visualizzato un errore "no permissions". Se l'indice non esiste, viene visualizzato un errore «tale indice è inesistente». Questa differenza nei messaggi di errore significa che puoi confermare l'esistenza di un indice se ti capita di indovinare il suo nome.

  Per ridurre al minimo il problema, non includere informazioni riservate nei nomi degli indici. Per negare tutti gli accessi a SQL, aggiungere il seguente elemento alla policy di accesso del dominio:

```
{
  "Effect": "Deny",
  "Principal": {
    "AWS": [
      "*"
    ]
  },
  "Action": [
    "es:*"
  ],
  "Resource": "arn:aws:es:us-east-1:123456789012:domain/my-domain/_plugins/_sql"
}
```

- Se la versione del tuo dominio è 2.3 o successiva e hai abilitato il controllo granulare degli accessi, l'impostazione su 1 `max_clause_count` causa problemi con il dominio. Ti consigliamo di impostare questo account su un numero più alto.
- Se stai abilitando il controllo granulare degli accessi in un dominio in cui non è impostato il controllo granulare degli accessi, per le fonti di dati create per l'interrogazione diretta, devi configurare tu stesso ruoli di controllo degli accessi granulari. Per ulteriori informazioni su come configurare ruoli di accesso granulari, consulta Creazione di integrazioni di origini dati di Amazon OpenSearch Service con Amazon S3.

Modifica dell'utente principale

Se si dimenticano i dettagli dell'utente master, è possibile riconfigurarlo utilizzando la console, AWS CLI o l'API di configurazione.

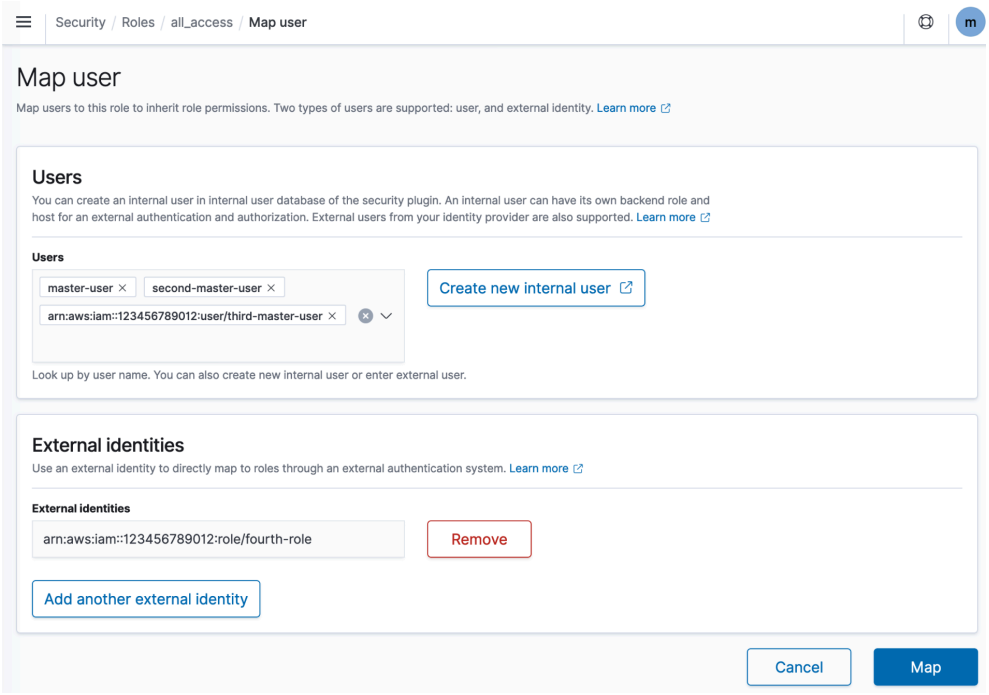
Per modificare l'utente master (console)

1. Accedi alla console di Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa/.
2. Seleziona il dominio e scegli Actions (Operazioni) quindi Edit security configuration (Modifica configurazione di sicurezza).
3. Scegliere Imposta ARN IAM come utente principale o Crea utente principale.
   - Se in precedenza è stato utilizzato un utente master IAM, il controllo granulare degli accessi esegue nuovamente la mappatura del ruolo `all_access` al nuovo ARN IAM specificato.
   - Se in precedenza è stato utilizzato il database utente interno, il controllo granulare degli accessi crea un nuovo utente principale. È possibile utilizzare il nuovo utente master per eliminare quello vecchio.
   - Il passaggio dal database utente interno a un utente principale IAM non elimina gli utenti dal database utente interno. Invece, disabilita semplicemente l'autenticazione di base HTTP. Eliminare manualmente gli utenti dal database utente interno o conservarli nel caso in cui sia necessario riabilitare l'autenticazione di base HTTP.
4. Scegli Save changes (Salva modifiche).

Utenti principali aggiuntivi

Si designa un utente master quando si crea un dominio, ma se si desidera, è possibile utilizzare questo utente master per creare ulteriori utenti master. Hai due opzioni: OpenSearch dashboard o API REST.

- In Dashboards scegliere Sicurezza, Ruoli, quindi associare il nuovo utente principale ai ruoli `all_access` e `security_manager`.



- Per utilizzare la REST API, inviare le seguenti richieste:

```
PUT _plugins/_security/api/rolesmapping/all_access
{
  "backend_roles": [
    "arn:aws:iam::123456789012:role/fourth-master-user"
  ],
  "hosts": [],
  "users": [
    "master-user",
    "second-master-user",
    "arn:aws:iam::123456789012:user/third-master-user"
  ]
}
```

```
PUT _plugins/_security/api/rolesmapping/security_manager
{
  "backend_roles": [
    "arn:aws:iam::123456789012:role/fourth-master-user"
  ],
  "hosts": [],
  "users": [
    "master-user",
    "second-master-user",
    "arn:aws:iam::123456789012:user/third-master-user"
```

```
  ]
}
```

Queste richieste sostituiscono le mappature dei ruoli correnti, quindi eseguire prima le richieste `GET` in modo da poter includere tutti i ruoli correnti nelle richieste `PUT`. La REST API è particolarmente utile se non è possibile accedere a Dashboards e si desidera mappare un ruolo IAM da Amazon Cognito al ruolo `all_access`.

Snapshot manuali

Il controllo granulare degli accessi introduce alcune complicazioni aggiuntive con l'acquisizione di snapshot manuali. Per registrare un repository di snapshot, anche se si utilizza l'autenticazione di base HTTP per tutti gli altri scopi, è necessario associare il ruolo `manage_snapshots` a un ruolo IAM che dispone delle autorizzazioni `iam:PassRole` per assumere `TheSnapshotRole`, come definito in the section called “Prerequisiti”.

Utilizzare quindi il ruolo IAM per inviare una richiesta firmata al dominio, come descritto in the section called “Registrazione di un repository di snapshot manuali”.

Integrazioni

Se utilizzi altri AWS servizi con OpenSearch Service, devi fornire i ruoli IAM per tali servizi con le autorizzazioni appropriate. Ad esempio, i flussi di distribuzione di Firehose utilizzano spesso un ruolo IAM chiamato. `firehose_delivery_role` In Dashboards, creare un ruolo per il controllo granulare degli accessi e mappare il ruolo IAM a tale ruolo. In questo caso, il nuovo ruolo richiede le seguenti autorizzazioni:

```
{
  "cluster_permissions": [
    "cluster_composite_ops",
    "cluster_monitor"
  ],
  "index_permissions": [{
    "index_patterns": [
      "firehose-index*"
    ],
    "allowed_actions": [
      "create_index",
      "manage",
```

```
      "crud"
    ]
  }]
}
```

Le autorizzazioni variano in base alle azioni eseguite da ciascun servizio. Una AWS IoT regola o una AWS Lambda funzione che indicizza i dati richiede probabilmente autorizzazioni simili a quelle di Firehose, mentre una funzione Lambda che esegue solo ricerche può utilizzare un set più limitato.

Differenze della REST API

L'API REST di controllo degli accessi a grana fine differisce leggermente a seconda della versione di OpenSearch/Elasticsearch . Prima di effettuare una richiesta PUT, effettuare una richiesta `GET` per verificare il corpo della richiesta previsto. Ad esempio, una richiesta `GET` per `_plugins/_security/api/user` restituisce tutti gli utenti, che è possibile modificare e utilizzare per effettuare richieste PUT valide.

In Elasticsearch 6.x, le richieste per creare utenti hanno il seguente aspetto:

```
PUT _opendistro/_security/api/user/new-user
{
  "password": "some-password",
  "roles": ["new-backend-role"]
}
```

Su OpenSearch Elasticsearch 7.x, le richieste hanno questo aspetto (modificale se usi Elasticsearch): `_plugins` `_opendistro`

```
PUT _plugins/_security/api/user/new-user
{
  "password": "some-password",
  "backend_roles": ["new-backend-role"]
}
```

Inoltre, i tenant sono proprietà dei ruoli in Elasticsearch 6.x:

```
GET _opendistro/_security/api/roles/all_access

{
```

```
  "all_access": {
    "cluster": ["UNLIMITED"],
    "tenants": {
      "admin_tenant": "RW"
    },
    "indices": {
      "*": {
        "*": ["UNLIMITED"]
      }
    },
    "readonly": "true"
  }
}
```

In OpenSearch Elasticsearch 7.x, sono oggetti con un proprio URI (da modificare se si utilizza Elasticsearch): `_plugins` `_opendistro`

```
GET _plugins/_security/api/tenants

{
  "global_tenant": {
    "reserved": true,
    "hidden": false,
    "description": "Global tenant",
    "static": false
  }
}
```

Per la documentazione sull'API OpenSearch REST, consulta il riferimento all'API del plug-in di sicurezza.

Tip

Se si utilizza il database utente interno, è possibile utilizzare curl per effettuare richieste e testare il dominio. Provare i seguenti comandi di esempio:

```
curl -XGET -u 'master-user:master-user-password' 'domain-endpoint/_search'
curl -XGET -u 'master-user:master-user-password' 'domain-endpoint/_plugins/
_security/api/user'
```

Tutorial: configurazione di un dominio con un utente master IAM e autenticazione Amazon Cognito

Questo tutorial illustra un popolare caso d'uso di Amazon OpenSearch Service per il controllo granulare degli accessi: un utente master IAM con autenticazione Amazon Cognito for Dashboards. OpenSearch

Nel tutorial configureremo un ruolo IAM master e un ruolo IAM limitato, che poi assoceremo agli utenti in Amazon Cognito. L'utente master può quindi accedere a OpenSearch Dashboards, mappare l'utente limitato a un ruolo e utilizzare un controllo granulare degli accessi per limitare le autorizzazioni dell'utente.

Sebbene questi passaggi utilizzino il bacino d'utenza di Amazon Cognito per l'autenticazione, questo stesso processo di base funziona per qualsiasi provider di autenticazione Cognito che consente di assegnare ruoli IAM differenti a utenti diversi.

In questo tutorial completerai le seguenti fasi:

1. Creazione di ruoli IAM master e limitati
2. Creazione di un dominio con l'autenticazione Cognito
3. Configurare un pool di utenti e un pool di identità di Cognito
4. Mappa i ruoli nelle dashboard OpenSearch
5. Test delle autorizzazioni

Fase1: Creazione di ruoli IAM master e limitati

Passa alla console AWS Identity and Access Management (IAM) e crea due ruoli separati:

- `MasterUserRole` - L'utente master, che disporrà di autorizzazioni complete per il cluster e gestirà ruoli e mappature dei ruoli.

- `LimitedUserRole` - Un ruolo più limitato, a cui come utente master concederai un accesso limitato.

Per istruzioni su come creare i ruoli, consulta Creazione di un ruolo utilizzando policy di fiducia personalizzate nella Guida per l'utente IAM.

Entrambi i ruoli devono disporre della seguente policy di attendibilità che consente al pool di identità Cognito di assumere i ruoli:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "Federated": "cognito-identity.amazonaws.com"
    },
    "Action": "sts:AssumeRoleWithWebIdentity",
    "Condition": {
      "StringEquals": {
        "cognito-identity.amazonaws.com:aud": "{identity-pool-id}"
      },
      "ForAnyValue:StringLike": {
        "cognito-identity.amazonaws.com:amr": "authenticated"
      }
    }
  }]
}
```

> **Note**
>
> Sostituzione di `identity-pool-id` con l'identificatore univoco del pool di identità Amazon Cognito. Ad esempio, `us-east-1:0c6cdba7-3c3c-443b-a958-fb9feb207aa6`.

Fase 2: Creazione di un dominio con l'autenticazione Cognito

Accedi alla console di Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa/ e crea un dominio con le seguenti impostazioni:

- OpenSearch 1.0 o versione successiva oppure Elasticsearch 7.8 o versione successiva
- Accesso pubblico
- Controllo granulare degli accessi abilitato con `MasterUserRole` come utente master (creato nella fase precedente)
- Autenticazione Amazon Cognito abilitata per le OpenSearch dashboard. Per le istruzioni sull'abilitazione dell'autenticazione Cognito e la selezione di un pool di utenti e identità, consulta the section called “Configurazione di un dominio per l'uso dell'autenticazione Amazon Cognito”.
- La seguente policy di accesso al dominio:

  JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
      },
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/{domain-name}/*"
    }
  ]
}
```

- HTTPS richiesto per tutto il traffico verso il dominio
- Node-to-node crittografia
- Crittografia dei dati a riposo

Fase 3: Configurazione degli utenti di Cognito

Durante la creazione del dominio, configura gli utenti master e limitati all'interno di Amazon Cognito seguendo la procedura Crea un pool di utenti nella Amazon Cognito Developer Guide. Infine, configura il tuo pool di identità seguendo i passaggi descritti in Creare un pool di identità in Amazon Cognito. Il bacino d'utenza e il pool di identità devono trovarsi nella stessa Regione AWS.

Fase 4: Mappa i ruoli nelle dashboard OpenSearch

Ora che gli utenti sono configurati, puoi accedere a OpenSearch Dashboards come utente principale e mappare gli utenti ai ruoli.

1. Torna alla console di OpenSearch servizio e vai all'URL delle OpenSearch dashboard per il dominio che hai creato. L'URL segue il seguente formato: *domain-endpoint*`/_dashboards/`.
2. Accedere con le credenziali `master-user`.
3. Scegliere Add sample data (Aggiungi dati di esempio) e aggiungere alcuni dati di volo di esempio.
4. Nel pannello di navigazione a sinistra, seleziona Security (Sicurezza), Roles (Ruoli) e Create role (Crea ruolo).
5. Denomina il ruolo `new-role`.
6. Per Index (Indice), specificare `opensearch_dashboards_sample_data_fli*` (`kibana_sample_data_fli*` sui domini Elasticsearch).
7. Per Index permissions (Autorizzazioni relative all'indice), scegliere read (lettura).
8. Per Sicurezza a livello di documento, specificare la seguente query:

```
{
  "match": {
    "FlightDelay": true
  }
}
```

9. Per la sicurezza a livello di campo, scegliere Escludi e specificare `FlightNum`.
10. Per Anonimizzazione, specificare `Dest`.
11. Scegliere Create (Crea) .
12. Scegliere Utenti mappati, Gestisci mappatura. Aggiungere il nome della risorsa Amazon (ARN) per `LimitedUserRole` come identità esterna e scegliere Map (Mappa).

13. Torna all'elenco di ruoli e scegli opensearch_dashboards_user. Scegliere Utenti mappati, Gestisci mappatura. Aggiungere l'ARN per `LimitedUserRole` come ruolo di back-end e scegliere Mappa.

Fase 5: Test delle autorizzazioni

Quando i ruoli sono mappati correttamente, puoi accedere come utente limitato e testare le autorizzazioni.

1. In una nuova finestra privata del browser, vai all'URL della OpenSearch dashboard per il dominio, accedi utilizzando le `limited-user` credenziali e scegli Esplora da solo.
2. Scegliere Strumenti di sviluppo ed eseguire la ricerca predefinita:

   ```
   GET _search
   {
     "query": {
       "match_all": {}
     }
   }
   ```

   Notare l'errore di autorizzazioni. `limited-user` non dispone delle autorizzazioni per eseguire ricerche a livello di cluster.
3. Esegui un'altra ricerca:

   ```
   GET opensearch_dashboards_sample_data_flights/_search
   {
     "query": {
       "match_all": {}
     }
   }
   ```

   Si noti che tutti i documenti corrispondenti hanno un campo `FlightDelay` di `true`, un campo `Dest` anonimizzato e nessun campo `FlightNum`.
4. Nella finestra del browser originale, accedere come `master-user`, scegliere Strumenti di sviluppo, e quindi eseguire le stesse ricerche. Nota la differenza tra autorizzazioni, numero di hit, documenti corrispondenti e campi inclusi.

Tutorial: Configurare un dominio con il database utente interno e l'autenticazione di base HTTP

Questo tutorial tratta un altro popolare caso d'uso del controllo granulare degli accessi: un utente principale nel database utenti interno e l'autenticazione di base HTTP per le dashboard. OpenSearch L'utente master può quindi accedere alle OpenSearch dashboard, creare un utente interno, mappare l'utente a un ruolo e utilizzare un controllo granulare degli accessi per limitare le autorizzazioni dell'utente.

In questo tutorial completerai le seguenti fasi:

1. Crea un dominio con un utente principale
2. Configura un utente interno nei OpenSearch pannelli di controllo
3. Mappa i ruoli nelle dashboard OpenSearch
4. Test delle autorizzazioni

Fase 1: Creazione di un dominio

Accedi alla console di Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa/ e crea un dominio con le seguenti impostazioni:

- OpenSearch 1.0 o versione successiva oppure Elasticsearch 7.9 o versione successiva
- Accesso pubblico
- Controllo granulare degli accessi con un utente principale nel database utente interno (`TheMasterUser` per il resto di questo tutorial)
- Autenticazione Amazon Cognito per Dashboards disabilitata
- La seguente policy di accesso:

  JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
      },
```

```
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/{domain-name}/*"
    }
  ]
}
```

- HTTPS richiesto per tutto il traffico verso il dominio
- Node-to-node crittografia
- Crittografia dei dati a riposo

Fase 2: Creare un utente interno nei OpenSearch pannelli di controllo

Ora che hai un dominio, puoi accedere alle OpenSearch dashboard e creare un utente interno.

1. Torna alla console di OpenSearch servizio e vai all'URL delle OpenSearch dashboard per il dominio che hai creato. L'URL segue il seguente formato: *domain-endpoint*`/_dashboards/`.
2. Accedi con. `TheMasterUser`
3. Scegliere Add sample data (Aggiungi dati di esempio) e aggiungere alcuni dati di volo di esempio.
4. Nel riquadro di navigazione a sinistra, scegli Sicurezza, Utenti interni, Crea utente interno.
5. Denominare l'utente `new-user` e specificare una password. Quindi, scegli Create (Crea).

Passaggio 3: mappare i ruoli nelle OpenSearch dashboard

Ora che l'utente è configurato, puoi mappare l'utente a un ruolo.

1. Resta nella sezione Sicurezza delle OpenSearch dashboard e scegli Ruoli, Crea ruolo.
2. Denomina il ruolo `new-role`.
3. Per Index, specifica `opensearch_dashboards_sample_data_fli*` (`kibana_sample_data_fli*`sui domini Elasticsearch) il modello di indice.
4. Per il gruppo di operazioni, scegliere read.
5. Per Sicurezza a livello di documento, specificare la seguente query:

```
{
  "match": {
```

```
    "FlightDelay": true
  }
}
```

6. Per la sicurezza a livello di campo, scegliere Escludi e specificare `FlightNum`.
7. Per Anonimizzazione, specificare `Dest`.
8. Scegliere Create (Crea) .
9. Scegliere Utenti mappati, Gestisci mappatura. Quindi aggiungere `new-user` a Utenti e scegliere Mappa.
10. Torna all'elenco di ruoli e scegli opensearch_dashboards_user. Scegliere Utenti mappati, Gestisci mappatura. Quindi aggiungere `new-user` a Utenti e scegliere Mappa.

Passaggio 4: Verifica le autorizzazioni

Quando i ruoli sono mappati correttamente, puoi accedere come utente limitato e testare le autorizzazioni.

1. In una nuova finestra privata del browser, vai all'URL della OpenSearch dashboard per il dominio, accedi utilizzando le `new-user` credenziali e scegli Esplora da solo.
2. Scegliere Strumenti di sviluppo ed eseguire la ricerca predefinita:

```
GET _search
{
  "query": {
    "match_all": {}
  }
}
```

   Notare l'errore di autorizzazioni. `new-user` non dispone delle autorizzazioni per eseguire ricerche a livello di cluster.
3. Esegui un'altra ricerca:

```
GET dashboards_sample_data_flights/_search
{
  "query": {
    "match_all": {}
  }
}
```

Si noti che tutti i documenti corrispondenti hanno un campo `FlightDelay` di `true`, un campo `Dest` anonimizzato e nessun campo `FlightNum`.

4. Nella finestra del browser originale, accedere come `TheMasterUser`, scegliere Strumenti di sviluppo ed eseguire le stesse ricerche. Nota la differenza tra autorizzazioni, numero di hit, documenti corrispondenti e campi inclusi.

Convalida della conformità per Amazon Service OpenSearch

I revisori di terze parti valutano la sicurezza e la conformità del OpenSearch servizio Amazon nell'ambito di diversi programmi di AWS conformità. Questi programmi includono SOC, PCI e HIPAA.

Se hai requisiti di conformità, prendi in considerazione l'utilizzo di qualsiasi versione di Elasticsearch 6.0 OpenSearch o successiva. Le versioni precedenti di Elasticsearch non offrono una combinazione di crittografia dei dati archiviati e node-to-node crittografia ed è improbabile che soddisfino le tue esigenze. Potresti anche prendere in considerazione l'utilizzo di qualsiasi versione di Elasticsearch 6.7 OpenSearch o successiva se il controllo granulare degli accessi è importante per il tuo caso d'uso. Indipendentemente da ciò, la scelta di una versione particolare OpenSearch o di Elasticsearch quando si crea un dominio non garantisce la conformità.

Per sapere se una Servizio AWS rientra nell'ambito di specifici programmi di conformità, consulta Servizi AWS la sezione Ambito per programma di conformitàServizi AWS e scegli il programma di conformità che ti interessa. Per informazioni generali, consulta Programmi di AWS conformità Programmi di di .

È possibile scaricare report di audit di terze parti utilizzando AWS Artifact. Per ulteriori informazioni, consulta Scaricamento dei report in AWS Artifact .

La vostra responsabilità di conformità durante l'utilizzo Servizi AWS è determinata dalla sensibilità dei dati, dagli obiettivi di conformità dell'azienda e dalle leggi e dai regolamenti applicabili. Per ulteriori informazioni sulla responsabilità di conformità durante l'utilizzo Servizi AWS, consulta la Documentazione AWS sulla sicurezza.

Resilienza in Amazon Service OpenSearch

L'infrastruttura AWS globale è costruita attorno a Regioni AWS zone di disponibilità. Regioni AWS forniscono più zone di disponibilità fisicamente separate e isolate, collegate con reti a bassa latenza,

ad alto throughput e altamente ridondanti. Con le zone di disponibilità, è possibile progettare e gestire applicazioni e database che eseguono il failover automatico tra zone di disponibilità senza interruzioni. Le zone di disponibilità sono più disponibili, tolleranti ai guasti e scalabili rispetto alle infrastrutture tradizionali a data center singolo o multiplo.

Per ulteriori informazioni sulle zone di disponibilità, vedere Global Regioni AWS Infrastructure.AWS

Oltre all'infrastruttura AWS globale, OpenSearch Service offre diverse funzionalità per supportare le esigenze di resilienza e backup dei dati:

- Domini AZ multipli e shard di replica
- Snapshot automatiche e manuali

Autenticazione e autorizzazione JWT per Amazon Service OpenSearch

Amazon OpenSearch Service ora consente di utilizzare JSON Web Tokens (JWTs) per l'autenticazione e l'autorizzazione. JWTs sono token di accesso basati su JSON utilizzati per concedere l'accesso Single Sign-On (SSO). È possibile utilizzare JWTs in OpenSearch Service per creare token Single Sign-on per convalidare le richieste al dominio del Servizio. OpenSearch Per utilizzarlo JWTs, è necessario che il controllo granulare degli accessi sia abilitato e fornire una chiave pubblica valida in formato RSA o ECDSA PEM. Per ulteriori informazioni sul controllo granulare degli accessi, consulta Controllo granulare degli accessi in Amazon Service. OpenSearch

Puoi configurare i token Web JSON utilizzando la console di OpenSearch servizio, il () o il. AWS Command Line Interface AWS CLI AWS SDKs

Considerazioni

Prima di JWTs utilizzarlo con Amazon OpenSearch Service devi considerare quanto segue:

- A causa delle dimensioni delle chiavi pubbliche RSA in formato PEM, consigliamo di utilizzare la AWS console per configurare l'autenticazione e l'autorizzazione JWT.
- È necessario fornire utenti e ruoli validi quando si specificano i campi soggetti e ruoli per le proprie richieste JWTs, altrimenti le richieste verranno respinte.
- OpenSearch 2.11 è la prima versione compatibile che può essere utilizzata per l'autenticazione JWT.

Modifica della policy di accesso al dominio

Prima di poter configurare il dominio per utilizzare l'autenticazione e l'autorizzazione JWT, è necessario aggiornare la politica di accesso al dominio per consentire agli utenti JWT di accedere al dominio. In caso contrario, tutte le richieste autorizzate JWT in entrata vengono rifiutate. La politica di accesso al dominio consigliata per fornire l'accesso completo alle risorse secondarie (/*) è:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
      {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
      "Action": "es:ESHttp*",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/domain-name/*"
    }
  ]
}
```

Configurazione dell'autenticazione e dell'autorizzazione JWT

È possibile abilitare l'autenticazione e l'autorizzazione JWT durante il processo di creazione del dominio o aggiornando un dominio esistente. I passaggi di configurazione variano leggermente a seconda dell'opzione scelta.

I passaggi seguenti spiegano come configurare un dominio esistente per l'autenticazione e l'autorizzazione JWT nella console di OpenSearch servizio:

1. In Configurazione del dominio, vai su Autenticazione e autorizzazione JWT per OpenSearch, seleziona Abilita autenticazione e autorizzazione JWT.
2. Configura la chiave pubblica da utilizzare per il tuo dominio. Per fare ciò, puoi caricare un file PEM contenente una chiave pubblica o inserirlo manualmente.

> **Note**
>
> Se la chiave caricata o inserita non è valida, verrà visualizzato un avviso sopra la casella di testo che specifica il problema.

3. (Facoltativo) In Impostazioni aggiuntive, puoi configurare i seguenti campi opzionali

   - Chiave dell'oggetto: puoi lasciare vuoto questo campo per utilizzare la `sub` chiave predefinita per il tuo JWTs.
   - Chiave dei ruoli: puoi lasciare vuoto questo campo per utilizzare la `roles` chiave predefinita per il tuo JWTs.

   Dopo aver apportato le modifiche, salva il dominio.

Utilizzo di un JWT per inviare una richiesta di test

Dopo aver creato un nuovo JWT con una coppia di soggetto e ruolo specificata, puoi inviare una richiesta di test. Per fare ciò, usa la chiave privata per firmare la tua richiesta tramite lo strumento che ha creato il JWT. OpenSearch Il servizio è in grado di convalidare la richiesta in arrivo verificando questa firma.

> **Note**
>
> Se hai specificato una chiave dell'oggetto o una chiave di ruolo personalizzata per il tuo JWT, devi utilizzare i nomi di claim corretti per il tuo JWT.

Di seguito è riportato un esempio di come utilizzare un token JWT per accedere al OpenSearch servizio tramite l'endpoint di ricerca del dominio:

```
curl -XGET "$search_endpoint" -H "Authorization: Bearer <JWT>"
```

Configurazione dell'autenticazione e dell'autorizzazione JWT ()AWS CLI

Il AWS CLI comando seguente abilita l'autenticazione e l'autorizzazione JWT a OpenSearch condizione che il dominio esista:

```
aws opensearch update-domain-config --domain-name <your_domain_name> --advanced-
security-options '{"JWTOptions":{"Enabled":true, "PublicKey": "<your_public_key>",
 "SubjectKey": "<your_subject_key>", "RolesKey": "<your_roles_key>"}}'
```

Configurazione dell'autenticazione e dell'autorizzazione JWT (configurazione tramite API)

La seguente richiesta all'API di configurazione abilita l'autenticazione e l'autorizzazione JWT OpenSearch su un dominio esistente:

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/config
{
  "AdvancedSecurityOptions": {
    "JWTOptions": {
      "Enabled": true,
      "PublicKey": "public-key",
      "RolesKey": "optional-roles-key",
      "SubjectKey": "optional-subject-key"
   }
  }
}
```

Generazione di una key pair

Per eseguire la configurazione JWTs per il tuo OpenSearch dominio, dovrai fornire una chiave pubblica in formato Privacy-Enhanced Mail (PEM). Amazon OpenSearch Service attualmente supporta due algoritmi di crittografia asimmetrica quando si utilizza JWTs: RSA ed ECDSA.

Per creare una coppia di key pair RSA utilizzando la libreria openssl comune, segui questi passaggi:

1. `openssl genrsa -out privatekey.pem 2048`
2. `openssl rsa -in privatekey.pem -pubout -out publickey.pem`

In questo esempio, il `publickey.pem` file contiene la chiave pubblica da utilizzare con Amazon OpenSearch Service, mentre `privatekey.pem` contiene quella privata per la firma dell' JWTs invio al servizio. Inoltre, hai la possibilità di convertire la chiave privata nel `pkcs8` formato comunemente usato se ne hai bisogno per generare il tuo JWTs.

Se utilizzi il pulsante di caricamento per aggiungere un file PEM direttamente alla console, il file deve avere un'`.pem`estensione o altre estensioni di file come `.crt.cert`, o non `.key` sono al momento supportate.

Sicurezza dell'infrastruttura in Amazon OpenSearch Service

In quanto servizio gestito, è protetto dalla sicurezza di rete AWS globale. Per informazioni sui servizi AWS di sicurezza e su come AWS protegge l'infrastruttura, consulta AWS Cloud Security. Per progettare il tuo AWS ambiente utilizzando le migliori pratiche per la sicurezza dell'infrastruttura, vedi Infrastructure Protection in Security Pillar AWS Well-Architected Framework.

Utilizzate chiamate API AWS pubblicate per accedere attraverso la rete. I client devono supportare quanto segue:

- Transport Layer Security (TLS). È richiesto TLS 1.2 ed è consigliato TLS 1.3.
- Suite di cifratura con Perfect Forward Secrecy (PFS), ad esempio Ephemeral Diffie-Hellman (DHE) o Elliptic Curve Ephemeral Diffie-Hellman (ECDHE). La maggior parte dei sistemi moderni, come Java 7 e versioni successive, supporta tali modalità.

Si utilizzano chiamate API AWS pubblicate per accedere all'API OpenSearch di configurazione del servizio tramite la rete. Per configurare la versione TLS minima richiesta da accettare, specificare il valore `TLSSecurityPolicy` nelle opzioni dell'endpoint di dominio:

```
aws opensearch update-domain-config --domain-name my-domain --domain-endpoint-options
 '{"TLSSecurityPolicy": "Policy-Min-TLS-1-2-2019-07"}'
```

Per maggiori dettagli, consulta la Guida di riferimento ai comandi della AWS CLI.

A seconda della configurazione del dominio, potrebbe essere necessario firmare le richieste anche a OpenSearch APIs. Per ulteriori informazioni, consulta the section called “Effettuazione e firma di richieste OpenSearch di assistenza”.

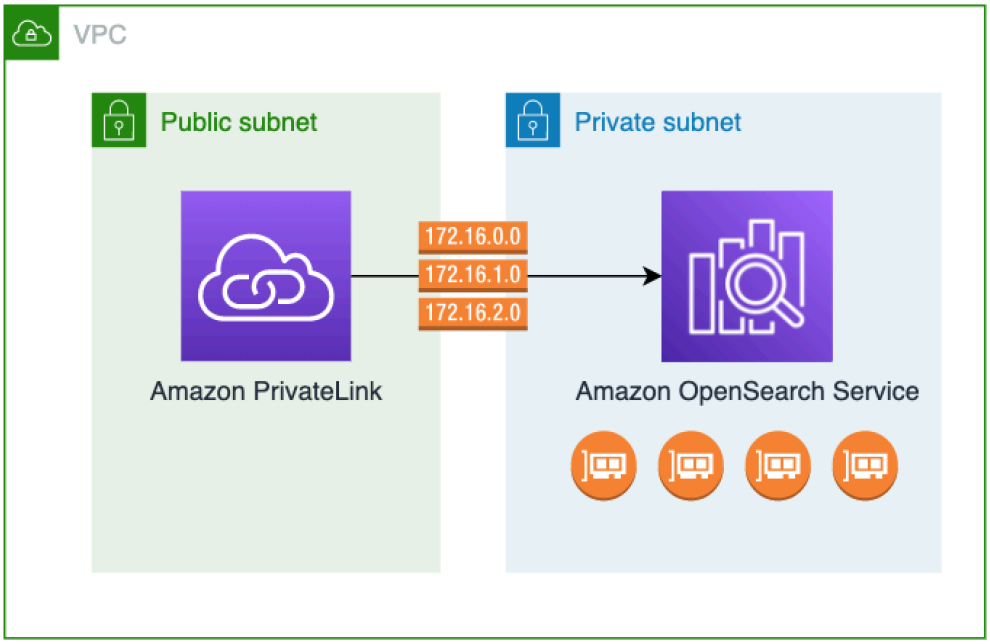
OpenSearch Il servizio supporta i domini di accesso pubblico, che possono ricevere richieste da qualsiasi dispositivo connesso a Internet, e i domini di accesso VPC, che sono isolati dalla rete Internet pubblica.

Accedi ad Amazon OpenSearch Service utilizzando un endpoint OpenSearch VPC gestito da Service ()AWS PrivateLink

Puoi accedere a un dominio Amazon OpenSearch Service configurando un endpoint OpenSearch VPC gestito dal servizio (fornito da). AWS PrivateLink Questi endpoint creano una connessione privata tra il tuo VPC e Amazon OpenSearch Service. Puoi accedere ai domini OpenSearch Service

VPC come se fossero nel tuo VPC, senza l'uso di un gateway Internet, un dispositivo NAT, una connessione VPN o una connessione. Direct Connect Le istanze nel tuo VPC non necessitano di indirizzi IP pubblici per accedere al OpenSearch servizio.

Puoi configurare i domini OpenSearch di servizio per esporre endpoint aggiuntivi in esecuzione su sottoreti pubbliche o private all'interno dello stesso VPC, VPC diverso o diverso. Account AWS Ciò consente di aggiungere un ulteriore livello di sicurezza per accedere ai domini indipendentemente da dove vengono eseguiti, senza alcuna infrastruttura da gestire. Il diagramma seguente illustra gli endpoint VPC OpenSearch gestiti dal servizio all'interno dello stesso VPC:



Si stabilisce questa connessione privata creando un endpoint VPC OpenSearch con interfaccia gestita dal servizio, alimentato da. AWS PrivateLink In ciascuna sottorete viene creata un'interfaccia di rete endpoint da abilitare per l'endpoint VPC dell'interfaccia. Si tratta di interfacce di rete gestite dal servizio che fungono da punto di ingresso per il traffico destinato al Servizio. OpenSearch I prezzi degli endpoint con AWS PrivateLink interfaccia standard si applicano agli endpoint VPC gestiti dal OpenSearch servizio fatturati in base al prezzo. AWS PrivateLink

Puoi creare endpoint VPC per domini che eseguono tutte le versioni precedenti di OpenSearch Elasticsearch. Per ulteriori informazioni, consulta la sezione Accesso a Servizi AWS tramite AWS PrivateLink nella Guida di AWS PrivateLink .

Considerazioni e limitazioni per il servizio OpenSearch

Prima di configurare un endpoint VPC di interfaccia per il OpenSearch servizio, consulta Accedere a un AWS servizio utilizzando un endpoint VPC di interfaccia nella Guida.AWS PrivateLink

Quando utilizzi gli endpoint OpenSearch VPC gestiti dal servizio, considera quanto segue:

- È possibile utilizzare solo gli endpoint VPC dell'interfaccia per connettersi ai domini VPC. I domini pubblici non sono supportati.
- Gli endpoint VPC possono connettersi solo ai domini all'interno della stessa Regione AWS.
- HTTPS è l'unico protocollo supportato per gli endpoint VPC. HTTP non è consentito.
- OpenSearch Il servizio supporta l'effettuazione di chiamate a tutte le operazioni OpenSearch API supportate tramite un endpoint VPC di interfaccia.
- Puoi configurare un massimo di 50 endpoint per account e un massimo di 10 endpoint per dominio. Un singolo dominio può avere un massimo di 10 principali autorizzati.
- Al momento non è possibile utilizzare AWS CloudFormation per creare endpoint VPC di interfaccia.
- Puoi creare endpoint VPC di interfaccia solo tramite la console di OpenSearch servizio o utilizzando l'OpenSearch API di servizio. Non puoi creare endpoint VPC di interfaccia per OpenSearch Service utilizzando la console Amazon VPC.
- OpenSearch Gli endpoint VPC gestiti dal servizio non sono accessibili da Internet. Un endpoint OpenSearch VPC gestito dai servizi è accessibile solo all'interno del VPC in cui viene fornito l'endpoint o tramite peering VPCs con il VPC in cui viene fornito l'endpoint, come consentito dalle tabelle di routing e dai gruppi di sicurezza.
- Le policy degli endpoint VPC non sono supportate per Service. OpenSearch È possibile associare un gruppo di sicurezza alle interfacce di rete degli endpoint per controllare il traffico verso il OpenSearch servizio tramite l'interfaccia VPC endpoint.
- Il tuo ruolo collegato al servizio deve trovarsi nello stesso AWS account che usi per creare l'endpoint VPC.
- Per creare, aggiornare ed eliminare l'endpoint OpenSearch Service VPC, devi disporre delle seguenti autorizzazioni Amazon EC2 oltre alle autorizzazioni Amazon Service: OpenSearch
  - `ec2:CreateVpcEndpoint`
  - `ec2:DescribeVpcEndpoints`
  - `ec2:ModifyVpcEndpoint`
  - `ec2:DeleteVpcEndpoints`
  - `ec2:CreateTags`
  - `ec2:DescribeTags`
  - `ec2:DescribeSubnets`
  - `ec2:DescribeSecurityGroups`

- `ec2:DescribeVpcs`

> **Note**
>
> Al momento, non puoi limitare la creazione di endpoint VPC a Service. OpenSearch Stiamo lavorando per renderlo possibile in un futuro aggiornamento.

Fornitura dell'accesso a un dominio

Se il VPC a cui desideri accedere al tuo dominio si trova in un altro Account AWS, devi autorizzarlo dall'account del proprietario prima di poter creare un endpoint VPC di interfaccia.

Per consentire a un VPC di un altro di accedere Account AWS al tuo dominio

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa/.
2. Nel pannello di navigazione, scegli Domains (Domini) e apri il dominio a cui desideri fornire l'accesso.
3. Vai alla scheda Endpoint VPC, che mostra gli account e i corrispondenti VPCs che hanno accesso al tuo dominio.
4. Scegli Authorize principal (Autorizza principale).
5. Inserisci l' Account AWS ID dell'account che accederà al tuo dominio. In questa fase autorizzi l'account specificato a creare endpoint VPC sul dominio.
6. Seleziona Authorize (Autorizza).

Creazione di un endpoint VPC dell'interfaccia per un dominio VPC

È possibile creare un endpoint VPC di interfaccia per OpenSearch Service utilizzando la console di OpenSearch servizio o (). AWS Command Line Interface AWS CLI

Per creare un endpoint VPC di interfaccia per un dominio di servizio OpenSearch

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa/.
2. Nel pannello di navigazione a sinistra, scegli Endpoint VPC.
3. Seleziona Crea endpoint.
4. Seleziona se connettere un dominio corrente Account AWS o un altro Account AWS.

5. Seleziona il dominio a cui ti connetti con questo endpoint. Se il dominio è nell'attuale Account AWS, usa il menu a discesa per scegliere il dominio. Se il dominio si trova in un altro account, immetti il nome della risorsa Amazon (ARN) del dominio a cui connettersi. Per scegliere un dominio in un altro account, il proprietario deve fornirti l'accesso al dominio.
6. Per VPC, seleziona il VPC da cui accederai a Service. OpenSearch
7. Per le sottoreti, seleziona una o più sottoreti da cui accedere al servizio. OpenSearch
8. Per Security groups (Gruppi di sicurezza), seleziona i gruppi di sicurezza da associare alle interfacce di rete dell'endpoint. Si tratta di una fase fondamentale per limitare le porte, i protocolli e le origini per il traffico in ingresso che si sta autorizzando per l'endpoint. Le regole del gruppo di sicurezza devono consentire alle risorse che utilizzeranno l'endpoint VPC di comunicare con il OpenSearch Servizio di comunicare con l'interfaccia di rete dell'endpoint.
9. Seleziona Crea endpoint. L'endpoint dovrebbe essere attivo entro 2-5 minuti.

Utilizzo degli endpoint OpenSearch VPC gestiti dal servizio utilizzando l'API di configurazione

Utilizza le seguenti operazioni API per creare e gestire endpoint OpenSearch VPC gestiti dal servizio.

- CreateVpcEndpoint
- ListVpcEndpoints
- UpdateVpcEndpoint
- DeleteVpcEndpoint

Utilizza le seguenti operazioni API per gestire l'accesso degli endpoint ai domini VPC:

- AuthorizeVpcEndpointAccess
- ListVpcEndpointAccess
- ListVpcEndpointsForDomain
- RevokeVpcEndpointAccess

Autenticazione SAML per dashboard OpenSearch

L'autenticazione SAML per OpenSearch dashboard ti consente di utilizzare il tuo provider di identità esistente per offrire Single Sign-On (SSO) per dashboard su domini Amazon OpenSearch Service

in esecuzione o Elasticsearch 6.7 o versione successiva. OpenSearch Per utilizzare l'autenticazione SAML, è necessario abilitare il controllo granulare degli accessi.

Invece di eseguire l'autenticazione tramite Amazon Cognito o il database utenti interno, l'autenticazione SAML OpenSearch per dashboard consente di utilizzare provider di identità di terze parti per accedere alle dashboard, gestire il controllo granulare degli accessi, cercare i dati e creare visualizzazioni. OpenSearch Il servizio supporta provider che utilizzano lo standard SAML 2.0, come Okta, Keycloak, Active Directory Federation Services (ADFS), Auth0 e. AWS IAM Identity Center

L'autenticazione SAML per le dashboard consente solo l'accesso alle dashboard tramite un browser Web. OpenSearch Le tue credenziali SAML non ti consentono di effettuare richieste HTTP dirette ai dashboard o. OpenSearch APIs

Panoramica della configurazione SAML

Questa documentazione presuppone che si disponga di un fornitore di identità esistente e che si abbia una certa familiarità con esso. Non possiamo fornire passaggi di configurazione dettagliati per il tuo provider esatto, ma solo per il tuo OpenSearch dominio di servizio.

Il flusso di accesso a OpenSearch Dashboards può assumere una delle due forme seguenti:

- Avviato da provider di servizi (SP): si passa a Dashboards (ad esempio `https://my-domain.us-east-1.es.amazonaws.com/_dashboards`), che reindirizza alla schermata di accesso. Dopo aver effettuato l'accesso, il provider di identità reindirizza l'utente da Dashboards.
- Provider di identità (IdP) avviato: accedi al tuo provider di identità, accedi e scegli OpenSearch Dashboard da una directory di applicazioni.

OpenSearch Il servizio offre due Single Sign-On URLs, avviato da SP e avviato da IdP, ma è necessario solo quello che corrisponde al flusso di accesso desiderato per Dashboards. OpenSearch

Indipendentemente dal tipo di autenticazione utilizzato, l'obiettivo è quello di accedere tramite il provider di identità e ricevere un'asserzione SAML contenente il nome utente (obbligatorio) e qualsiasi ruolo di back-end (facoltativo, ma consigliato). Queste informazioni consentono il controllo granulare degli accessi per assegnare le autorizzazioni agli utenti SAML. Nei provider di identità esterni, i ruoli back-end sono in genere denominati "ruoli" o "gruppi".

Considerazioni

Durante la configurazione dell'autenticazione SAML tieni presente quanto segue:

- A causa delle dimensioni del file di metadati IdP, consigliamo vivamente di utilizzare la AWS console per configurare l'autenticazione SAML.
- I domini supportano solo un metodo di autenticazione Dashboards alla volta. Se hai abilitato l'autenticazione Amazon Cognito per OpenSearch dashboard, devi disabilitarla prima di poter abilitare l'autenticazione SAML.
- Se utilizzi un sistema di bilanciamento del carico di rete con SAML, devi prima creare un endpoint personalizzato. Per ulteriori informazioni, consulta ???.
- Le policy di controllo dei servizi (SCP) non saranno applicabili o valutate nel caso di identità non IAM (come SAML in Amazon OpenSearch Serverless e SAML e l'autorizzazione utente interna di base per Amazon Service). OpenSearch

Autenticazione SAML per domini VPC

SAML non richiede una comunicazione diretta tra il fornitore di identità e il provider di servizi. Pertanto, anche se il tuo OpenSearch dominio è ospitato all'interno di un VPC privato, puoi comunque utilizzare SAML purché il tuo browser sia in grado di comunicare sia con il tuo OpenSearch cluster che con il tuo provider di identità. Il browser agisce essenzialmente come intermediario tra il provider di identità e il provider di servizi. Per un diagramma utile che spiega il flusso di autenticazione SAML, consulta la Documentazione Okta.

Modifica della policy di accesso al dominio

Prima di configurare l'autenticazione SAML è necessario aggiornare la policy di accesso al dominio per consentire agli utenti SAML di accedervi. Altrimenti, saranno restituiti errori di accesso negato.

Consigliamo la seguente policy di accesso al dominio, che fornisce l'accesso completo alle risorse secondarie (`/*`) del dominio:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
```

```
      "Action": "es:ESHttp*",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/domain-name/*"
    }
  ]
}
```

Per rendere la politica più restrittiva, puoi aggiungere una condizione relativa all'indirizzo IP alla politica. Questa condizione limita l'accesso solo all'intervallo di indirizzi IP o alla sottorete specificati. Ad esempio, la seguente politica consente l'accesso solo dalla sottorete 192.0.2.0/24:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
      "Action": [
        "es:ESHttp*"
      ],
      "Condition": {
        "IpAddress": {
          "aws:SourceIp": [
            "192.0.2.0/24"
          ]
        }
      },
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/domain-name/*"
    }
  ]
}
```

> **Note**
>
> Una politica di accesso al dominio aperta richiede l'attivazione di un controllo granulare degli accessi sul dominio, altrimenti viene visualizzato il seguente errore:

```
To protect domains with public access, a restrictive policy or fine-
grained access control is required.
```

Se hai un utente principale o un utente interno configurato con una password robusta, mantenere aperta la policy utilizzando un controllo granulare degli accessi potrebbe essere accettabile dal punto di vista della sicurezza. Per ulteriori informazioni, consulta ???.

Configurazione dell'autenticazione avviata da SP o da IdP

Questi passaggi spiegano come abilitare l'autenticazione SAML con l'autenticazione avviata da SP o IdP per i dashboard. OpenSearch Per il passaggio aggiuntivo richiesto per abilitare entrambi, consulta Configurazione dell'autenticazione avviata da SP e avviata da IdP.

Fase 1: Abilitazione dell'autenticazione SAML

Puoi abilitare l'autenticazione SAML durante la creazione del dominio o scegliendo Actions (Operazioni), Edit security configuration (Modifica configurazione di sicurezza) su un dominio esistente. I seguenti passaggi variano leggermente a seconda della scelta effettuata.

All'interno della configurazione del dominio, in Autenticazione SAML per Dashboards/Kibana, seleziona Abilita l'autenticazione SAML. OpenSearch

Fase 2: Configurazione del fornitore di identità

Esegui i seguenti passaggi a seconda di quando viene configurata l'autenticazione SAML.

Se stai creando un nuovo dominio

Se stai creando un nuovo dominio, Service non è ancora in grado di generare un ID di entità OpenSearch del fornitore di servizi o SSO. URLs Il tuo fornitore di identità richiede questi valori per abilitare correttamente l'autenticazione SAML, ma possono essere generati solo dopo la creazione del dominio. Per ovviare a questa interdipendenza durante la creazione del dominio, puoi fornire valori temporanei nella configurazione dell'IdP per generare i metadati richiesti e quindi aggiornarli una volta che il dominio è attivo.

Se utilizzi un endpoint personalizzato, puoi dedurre quale sarà. URLs Ad esempio, se l'endpoint personalizzato è `www.custom-endpoint.com`, l'ID dell'entità del fornitore di servizi sarà `www.custom-endpoint.com`, l'URL SSO avviato dall'IDP sarà `www.custom-endpoint.com/_dashboards/_opendistro/_security/saml/acs/idpinitiated` e l'URL SSO avviato da SP sarà `www.custom-endpoint.com/_dashboards/_opendistro/_security/saml/acs`.

Puoi utilizzare i valori per configurare il tuo fornitore di identità prima della creazione del dominio. Per gli esempi consultare la prossima sezione.

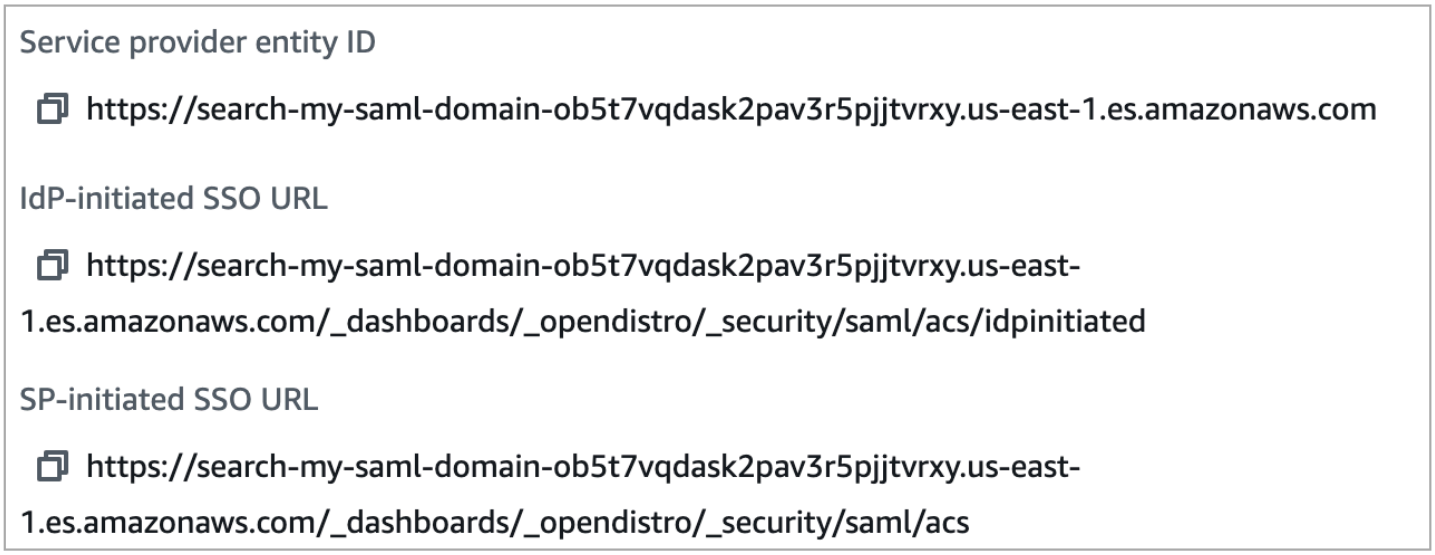
> **Note**
>
> Non puoi accedere con un endpoint dual stack perché il nome di dominio completo di una richiesta HTTP è diverso dal nome di dominio completo di una richiesta SAML. Un OpenSearch amministratore dovrà configurare un endpoint personalizzato e impostare il valore CNAME su un endpoint dual stack se desideri accedere utilizzando un endpoint dual stack.

Se non utilizzi un endpoint personalizzato, puoi inserire valori temporanei nel tuo IdP per generare i metadati richiesti e quindi aggiornarli in un secondo momento dopo l'attivazione del dominio.

Ad esempio, all'interno di Okta, puoi inserire `https://temp-endpoint.amazonaws.com` nei campi Single sign on URL (URL Single Sign On) e Audience URI (SP Entity ID) (URI del pubblico [ID entità SP]), che consentono di generare i metadati. Quindi, dopo che il dominio è attivo, puoi recuperare i valori corretti da OpenSearch Service e aggiornarli in Okta. Per istruzioni, consulta the section called “Passaggio 6: aggiorna il tuo IdP URLs”.

Se stai modificando un dominio esistente

Se stai abilitando l'autenticazione SAML su un dominio esistente, copia l'ID dell'entità del fornitore di servizi e uno degli SSO. URLs Per indicazioni sull'URL da utilizzare, consulta the section called “Panoramica della configurazione SAML”.



Utilizzare questi valori per configurare il fornitore di identità. Questa è la parte più complessa del processo e, sfortunatamente, la terminologia e i passaggi variano notevolmente a seconda del provider. Consultare la documentazione del provider.

In Okta, ad esempio, crei un'applicazione web SAML 2.0. Per Single sign on URL (URL Single Sign-On), specifica l'URL SSO. Per URI di destinazione (ID entità SP), specificare l'ID entità SP.

Piuttosto che utenti e ruoli di back-end, Okta utilizza utenti e gruppi. Per Group Attribute Statements (Istruzioni degli attributi di gruppo), consigliamo di aggiungere `role` al campo Name (Nome) e l'espressione regolare `.+` al campo Filter (Filtro). Questa istruzione indica al provider di identità Okta di includere tutti i gruppi di utenti sotto il campo `role` dell'asserzione SAML dopo l'autenticazione di un utente.

In IAM Identity Center, si specifica l'ID dell'entità SP come audience SAML dell'applicazione. È inoltre necessario specificare le seguenti mappature degli attributi: e. `Subject=${user:subject}:format=unspecified` `Role=${user:groups}:format=uri`

In Auth0, crei una normale applicazione web e abiliti il componente aggiuntivo SAML 2.0. In Keycloak, crei un client.

Fase 3: Importazione dei metadati dell'IdP

Dopo aver configurato il provider di identità, viene generato un file di metadati IdP. Questo file XML contiene informazioni sul provider, ad esempio un certificato TLS, endpoint Single Sign-On e l'ID entità del provider di identità.

Copia il contenuto del file di metadati IdP e incollalo nel campo Metadati da IdP della console di servizio. OpenSearch In alternativa, scegliere Importa da file XML e caricare il file. Il file dei metadati dovrebbe avere un aspetto simile al seguente:

```
<?xml version="1.0" encoding="UTF-8"?>
<md:EntityDescriptor entityID="entity-id"
 xmlns:md="urn:oasis:names:tc:SAML:2.0:metadata">
  <md:IDPSSODescriptor WantAuthnRequestsSigned="false"
 protocolSupportEnumeration="urn:oasis:names:tc:SAML:2.0:protocol">
    <md:KeyDescriptor use="signing">
      <ds:KeyInfo xmlns:ds="http://www.w3.org/2000/09/xmldsig#">
        <ds:X509Data>
          <ds:X509Certificate>tls-certificate</ds:X509Certificate>
        </ds:X509Data>
```

```
      </ds:KeyInfo>
    </md:KeyDescriptor>
    <md:NameIDFormat>urn:oasis:names:tc:SAML:1.1:nameid-format:unspecified</
md:NameIDFormat>
    <md:NameIDFormat>urn:oasis:names:tc:SAML:1.1:nameid-format:emailAddress</
md:NameIDFormat>
    <md:SingleSignOnService Binding="urn:oasis:names:tc:SAML:2.0:bindings:HTTP-POST"
 Location="idp-sso-url"/>
    <md:SingleSignOnService Binding="urn:oasis:names:tc:SAML:2.0:bindings:HTTP-
Redirect" Location="idp-sso-url"/>
  </md:IDPSSODescriptor>
</md:EntityDescriptor>
```

Fase 4: Configurazione dei campi SAML

Dopo aver inserito i metadati IdP, configura i seguenti campi aggiuntivi nella OpenSearch console di servizio:

- IdP entity ID (ID entità IdP): copiare il valore della proprietà `entityID` dal file di metadati e incollarlo in questo campo. Molti provider di identità visualizzano questo valore anche come parte di un riepilogo successivo alla configurazione. Alcuni fornitori lo chiamano "emittente".
- Nome utente principale SAML e ruolo di backend principale SAML: il ruolo utente di and/or backend specificato riceve le autorizzazioni complete per il cluster, equivalenti a quelle di un nuovo utente master, ma può utilizzare tali autorizzazioni solo all'interno delle dashboard. OpenSearch

  Ad esempio, in Okta è possibile avere un utente `jdoe` che appartiene al gruppo `admins`. Se si aggiunge `jdoe` al campo Nome utente master SAML, solo quell'utente riceverà le autorizzazioni complete. Se si aggiunge `admins` al campo SAML master backend role (Ruolo back-end master SAML), qualsiasi utente che appartiene al gruppo `admins` riceverà le autorizzazioni complete.

> **Note**
>
> Il contenuto dell'asserzione SAML deve corrispondere esattamente alle stringhe utilizzate per il nome utente master SAML e/o il ruolo di master SAML. Alcuni provider di identità aggiungono un prefisso prima dei nomi utente, il che può causare una mancata corrispondenza. hard-to-diagnose Nell'interfaccia utente del provider di identità, è possibile che venga visualizzato `jdoe`, ma l'asserzione SAML potrebbe contenere `auth0|jdoe`. Utilizzare sempre la stringa dall'asserzione SAML.

Molti provider di identità consentono di visualizzare un'asserzione di esempio durante il processo di configurazione e strumenti come Tracer SAML possono aiutare a esaminare e risolvere i problemi del contenuto di asserzioni reali. Le asserzioni hanno il seguente aspetto:

```
<?xml version="1.0" encoding="UTF-8"?>
<saml2:Assertion ID="id67229299299259351343340162"
 IssueInstant="2020-09-22T22:03:08.633Z" Version="2.0"
  xmlns:saml2="urn:oasis:names:tc:SAML:2.0:assertion">
  <saml2:Issuer Format="urn:oasis:names:tc:SAML:2.0:nameid-format:entity">idp-issuer</
saml2:Issuer>
  <saml2:Subject>
    <saml2:NameID Format="urn:oasis:names:tc:SAML:1.1:nameid-
format:unspecified">username</saml2:NameID>
    <saml2:SubjectConfirmation Method="urn:oasis:names:tc:SAML:2.0:cm:bearer">
      <saml2:SubjectConfirmationData NotOnOrAfter="2020-09-22T22:08:08.816Z"
 Recipient="domain-endpoint/_dashboards/_opendistro/_security/saml/acs"/>
    </saml2:SubjectConfirmation>
  </saml2:Subject>
  <saml2:Conditions NotBefore="2020-09-22T21:58:08.816Z"
 NotOnOrAfter="2020-09-22T22:08:08.816Z">
    <saml2:AudienceRestriction>
      <saml2:Audience>domain-endpoint</saml2:Audience>
    </saml2:AudienceRestriction>
  </saml2:Conditions>
  <saml2:AuthnStatement AuthnInstant="2020-09-22T19:54:37.274Z">
    <saml2:AuthnContext>

 <saml2:AuthnContextClassRef>urn:oasis:names:tc:SAML:2.0:ac:classes:PasswordProtectedTransport<
saml2:AuthnContextClassRef>
    </saml2:AuthnContext>
  </saml2:AuthnStatement>
  <saml2:AttributeStatement>
    <saml2:Attribute Name="role" NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-
format:unspecified">
      <saml2:AttributeValue
        xmlns:xs="http://www.w3.org/2001/XMLSchema"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:type="xs:string">GroupName Match Matches regex ".+" (case-sensitive)
      </saml2:AttributeValue>
    </saml2:Attribute>
  </saml2:AttributeStatement>
</saml2:Assertion>
```

Fase 5: (facoltativo) configurazione delle impostazioni aggiuntive

In Additional settings (Impostazioni aggiuntive), configura i seguenti campi facoltativi:

- Subject key (Chiave oggetto): è possibile lasciare questo campo vuoto in modo da utilizzare l'elemento `NameID` dell'asserzione SAML per il nome utente. Se l'asserzione non utilizza questo elemento standard e include invece il nome utente come attributo personalizzato, specificare tale attributo qui.
- Roles key)Chiave ruoli): Se si desidera utilizzare i ruoli back-end (scelta consigliata), specificare un attributo dall'asserzione in questo campo, ad esempio `role` o `group`. Questa è un'altra situazione in cui strumenti come i tracer SAML possono aiutare.
- Durata della sessione: per impostazione predefinita, OpenSearch Dashboards disconnette gli utenti dopo 24 ore. È possibile configurare questo valore su qualsiasi numero compreso tra 60 e 1.440 (24 ore) specificando un nuovo valore.

Se la configurazione ti soddisfa, salva il dominio.

Passaggio 6: aggiorna il tuo IdP URLs

Se hai abilitato l'autenticazione SAML durante la creazione di un dominio, dovevi specificare il valore temporaneo URLs all'interno del tuo IdP per generare il file di metadati XML. Dopo che lo stato del dominio è cambiato in`Active`, puoi ottenere il tuo URLs IdP corretto e modificarlo.

Per recuperarlo URLs, seleziona il dominio e scegli Azioni, Modifica configurazione di sicurezza. In Autenticazione SAML per OpenSearch Dashboards/Kibana, puoi trovare l'ID dell'entità del fornitore di servizi e l'SSO corretti. URLs Copia i valori e usali per configurare il tuo provider di identità, sostituendo il valore temporaneo fornito nel passaggio 2. URLs

Fase 7: Associazione degli utenti SAML ai ruoli

Una volta che lo stato del dominio è Attivo e il tuo IdP è configurato correttamente, accedi a OpenSearch Dashboards.

- Se è stato scelto l'URL avviato da SP, passare a *`domain-endpoint`*`/_dashboards`. Per accedere direttamente a un tenant specifico, aggiungere `?security_tenant=`*`tenant-name`* all'URL.
- Se è stato scelto l'URL avviato dall'IdP, passare alla directory dell'applicazione del provider di identità.

In entrambi i casi, accedere come utente principale SAML o come utente appartenente al ruolo back-end principale SAML. Per continuare l'esempio dalla fase 7, accedere come `jdoe` o come membro del gruppo `admins`.

Dopo il caricamento di OpenSearch Dashboards, scegli Sicurezza, Ruoli. Quindi, mappa i ruoli per consentire ad altri utenti di accedere alle OpenSearch dashboard.

Ad esempio, è possibile mappare il collega fidato `jroe` ai ruoli `all_access` e `security_manager`. È inoltre possibile mappare il ruolo back-end `analysts` ai ruoli `readall` e `opensearch_dashboards_user`.

Se preferisci utilizzare l'API anziché le OpenSearch dashboard, consulta la seguente richiesta di esempio:

```
PATCH _plugins/_security/api/rolesmapping
[
  {
    "op": "add", "path": "/security_manager", "value": { "users": ["master-user",
 "jdoe", "jroe"], "backend_roles": ["admins"] }
  },
  {
    "op": "add", "path": "/all_access", "value": { "users": ["master-user", "jdoe",
 "jroe"], "backend_roles": ["admins"] }
  },
  {
    "op": "add", "path": "/readall", "value": { "backend_roles": ["analysts"] }
  },
  {
    "op": "add", "path": "/opensearch_dashboards_user", "value": { "backend_roles":
 ["analysts"] }
  }
]
```

Configurazione dell'autenticazione avviata da SP e avviata da IdP

Se si desidera configurare sia l'autenticazione avviata da SP che quella avviata da IdP, è necessario farlo tramite il provider di identità. Ad esempio, in Okta, puoi eseguire le operazioni seguenti:

1. All'interno dell'applicazione SAML, vai su General (Generale), SAML settings (Impostazioni SAML).

2. Per Single sign on URL (URL Single Sign On), fornisci l'URL SSO avviato da IdP. Ad esempio, `https://search-domain-hash/_dashboards/_opendistro/_security/saml/acs/idpinitiated`.
3. Abilita Consenti a questa app di richiedere altri URLs SSO.
4. In Requestable SSO URLs, aggiungi uno o più SSO avviati dall'SP. URLs Ad esempio, `https://search-domain-hash/_dashboards/_opendistro/_security/saml/acs`.

Configurazione dell'autenticazione SAML (AWS CLI)

Il AWS CLI comando seguente abilita l'autenticazione SAML per OpenSearch i dashboard su un dominio esistente:

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --advanced-security-options '{"SAMLOptions":{"Enabled":true,"MasterUserName":"my-idp-user","MasterBackendRole":"my-idp-group-or-role","Idp":{"EntityId":"entity-id","MetadataContent":"metadata-content-with-quotes-escaped"},"RolesKey":"optional-roles-key","SessionTimeoutMinutes":180,"SubjectKey":"optional-subject-key"}}'
```

È necessario eseguire l'escape di tutte le virgolette e i caratteri newline nel file XML dei metadati. Ad esempio, utilizzare `<KeyDescriptor use=\"signing\">\n` invece di `<KeyDescriptor use="signing">` e un'interruzione di riga. Per informazioni dettagliate sull'utilizzo della AWS CLI, consultare Riferimento ai comandi AWS CLI.

Configurazione dell'autenticazione SAML (API di configurazione)

La seguente richiesta all'API di configurazione abilita l'autenticazione SAML per le OpenSearch dashboard su un dominio esistente:

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/config
{
  "AdvancedSecurityOptions": {
    "SAMLOptions": {
      "Enabled": true,
      "MasterUserName": "my-idp-user",
      "MasterBackendRole": "my-idp-group-or-role",
      "Idp": {
        "EntityId": "entity-id",
        "MetadataContent": "metadata-content-with-quotes-escaped"
```

```
      },
      "RolesKey": "optional-roles-key",
      "SessionTimeoutMinutes": 180,
      "SubjectKey": "optional-subject-key"
    }
  }
}
```

È necessario eseguire l'escape di tutte le virgolette e i caratteri newline nel file XML dei metadati. Ad esempio, utilizzare `<KeyDescriptor use=\"signing\">\n` invece di `<KeyDescriptor use="signing">` e un'interruzione di riga. Per informazioni dettagliate sull'utilizzo dell'API di configurazione, consulta il riferimento all'API OpenSearch di servizio.

Risoluzione dei problemi SAML

Errore	Informazioni
La tua richiesta: '*/some/path* ' non è consentita.	Verificare di aver fornito al provider di identità l'URL SSO corretto (fase 3).
Fornisci un documento valido per i metadati del provider di identità per abilitare SAML.	Il file di metadati IdP non è conforme allo standard SAML 2.0. Verificare la presenza di errori utilizzando uno strumento di convalida.
Le opzioni di configurazione SAML non sono visibili nella console.	Effettuare l'aggiornamento alla versione più recente del software di assistenza.
Errore di configurazione SAML: si è verificato un errore durante il recupero della configurazione SAML, controlla le impostazioni.	Questo errore generico può verificarsi per molti motivi. • Verificare di aver fornito al provider di identità l'ID entità SP e l'URL SSO corretti. • Rigenerare il file di metadati IdP e verificare l'ID entità IdP. Aggiungere tutti i metadati aggiornati nella console AWS . • Verifica che la politica di accesso al dominio consenta l'accesso alle OpenSearch dashboard e_plugins/_security/* . In generale, per domini che utilizzano un controllo granulare degli

Errore	Informazioni
	accessi è preferibile utilizzare una policy di accesso aperta. • Consultare la documentazione del provider di identità per le istruzioni sulla configurazione di SAML.
Ruolo mancante: nessun ruolo disponibile per questo utente, contatta l'amministratore di sistema.	L'autenticazione è stata eseguita correttamente, ma il nome utente e gli eventuali ruoli back-end dell'asserzione SAML non sono mappati ad alcun ruolo e quindi non dispongono di autorizzazioni. Queste mappature distinguono tra maiuscole e minuscole. L'amministratore di sistema può verificare il contenuto dell'asserzione SAML utilizzando uno strumento come SAML-Tracer, quindi verificare la mappatura dei ruoli utilizzando la seguente richiesta: `GET _plugins/_security/api/rolesmapping`
Il browser reindirizza o riceve continuamente errori HTTP 500 quando tenta di accedere alle dashboard. OpenSearch	Questi errori possono verificarsi se l'asserzione SAML contiene un numero elevato di ruoli per un totale di circa 1.500 caratteri. Ad esempio, se si superano 80 ruoli, la cui lunghezza media è di 20 caratteri, è possibile che il limite di dimensione per i cookie nel browser Web venga superato. A partire dalla OpenSearch versione 2.7, l'asserzione SAML supporta ruoli fino a 5000 caratteri.
Non è possibile disconnettersi da ADFS.	ADFS richiede che tutte le richieste di disconnessione siano firmate, cosa che OpenSearch Service non supporta. Rimuovi `<SingleLogoutService />` dal file di metadati IdP per forzare il OpenSearch servizio a utilizzare il proprio meccanismo di disconnessione interno.

Errore	Informazioni
`Could not find entity descriptor for __PATH__.`	L'ID dell'entità dell'IdP fornito nei metadati XML to OpenSearch Service è diverso da quello nella risposta SAML. Per risolvere questo problema, assicurati che corrispondano. Abilita i log degli errori dell'appl icazione CW sul tuo dominio per trovare il messaggio di errore per il debug del problema di integrazione SAML.
`Signature validation failed. SAML response rejected.`	OpenSearch Il servizio non è in grado di verificare la firma nella risposta SAML utilizzando il certificato dell'IdP fornito nei metadati XML. Potrebbe trattarsi di un errore manuale o il certificato del tuo IdP potrebbe aver modificato il certificato. Aggiorna il certificato più recente del tuo IdP nei metadati XML forniti al OpenSearch Servizio tramite. Console di gestione AWS
`__PATH__ is not a valid audience for this response.`	Il campo audience nella risposta SAML non corrispon de all'endpoint del dominio. Per correggere questo errore, aggiorna il campo SP audience in modo che corrisponda all'endpoint del tuo dominio. Se hai abilitato gli endpoint personalizzati, il campo audience deve corrispondere al tuo endpoint personalizzato. Abilita i registri degli errori dell'applicazione CW sul tuo dominio per trovare il messaggio di errore per eseguire il debug del problema di integrazione SAML.
Il tuo browser riceve un errore HTTP 400 nella risposta. `Invalid Request Id`	Questo errore si verifica in genere se hai configura to l'URL avviato dall'IdP con il formato. *<DashboardsURL>* `/_opendistro/_security/saml/acs` Configura invece l'URL con il formato. *<DashboardsURL>* `/_opendistro/_security/saml/acs/idpinitiated`

Errore	Informazioni
La risposta è stata ricevuta al __PATH__ posto di__PATH__.	Il campo di destinazione nella risposta SAML non corrisponde a uno dei seguenti formati URL: • *<DashboardsURL>* /_opendistro/_security/saml/acs • *<DashboardsURL>* /_opendistro/_security/saml/acs/idpinitiated . A seconda del flusso di accesso utilizzato (avviato da SP o avviato da IDP), inserisci un campo di destinazione che corrisponda a uno dei seguenti. OpenSearch URLs
La risposta ha un `InResponseTo` attributo, mentre non era previsto. `InResponseTo`	Stai utilizzando l'URL avviato dall'IdP per un flusso di accesso avviato da SP. Utilizza invece l'URL avviato da SP.

Disabilitazione dell'autenticazione SAML

Per disabilitare l'autenticazione SAML per OpenSearch Dashboards (console)

1. Scegli il dominio, Operazioni quindi Modifica configurazione di sicurezza.
2. Deselezionare Abilita autenticazione SAML.
3. Scegli Save changes (Salva modifiche).
4. Al termine dell'elaborazione del dominio, verificare la mappatura dei ruoli del controllo granulare degli accessi con la seguente richiesta:

   ```
   GET _plugins/_security/api/rolesmapping
   ```

   La disabilitazione dell'autenticazione SAML per le dashboard non rimuove le mappature per il nome utente principale SAML, il ruolo di backend principale SAML. and/or Se si desidera rimuovere queste mappature, accedere a Dashboards utilizzando il database utente interno (se abilitato) oppure l'API:

```
PUT _plugins/_security/api/rolesmapping/all_access
{
  "users": [
    "master-user"
  ]
}
```

IAM Identity Center Trusted Identity Propagation Support per OpenSearch

Ora puoi utilizzare i principali (utenti e gruppi) del tuo AWS IAM Identity Center configurati centralmente tramite Trusted Identity Propagation per accedere ai domini Amazon OpenSearch Service tramite OpenSearch le applicazioni di servizio. Per abilitare il supporto di IAM Identity Center per OpenSearch, dovrai abilitare l'uso di IAM Identity Center. Per saperne di più su come eseguire questa operazione, consulta Cos'è IAM Identity Center? . Vedi Come associare il OpenSearch dominio come origine dati nelle OpenSearch applicazioni? per i dettagli.

Puoi configurare IAM Identity Center utilizzando la console di OpenSearch servizio, il AWS Command Line Interface (AWS CLI) o il AWS SDKs.

> **Note**
>
> I principali di IAM Identity Center non sono supportati tramite dashboard (collocati insieme al cluster). Sono supportati solo tramite un'interfaccia OpenSearch utente centralizzata (dashboard).

Considerazioni

Prima di utilizzare IAM Identity Center con Amazon OpenSearch Service, devi considerare quanto segue:

- IAM Identity Center è abilitato nell'account.
- IAM Identity Center è disponibile nella tua regione.
- La versione del OpenSearch dominio è 1.3 o successiva.
- Il controllo granulare degli accessi è abilitato sul dominio.

- Il dominio si trova nella stessa regione dell'istanza IAM Identity Center.
- Il dominio e OpenSearch l'applicazione appartengono allo stesso AWS account.

Modifica della policy di accesso al dominio

Prima di configurare IAM Identity Center, è necessario aggiornare la policy di accesso al dominio o le autorizzazioni del ruolo IAM configurato nelle OpenSearch applicazioni per Trusted Identity Propagation.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::111122223333:role/OpenSearchRole"
            },
            "Action": "es:ESHttp*",
            "Resource": "arn:aws:es:us-east-1:111122223333:domain/example-domain/
*"
        }
    ]
}
```

Configurazione dell'autenticazione e dell'autorizzazione di IAM Identity Center (console)

Puoi abilitare l'autenticazione e l'autorizzazione di IAM Identity Center durante il processo di creazione del dominio o aggiornando un dominio esistente. I passaggi di configurazione variano leggermente a seconda dell'opzione scelta.

I passaggi seguenti spiegano come configurare un dominio esistente per l'autenticazione e l'autorizzazione di IAM Identity Center nella console OpenSearch di Amazon Service:

1. In Configurazione del dominio, vai a Configurazione di sicurezza, scegli Modifica e vai alla sezione Autenticazione IAM Identity Center e seleziona Abilita l'accesso all'API autenticato con IAM Identity Center.
2. Seleziona il tasto SubjectKey and Roles come segue.
   - Chiave dell'oggetto: scegli uno tra UserId (predefinito) UserName ed Email per utilizzare l'attributo corrispondente come principale di accesso al dominio.
   - Chiave dei ruoli: scegli uno tra GroupId (impostazione predefinita) e GroupName utilizza i valori degli attributi corrispondenti come ruolo di backend fine-grained-access-controlper tutti i gruppi associati al principale di IAM Identity Center.

Dopo aver apportato le modifiche, salva il dominio.

Configurazione del controllo granulare degli accessi

Dopo aver abilitato l'opzione IAM Identity Center sul tuo OpenSearch dominio, puoi configurare l'accesso ai principali di IAM Identity Center creando la mappatura dei ruoli al ruolo di backend. Il valore del ruolo di backend per il principale si basa sull'appartenenza al gruppo del responsabile di IAM Identity Center e sulla RolesKey configurazione di o. GroupId GroupName

> **Note**
>
> Amazon OpenSearch Service può supportare fino a 100 gruppi per un singolo utente. Se tenti di utilizzare un numero di istanze superiore a quello consentito, potresti riscontrare un'incoerenza nell'elaborazione delle fine-grained-access-control autorizzazioni e riceverai un messaggio di errore 403.

Configurazione dell'autenticazione e dell'autorizzazione (CLI) di IAM Identity Center

```
  aws opensearch update-domain-config \
      --domain-name my-domain \
      --identity-center-options '{"EnabledAPIAccess": true,
 "IdentityCenterInstanceARN": "instance arn",  "SubjectKey": "UserId/UserName/
UserEmail" , "RolesKey": "GroupId/GroupName"}'
```

Disabilitazione dell'autenticazione IAM Identity Center sul dominio

Per disabilitare IAM Identity Center sul tuo OpenSearch dominio:

1. Scegli il dominio, Operazioni quindi Modifica configurazione di sicurezza.
2. Deseleziona Abilita l'accesso all'API autenticato con IAM Identity Center.
3. Scegli Save changes (Salva modifiche).
4. Al termine dell'elaborazione del dominio, rimuovi le mappature dei ruoli aggiunte per i principali di IAM Identity Center

Per disabilitare IAM Identity Center tramite CLI, puoi usare quanto segue

```
aws opensearch update-domain-config \
    --domain-name my-domain \
    --identity-center-options '{"EnabledAPIAccess": false}'
```

Configurazione dell'autenticazione Amazon Cognito per dashboard OpenSearch

Puoi autenticare e proteggere l'installazione predefinita di OpenSearch dashboard di Amazon OpenSearch Service utilizzando Amazon Cognito. L'autenticazione Amazon Cognito è facoltativa e disponibile solo per i domini che utilizzano Elasticsearch 5.1 OpenSearch o versione successiva. Se non si configura l'autenticazione Amazon Cognito, è comunque possibile proteggere Dashboards usando una policy di accesso basata su IP e un server proxy, autenticazione di base HTTP o SAML.

Gran parte del processo di autenticazione avviene in Amazon Cognito, ma questa sezione offre linee guida e requisiti per configurare le risorse di Amazon Cognito in modo che funzionino con i domini di servizio. OpenSearch A tutte le risorse Amazon Cognito si applicano i prezzi standard.

> **Tip**
>
> La prima volta che configuri un dominio per utilizzare l'autenticazione Amazon Cognito per le OpenSearch dashboard, ti consigliamo di utilizzare la console. Le risorse Amazon Cognito

sono estremamente personalizzabili e la console può aiutare a identificare e comprendere le funzionalità importanti più adatte al proprio caso specifico.

Argomenti

- Prerequisiti
- Configurazione di un dominio per l'uso dell'autenticazione Amazon Cognito
- Concessione del ruolo autenticato
- Configurazione dei provider di identità
- (Facoltativo) Configurazione dell'accesso granulare
- (Facoltativo) Personalizzazione della pagina di accesso
- (Facoltativo) Configurazione della sicurezza avanzata
- Test in corso
- Quote
- Problemi di configurazione comuni
- Disattivazione dell'autenticazione Amazon Cognito per dashboard OpenSearch
- Eliminazione di domini che utilizzano l'autenticazione Amazon Cognito per dashboard OpenSearch

Prerequisiti

Prima di poter configurare l'autenticazione Amazon Cognito per le OpenSearch dashboard, devi soddisfare diversi prerequisiti. La console OpenSearch di servizio aiuta a semplificare la creazione di queste risorse, ma la comprensione dello scopo di ciascuna risorsa aiuta nella configurazione e nella risoluzione dei problemi. L'autenticazione Amazon Cognito per Dashboards richiede le risorse seguenti:

- Bacino d'utenza di Amazon Cognito
- Pool di identità di Amazon Cognito
- Ruolo IAM a cui è collegata la policy `AmazonOpenSearchServiceCognitoAccess` (`CognitoAccessForAmazonOpenSearch`)

> **Note**
>
> Il bacino d'utenza e il pool di identità devono trovarsi nella stessa Regione AWS. Puoi utilizzare lo stesso pool di utenti, pool di identità e ruolo IAM per aggiungere l'autenticazione Amazon Cognito per dashboard a più OpenSearch domini di servizio. Per ulteriori informazioni, consulta the section called “Quote”.

Informazioni sul bacino d'utenza

I pool di utenti hanno due caratteristiche principali: creano e gestiscono una directory di utenti e permettono agli utenti di registrarsi e accedere. Per istruzioni su come creare un pool di utenti, consulta la Guida introduttiva ai pool di utenti nella Amazon Cognito Developer Guide.

Quando crei un pool di utenti da utilizzare con OpenSearch Service, considera quanto segue:

- Il tuo pool di utenti Amazon Cognito deve avere un nome di dominio. OpenSearch Il servizio utilizza questo nome di dominio per reindirizzare gli utenti a una pagina di accesso per accedere alle dashboard. Oltre a un nome di dominio, il pool di utenti non richiede altre configurazioni non predefinite.
- È necessario specificare gli attributi standard obbligatori del pool, ovvero attributi come nome, data di nascita, indirizzo e-mail e numero di telefono. Poiché non puoi modificare questi attributi dopo che crei il pool di utenti, scegli quelli più importanti a questo punto.
- Durante la creazione del pool di utenti, scegli se gli utenti possono creare i propri account, la complessità minima delle password per gli account e se abilitare l'autenticazione a più fattori. Se prevedi di usare un provider di identità esterno, queste impostazioni non si escludono a vicenda. Tecnicamente, puoi abilitare il pool di utenti come provider di identità e abilitare un provider di identità esterno, ma la maggior parte degli utenti preferisce scegliere solo uno dei due approcci.

Il pool di utenti IDs assume la forma di. *region*_*ID* Se prevedi di utilizzare la AWS CLI o un AWS SDK per configurare il OpenSearch servizio, prendi nota dell'ID.

Informazioni sul pool di identità

I pool di identità permettono di assegnare ruoli temporanei con privilegi limitati agli utenti dopo che questi accedono. Per istruzioni sulla creazione di un pool di identità, consulta la panoramica della console dei pool di identità nella Amazon Cognito Developer Guide. Quando crei un pool di identità da utilizzare con OpenSearch Service, considera quanto segue:

- Se si utilizza la console Amazon Cognito, è necessario selezionare la casella di controllo Consenti l'accesso a identità non autenticate per creare il pool di identità. Dopo aver creato il pool di identità e configurato il dominio del OpenSearch servizio, Amazon Cognito disabilita questa impostazione.
- Non devi aggiungere provider di identità esterni al pool di identità. Quando configuri OpenSearch Service per utilizzare l'autenticazione Amazon Cognito, configura il pool di identità per utilizzare il pool di utenti che hai appena creato.
- Dopo aver creato il pool di identità, devi scegliere i ruoli IAM non autenticati e autenticati. Questi ruoli specificano le policy d'accesso associate agli utenti prima e dopo l'accesso. Se si utilizza la console Amazon Cognito, questa può creare i ruoli per conto dell'utente. Dopo aver creato il ruolo autenticato, annota l'ARN, che usa il formato `arn:aws:iam::123456789012:role/Cognito_identitypoolnameAuth_Role`.

Il pool di identità IDs assume la forma di. *region*`:`*ID*`-`*ID*`-`*ID*`-`*ID*`-`*ID* Se prevedi di utilizzare la AWS CLI o un AWS SDK per configurare il OpenSearch servizio, prendi nota dell'ID.

Informazioni sul ruolo CognitoAccessForAmazonOpenSearch

OpenSearch Il servizio richiede le autorizzazioni per configurare i pool di utenti e identità di Amazon Cognito e utilizzarli per l'autenticazione. È possibile utilizzare`AmazonOpenSearchServiceCognitoAccess`, che è una politica AWS gestita, per questo scopo. `AmazonESCognitoAccess`è una politica precedente che è stata sostituita da `AmazonOpenSearchServiceCognitoAccess` quando il servizio è stato rinominato Amazon OpenSearch Service. Entrambe le policy forniscono le autorizzazioni minime di Amazon Cognito necessarie per abilitare l'autenticazione Amazon Cognito. Per i dettagli sulle politiche, consulta AmazonOpenSearchServiceCognitoAccessla AWS Managed Policy Reference Guide.

Se utilizzi la console per creare o configurare il tuo dominio di OpenSearch servizio, questa crea per te un ruolo IAM e allega la `AmazonOpenSearchServiceCognitoAccess` policy (o la `AmazonESCognitoAccess` policy se si tratta di un dominio Elasticsearch) al ruolo. Il nome predefinito per questo ruolo è `CognitoAccessForAmazonOpenSearch`.

`AmazonESCognitoAccess`Entrambe le politiche `AmazonOpenSearchServiceCognitoAccess` relative alle autorizzazioni dei ruoli consentono a OpenSearch Service di completare le seguenti azioni su tutti i pool di identità e utenti:

- Operazione: `cognito-idp:DescribeUserPool`
- Operazione: `cognito-idp:CreateUserPoolClient`

- Operazione: `cognito-idp:DeleteUserPoolClient`
- Operazione: `cognito-idp:UpdateUserPoolClient`
- Operazione: `cognito-idp:DescribeUserPoolClient`
- Operazione: `cognito-idp:AdminInitiateAuth`
- Operazione: `cognito-idp:AdminUserGlobalSignOut`
- Operazione: `cognito-idp:ListUserPoolClients`
- Operazione: `cognito-identity:DescribeIdentityPool`
- Operazione: `cognito-identity:SetIdentityPoolRoles`
- Operazione: `cognito-identity:GetIdentityPoolRoles`

Se si utilizza il AWS CLI o uno dei ruoli AWS SDKs, è necessario creare il proprio ruolo, allegare la policy e specificare l'ARN per questo ruolo quando si configura il dominio di OpenSearch servizio. Il ruolo deve avere la relazione di trust seguente:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "opensearchservice.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

Per istruzioni, consulta Creare un ruolo per delegare le autorizzazioni a un AWS servizio e Aggiungere e rimuovere i permessi di identità IAM nella Guida per l'utente IAM.

Configurazione di un dominio per l'uso dell'autenticazione Amazon Cognito

Dopo aver completato i prerequisiti, puoi configurare un dominio di OpenSearch servizio per utilizzare Amazon Cognito for Dashboards.

> **Note**
>
> Amazon Cognito non è disponibile in tutti i casi. Regioni AWS Per un elenco delle regioni supportate, consulta Endpoint di servizio per Amazon Cognito. Non è necessario utilizzare la stessa regione per Amazon Cognito utilizzata per OpenSearch Service.

Configurazione dell'autenticazione Amazon Cognito (console)

Poiché crea il `CognitoAccessForAmazonOpenSearch` ruolo per te, la console offre l'esperienza di configurazione più semplice. Oltre alle autorizzazioni standard del OpenSearch Servizio, è necessario il seguente set di autorizzazioni per utilizzare la console per creare un dominio che utilizza l'autenticazione Amazon Cognito per le dashboard. OpenSearch

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeVpcs",
        "cognito-identity:ListIdentityPools",
        "cognito-idp:ListUserPools",
        "iam:CreateRole",
        "iam:AttachRolePolicy"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:GetRole",
        "iam:PassRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/service-
role/CognitoAccessForAmazonOpenSearch"
    }
  ]
}
```

Per istruzioni per aggiungere autorizzazioni per un'identità (utente, gruppo di utenti o ruolo), consulta Aggiunta di autorizzazioni per identità IAM (console).

Se `CognitoAccessForAmazonOpenSearch` esiste già, è necessario un numero minore di autorizzazioni:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeVpcs",
        "cognito-identity:ListIdentityPools",
        "cognito-idp:ListUserPools"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:GetRole",
        "iam:PassRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/service-
role/CognitoAccessForAmazonOpenSearch"
    }
  ]
}
```

Come configurare l'autenticazione Amazon Cognito per Dashboards (console)

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa/.
2. In Domini, seleziona il dominio che desideri configurare.
3. Scegli Operazioni, quindi Modifica configurazione di sicurezza.
4. SelezionaAbilitare l'autenticazione Amazon Cognito.
5. Per Regione, seleziona Regione AWS quella che contiene il pool di utenti e il pool di identità di Amazon Cognito.

6. Per Bacino d'utenza Cognito, seleziona un pool di utenti o creane uno. Per ulteriori informazioni, consulta the section called “Informazioni sul bacino d'utenza”.
7. Per Pool di identità Cognito, seleziona un pool di identità o creane uno. Per ulteriori informazioni, consulta the section called “Informazioni sul pool di identità”.

   > **Note**
   >
   > I link Crea bacino d'utenza e Crea pool di identità reindirizzano alla console Amazon Cognito e richiedono la creazione manuale di queste risorse. Il processo non è automatico. Per ulteriori informazioni, consulta the section called “Prerequisiti”.

8. Per Nome ruolo IAM, utilizza il valore di default `CognitoAccessForAmazonOpenSearch` (consigliato) o specifica un nuovo nome. Per ulteriori informazioni, consulta the section called “Informazioni sul ruolo CognitoAccessForAmazonOpenSearch”.
9. Scegli Salva modifiche.

Al termine dell'elaborazione del dominio, consulta the section called “Concessione del ruolo autenticato” e the section called “Configurazione dei provider di identità” per informazioni sulle altre fasi di configurazione.

Configurazione dell'autenticazione Amazon Cognito (AWS CLI)

Usa il `--cognito-options` parametro per configurare il tuo dominio OpenSearch di servizio. La sintassi seguente viene usata dai comandi `create-domain` e `update-domain-config`:

```
--cognito-options Enabled=true,UserPoolId="user-pool-id",IdentityPoolId="identity-pool-id",RoleArn="arn:aws:iam::123456789012:role/CognitoAccessForAmazonOpenSearch"
```

Esempio

Nell'esempio seguente viene creato un dominio nella regione `us-east-1` che abilita l'autenticazione Amazon Cognito per Dashboards tramite il ruolo `CognitoAccessForAmazonOpenSearch` e fornisce a `Cognito_Auth_Role` l'accesso al dominio:

```
aws opensearch create-domain --domain-name my-domain --
region us-east-1 --access-policies '{ "Version": "2012-10-17",
        "Statement":[{"Effect":"Allow","Principal":{"AWS":
 ["arn:aws:iam::123456789012:role/
Cognito_Auth_Role"]},"Action":"es:ESHttp*","Resource":"arn:aws:es:us-
```

```
east-1:123456789012:domain/*" }]}' --engine-version "OpenSearch_1.0"
 --cluster-config InstanceType=m4.xlarge.search,InstanceCount=1
 --ebs-options EBSEnabled=true,VolumeSize=10 --cognito-options
 Enabled=true,UserPoolId="us-east-1_123456789",IdentityPoolId="us-
east-1:12345678-1234-1234-1234-123456789012",RoleArn="arn:aws:iam::123456789012:role/
CognitoAccessForAmazonOpenSearch"
```

Al termine dell'elaborazione del dominio, consulta the section called “Concessione del ruolo autenticato” e the section called “Configurazione dei provider di identità” per informazioni sulle altre fasi di configurazione.

Configurazione dell'autenticazione Amazon Cognito ()AWS SDKs

AWS SDKs (eccetto Android e iOS SDKs) supportano tutte le operazioni definite nell'Amazon OpenSearch Service API Reference, incluso il `CognitoOptions` parametro per le `UpdateDomainConfig` operazioni `CreateDomain` and. Per ulteriori informazioni sull'installazione e l'utilizzo di AWS SDKs, consulta AWS Software Development Kits.

Al termine dell'elaborazione del dominio, consulta the section called “Concessione del ruolo autenticato” e the section called “Configurazione dei provider di identità” per informazioni sulle altre fasi di configurazione.

Concessione del ruolo autenticato

Per impostazione predefinita, il ruolo IAM autenticato che hai configurato seguendo le linee guida contenute the section called “Informazioni sul pool di identità” non dispone dei privilegi necessari per accedere alle dashboard. OpenSearch Devi fornire al ruolo autorizzazioni aggiuntive.

> **Note**
>
> Se hai configurato un controllo granulare degli accessi e utilizzi una policy di accesso aperta o basata su IP, puoi saltare questo passaggio.

È possibile includere queste autorizzazioni in una policy basata sull'identità, ma a meno che non si desideri che gli utenti autenticati abbiano accesso a tutti i domini di OpenSearch servizio, una policy basata sulle risorse collegata a un singolo dominio è l'approccio migliore.

Per `Principal`, specifica l'ARN del ruolo autenticato di Cognito che hai configurato con le linee guida in the section called “Informazioni sul pool di identità”.

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "AWS":[
               "arn:aws:iam::111122223333:role/Cognito_identitypoolname/
Auth_Role"
            ]
         },
         "Action":[
            "es:ESHttp*"
         ],
         "Resource":"arn:aws:es:us-east-1:111122223333:domain/domain-name/*"
      }
   ]
}
```

Per istruzioni sull'aggiunta di una politica basata sulle risorse a un dominio di servizio, consulta. OpenSearch the section called “Configurazione delle policy di accesso”

Configurazione dei provider di identità

Quando configuri un dominio per utilizzare l'autenticazione Amazon Cognito per dashboard, OpenSearch Service aggiunge un client di app al pool di utenti e aggiunge il pool di utenti al pool di identità come provider di autenticazione.

> **Warning**
>
> Non rinominare o eliminare il client app.

A seconda di come hai configurato il pool di utenti, potresti dover creare gli account utente manualmente oppure gli utenti potrebbero essere in grado di creare i propri account. Se queste impostazioni sono accettabili, non devi eseguire altre operazioni. Molte persone, tuttavia, preferiscono usare provider di identità esterni.

Per abilitare un provider di identità SAML 2.0, devi fornire un documento di metadati SAML. Per abilitare provider di identità social come Login with Amazon, Facebook e Google, devi ottenere un ID app e un segreto dell'app da questi provider. Puoi abilitare qualsiasi combinazione di provider di identità.

Il metodo più semplice per configurare il bacino d'utenza è usare la console Amazon Cognito. Per istruzioni, consulta Accesso al pool di utenti con provider di identità di terze parti e Impostazioni specifiche dell'applicazione con client di app nella Amazon Cognito Developer Guide.

(Facoltativo) Configurazione dell'accesso granulare

Avrai notato che le impostazioni predefinite del pool di identità assegnano a ogni utente che accede lo stesso ruolo IAM (`Cognito_`*`identitypool`*`Auth_Role`), il che significa che ogni utente può accedere alle stesse risorse. AWS Se, ad esempio, si desidera utilizzare il controllo granulare degli accessi con Amazon Cognito, ad esempio se si desidera che gli analisti dell'organizzazione abbiano accesso in sola lettura a diversi indici, ma gli sviluppatori abbiano accesso in scrittura a tutti gli indici, sono disponibili due opzioni:

- Puoi creare gruppi di utenti e configurare il provider di identità in modo da scegliere il ruolo IAM in base al token di autenticazione dell'utente (consigliato).
- Puoi configurare il provider di identità in modo da scegliere il ruolo IAM in base a uno o più ruoli.

Per una procedura dettagliata che include il controllo granulare degli accessi, consultare the section called “Tutorial: controllo granulare degli accessi con autenticazione Cognito”.

> **Important**
>
> Proprio come il ruolo di default, Amazon Cognito deve far parte della relazione di trust di ogni ruolo aggiuntivo. Per i dettagli, consulta Creazione di ruoli per la mappatura dei ruoli nella Amazon Cognito Developer Guide.

Gruppi di utenti e token

Quando crei un gruppo di utenti, devi scegliere un ruolo IAM per i membri del gruppo. Per informazioni sulla creazione di gruppi, consulta Aggiungere gruppi a un pool di utenti nella Amazon Cognito Developer Guide.

Dopo aver creato uno o più gruppi di utenti, puoi configurare il provider di autenticazione in modo da assegnare agli utenti i ruoli dei rispettivi gruppi anziché il ruolo predefinito del pool di identità. Selezionare Scegli ruolo da token, quindi selezionare Usa ruolo autenticato predefinito o DENY per specificare il modo in cui il pool di identità gestisce gli utenti che non fanno parte di un gruppo.

Regolamento

Le regole sono essenzialmente una serie di istruzioni `if` che Amazon Cognito valuta in sequenza. Ad esempio, se l'indirizzo e-mail di un utente contiene `@corporate`, Amazon Cognito assegna all'utente il ruolo `Role_A`. Se l'indirizzo e-mail di un utente contiene `@subsidiary`, assegna all'utente il ruolo `Role_B`. In caso contrario, assegna all'utente il ruolo autenticato predefinito.

Per ulteriori informazioni, consulta Usare la mappatura basata su regole per assegnare ruoli agli utenti nella Amazon Cognito Developer Guide.

(Facoltativo) Personalizzazione della pagina di accesso

Puoi utilizzare la console Amazon Cognito per caricare un logo personalizzato e apportare modifiche CSS alla pagina di accesso. Per istruzioni e un elenco completo delle proprietà CSS, consulta Personalizzazione del marchio dell'interfaccia utente ospitata (classica) nella Amazon Cognito Developer Guide.

(Facoltativo) Configurazione della sicurezza avanzata

I bacini d'utenza di Amazon Cognito supportano funzionalità di sicurezza avanzate quali l'autenticazione a più fattori, la verifica di credenziali compromesse e l'autenticazione adattiva. Per ulteriori informazioni, consulta la sezione Utilizzo delle funzionalità di sicurezza dei pool di utenti di Amazon Cognito nella Amazon Cognito Developer Guide.

Test in corso

Quando la configurazione sembra soddisfacente, verificare che l'esperienza utente soddisfi le aspettative.

Per accedere alle dashboard OpenSearch

1. Passare `https://`*`opensearch-domain`*`/_dashboards` in un Web browser. Per accedere direttamente a un tenant specifico, aggiungere `?security_tenant=`*`tenant-name`* all'URL.
2. Accedere usando le credenziali preferite.

3. Dopo il caricamento di OpenSearch Dashboards, configura almeno un modello di indice. Dashboards usa questi modelli per identificare gli indici da analizzare. Immettere *, scegliere Next step (Fase successiva) e quindi Create index pattern (Crea modello di indice).
4. Per cercare o esplorare i dati, scegliere Discover (Individua).

Se qualsiasi fase di questo processo non riesce, consulta the section called “Problemi di configurazione comuni” per informazioni sulla risoluzione dei problemi.

Quote

Amazon Cognito applica limiti flessibili a molte delle risorse. Se desideri abilitare l'autenticazione delle dashboard per un gran numero di domini di OpenSearch servizio, consulta Quotas in Amazon Cognito e richiedi l'aumento dei limiti, se necessario.

Ogni dominio OpenSearch di servizio aggiunge un client di app al pool di utenti, che aggiunge un provider di autenticazione al pool di identità. Se OpenSearch abiliti l'autenticazione Dashboards per più di 10 domini, potresti incontrare il limite del «numero massimo di provider del pool di utenti di Amazon Cognito per pool di identità». Se superi un limite, tutti i domini di OpenSearch servizio che tenti di configurare per utilizzare l'autenticazione Amazon Cognito per dashboard possono rimanere bloccati in uno stato di configurazione di Elaborazione.

Problemi di configurazione comuni

La tabella seguente elenca i problemi di configurazione più comuni e le relative soluzioni.

Servizio di configurazione OpenSearch

Problema	Soluzione
`OpenSearch Service can't create the role` (console)	Non devi avere le autorizzazioni IAM corrette. Aggiungi le autorizzazioni specificate in the section called “Configurazione dell'autenticazione Amazon Cognito (console)”.
`User is not authorized to perform: iam:PassRole on resource CognitoAccessForAmazonOpenSearch` (console)	Non disponi `iam:PassRole` delle autorizzazioni per il CognitoAccessForAmazonOpenSearchruolo. Collega la policy seguente al tuo account:

Problema	Soluzione
	JSON `{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Action": [ "iam:PassRole" ], "Resource": "arn:aws: iam:: 123456789012:role/service-role/ CognitoAccessForAmazonOpenSearch " } ] }` In alternativa, puoi collegare la policy `IAMFullAccess` .
`User is not authorize d to perform: cognito- identity:ListIdenti tyPools on resource`	Non è necessario disporre di autorizzazioni di lettura per Amazon Cognito. Collega la policy `AmazonCog nitoReadOnly` al tuo account.
`An error occurred (Validati onException) when calling the CreateDomain operation : OpenSearch Service must be allowed to use the passed role`	OpenSearch Il servizio non è specificato nella relazione di fiducia del `CognitoAccessForAmazonOpenS earch` ruolo. Controlla che il ruolo usi la relazione di trust specificata in the section called “Informazioni sul ruolo CognitoAccessForAmazonOpenSearch”. In alternativa, per configurare l'autenticazione Amazon Cognito utilizzare la console. La console crea un ruolo per te.

Problema	Soluzione
`An error occurred (ValidationException) when calling the CreateDomain operation: User is not authorized to perform: cognito-idp: action on resource: user pool`	Il ruolo specificato in `--cognito-options` non ha le autorizzazioni per accedere ad Amazon Cognito. Verifica che al ruolo sia allegata la `AmazonOpenSearchServiceCognitoAccess` policy AWS gestita. In alternativa, per configurare l'autenticazione Amazon Cognito utilizzare la console. La console crea un ruolo per te.
`An error occurred (ValidationException) when calling the CreateDomain operation: User pool does not exist`	OpenSearch Il servizio non riesce a trovare il pool di utenti. Conferma di averne creato uno e di avere l'ID corretto. Per trovare l'ID, puoi utilizzare la console Amazon Cognito o il seguente AWS CLI comando: `aws cognito-idp list-user-pools --max-results 60 --region region`
`An error occurred (ValidationException) when calling the CreateDomain operation: IdentityPool not found`	OpenSearch Il servizio non riesce a trovare il pool di identità. Conferma di averne creato uno e di avere l'ID corretto. Per trovare l'ID, puoi utilizzare la console Amazon Cognito o il seguente AWS CLI comando: `aws cognito-identity list-identity-pools --max-results 60 --region region`
`An error occurred (ValidationException) when calling the CreateDomain operation: Domain needs to be specified for user pool`	Il pool di utenti non ha un nome di dominio. È possibile configurarne uno tramite la console Amazon Cognito o con il seguente comando della AWS CLI : `aws cognito-idp create-user-pool-domain --domain name --user-pool-id id`

Accesso ai pannelli OpenSearch di controllo

Problema	Soluzione
La pagina di accesso non mostra i provider di identità preferiti.	Verifica di aver abilitato il provider di identità per il client dell'app OpenSearch Service come specificato inthe section called “Configurazione dei provider di identità”.
La pagina di accesso non sembra associata all'organizzazione.	Per informazioni, consulta the section called “(Facoltativo) Personalizzazione della pagina di accesso”.
Le credenziali di accesso non funzionano.	Verifica di aver configurato il provider di identità come indicato in the section called “Configurazione dei provider di identità”. Se utilizzi il pool di utenti come provider di identità, verifica che l'account esista nella console Amazon Cognito.
OpenSearch Le dashboard non si caricano affatto o non funzionano correttamente.	Il ruolo autenticato con Amazon Cognito deve disporre delle autorizzazioni `es:ESHttp*` per il dominio (`/*`) per poter accedere e utilizzare Dashboards. Verifica di aver aggiungo una policy d'accesso come indicato in the section called “Concessione del ruolo autenticato”.
Quando esco dai OpenSearch dashboard da una scheda, nelle schede rimanenti viene visualizzato un messaggio che indica che il token di aggiornamento è stato revocato.	Quando esci da una sessione di OpenSearch Dashboard s mentre utilizzi l'autenticazione Amazon Cognito OpenSearch , Service esegue AdminUserGlobalSignOutun'operazione che ti disconnette da tutte le sessioni di Dashboards OpenSearch attive.
`Invalid identity pool configuration. Check assigned IAM roles for this pool.`	Amazon Cognito non dispone delle autorizzazioni per assumere il ruolo IAM per conto dell'utente autenticato. Modificare la relazione di trust per il ruolo per includere: JSON `{ "Version":"2012-10-17",`

Problema	Soluzione
	`"Statement": [{ "Effect": "Allow", "Principal": { "Federated": "cognito-identity.amazonaws.com" }, "Action": "sts:AssumeRoleWithWebIdentity", "Condition": { "StringEquals": { "cognito-identity.amazonaws.com:aud": " identity-pool-id " }, "ForAnyValue:StringLike": { "cognito-identity.amazonaws.com:amr": "authenticated" } } }] }`
`Token is not from a supported provider of this identity pool.`	Questo errore poco comune può verificarsi quando rimuovi il client app dal pool di utenti. Provare ad aprire Dashboards in una nuova sessione del browser.

Disattivazione dell'autenticazione Amazon Cognito per dashboard OpenSearch

Usa la procedura seguente per disabilitare l'autenticazione Amazon Cognito per Dashboards.

Come disabilitare l'autenticazione Amazon Cognito per Dashboards (console)

1. Apri la console Amazon OpenSearch Service.
2. In Domini, scegli il dominio che desideri configurare.
3. Scegli Operazioni, quindi Modifica configurazione di sicurezza.
4. Deselezionare Abilita autenticazione Amazon Cognito.

5. Scegli Save changes (Salva modifiche).

> **Important**
>
> Se il bacino d'utenza e il pool di identità di Amazon Cognito non sono più necessari, è possibile eliminarli. Altrimenti, continuano a esserti addebitati i costi.

Eliminazione di domini che utilizzano l'autenticazione Amazon Cognito per dashboard OpenSearch

Per evitare che i domini che utilizzano l'autenticazione Amazon Cognito per dashboard rimangano bloccati in uno stato di configurazione di Elaborazione, OpenSearch elimina i domini di servizio prima di eliminare i pool di utenti e identità di Amazon Cognito associati.

Utilizzo di ruoli collegati ai servizi per Amazon Service OpenSearch

Amazon OpenSearch Service utilizza ruoli collegati ai servizi AWS Identity and Access Management (IAM). Un ruolo collegato al servizio è un tipo unico di ruolo IAM collegato direttamente al servizio. OpenSearch I ruoli collegati ai servizi sono predefiniti da OpenSearch Service e includono tutte le autorizzazioni richieste dal servizio per chiamare altri servizi per tuo conto. AWS

Un ruolo collegato al servizio semplifica la configurazione del OpenSearch servizio perché non è necessario aggiungere manualmente le autorizzazioni necessarie. OpenSearch Il servizio definisce le autorizzazioni dei suoi ruoli collegati al servizio e, se non diversamente definito, solo OpenSearch il servizio può assumerne i ruoli. Le autorizzazioni definite includono la policy di attendibilità e la policy delle autorizzazioni. Una policy delle autorizzazioni specifica non può essere collegata a un'altra entità IAM. Per gli aggiornamenti ai ruoli collegati ai servizi e alle politiche di autorizzazione, consulta la Cronologia dei documenti per Amazon Service. OpenSearch

Per informazioni su altri servizi che supportano i ruoli collegati ai servizi, consulta i AWS servizi che funzionano con IAM e cerca i servizi con Sì nella colonna Ruoli collegati ai servizi. Scegli Sì in corrispondenza di un link per visualizzare la documentazione relativa al ruolo collegato al servizio per tale servizio.

Argomenti

- Utilizzo di ruoli collegati ai servizi per creare domini VPC e fonti di dati di interrogazione diretta

- Utilizzo di ruoli collegati ai servizi per creare raccolte Serverless OpenSearch
- Utilizzo di ruoli collegati ai servizi per creare pipeline di ingestione OpenSearch

Utilizzo di ruoli collegati ai servizi per creare domini VPC e fonti di dati di interrogazione diretta

Amazon OpenSearch Service utilizza ruoli collegati ai servizi AWS Identity and Access Management (IAM). Un ruolo collegato al servizio è un tipo unico di ruolo IAM collegato direttamente al servizio. OpenSearch I ruoli collegati ai servizi sono predefiniti da OpenSearch Service e includono tutte le autorizzazioni richieste dal servizio per chiamare altri servizi per tuo conto. AWS

OpenSearch Il servizio utilizza il ruolo collegato al servizio denominato AWSServiceRoleForAmazonOpenSearchService, che fornisce le autorizzazioni minime di Amazon EC2 ed Elastic Load Balancing necessarie affinché il ruolo consenta l'accesso VPC a un dominio o a un'origine dati di query diretta.

Ruolo di Elasticsearch legacy

Amazon OpenSearch Service utilizza un ruolo collegato al servizio chiamato. `AWSServiceRoleForAmazonOpenSearchService` Gli account potrebbero anche contenere un ruolo collegato al servizio legacy denominato `AWSServiceRoleForAmazonElasticsearchService`, che funziona con gli endpoint dell'API Amazon Elasticsearch Service obsoleti.

Se il ruolo legacy di Elasticsearch non esiste nel tuo account, OpenSearch Service crea automaticamente un nuovo ruolo OpenSearch collegato al servizio la prima volta che crei un dominio. OpenSearch In caso contrario, l'account continua a utilizzare il ruolo Elasticsearch. Perché questa operazione di creazione automatica riesca, devi disporre delle autorizzazioni per l'operazione `iam:CreateServiceLinkedRole`.

Permissions

Ai fini dell'assunzione del ruolo, il ruolo collegato al servizio `AWSServiceRoleForAmazonOpenSearchService` considera attendibili i seguenti servizi:

- `opensearchservice.amazonaws.com`

La politica di autorizzazione dei ruoli denominata AmazonOpenSearchServiceRolePolicyconsente a OpenSearch Service di completare le seguenti azioni sulle risorse specificate:

- Operazione: `acm:DescribeCertificate` su *
- Operazione: `cloudwatch:PutMetricData` su *
- Operazione: `ec2:CreateNetworkInterface` su *
- Operazione: `ec2:DeleteNetworkInterface` su *
- Operazione: `ec2:DescribeNetworkInterfaces` su *
- Operazione: `ec2:ModifyNetworkInterfaceAttribute` su *
- Operazione: `ec2:DescribeSecurityGroups` su *
- Operazione: `ec2:DescribeSubnets` su *
- Operazione: `ec2:DescribeVpcs` su *
- Azione: `ec2:CreateTags` su tutte le interfacce di rete e gli endpoint VPC
- Operazione: `ec2:DescribeTags` su *
- Azione: `ec2:CreateVpcEndpoint` su tutti i gruppi di sicurezza VPCs, le sottoreti e le tabelle di routing, nonché su tutti gli endpoint VPC quando la richiesta contiene il tag `OpenSearchManaged=true`
- Azione: `ec2:ModifyVpcEndpoint` su tutti i gruppi di sicurezza VPCs, le sottoreti e le tabelle di routing, nonché su tutti gli endpoint VPC quando la richiesta contiene il tag `OpenSearchManaged=true`
- Azione: `ec2:DeleteVpcEndpoints` su tutti gli endpoint quando la richiesta contiene il tag `OpenSearchManaged=true`
- Operazione: `ec2:AssignIpv6Addresses` su *
- Operazione: `ec2:UnAssignIpv6Addresses` su *
- Operazione: `elasticloadbalancing:AddListenerCertificates` su *
- Operazione: `elasticloadbalancing:RemoveListenerCertificates` su *

Per consentire a un'entità IAM (come un utente, un gruppo o un ruolo) di creare, modificare o eliminare un ruolo collegato al servizio è necessario configurare le relative autorizzazioni. Per ulteriori informazioni, consulta Autorizzazioni del ruolo collegato ai servizi nella Guida per l'utente di IAM.

Creazione del ruolo collegato ai servizi

Non hai bisogno di creare manualmente un ruolo collegato ai servizi. Quando crei un dominio abilitato per VPC o un'origine dati di query diretta utilizzando Console di gestione AWS, OpenSearch Service crea automaticamente il ruolo collegato al servizio. Perché questa operazione di creazione automatica riesca, devi disporre delle autorizzazioni per l'operazione `iam:CreateServiceLinkedRole`.

Per creare manualmente un ruolo collegato ai servizi, è possibile utilizzare la console IAM, la CLI IAM o l'API IAM. Per ulteriori informazioni, consultare Creazione di un ruolo collegato ai servizi nella Guida per l'utente di IAM.

Modifica del ruolo collegato ai servizi

OpenSearch Il servizio non ti consente di modificare il ruolo collegato al servizio. `AWSServiceRoleForAmazonOpenSearchService` Dopo aver creato un ruolo collegato al servizio, non è possibile modificarne il nome, perché potrebbero farvi riferimento diverse entità. È possibile tuttavia modificarne la descrizione utilizzando IAM. Per ulteriori informazioni, consulta Modifica di un ruolo collegato al servizio nella Guida per l'utente di IAM.

Eliminazione del ruolo collegato ai servizi

Se non è più necessario utilizzare una funzionalità o un servizio che richiede un ruolo collegato al servizio, ti consigliamo di eliminare il ruolo. In questo modo non sarà più presente un'entità non utilizzata che non viene monitorata e gestita attivamente. Tuttavia, è necessario effettuare la pulizia delle risorse associate al ruolo collegato ai servizi prima di poterlo eliminare manualmente.

Pulizia del ruolo collegato ai servizi

Prima di utilizzare IAM per eliminare un ruolo collegato ai servizi, devi innanzitutto verificare che il ruolo non abbia sessioni attive ed eliminare tutte le risorse utilizzate dal ruolo.

Per verificare se il ruolo collegato ai servizi dispone di una sessione attiva nella console IAM

1. Accedi Console di gestione AWS e apri la console IAM all'indirizzo. https://console.aws.amazon.com/iam/
2. Nel pannello di navigazione della console IAM seleziona Ruoli. Quindi, scegli il nome (non la casella di controllo) del ruolo `AWSServiceRoleForAmazonOpenSearchService`.
3. Nella pagina Summary (Riepilogo) per il ruolo selezionato, scegliere la scheda Access Advisor (Consulente accessi).

4. Nella scheda Access Advisor (Consulente accessi), esamina l'attività recente per il ruolo collegato ai servizi.

> **Note**
>
> Se non sei sicuro che OpenSearch Service stia utilizzando il `AWSServiceRoleForAmazonOpenSearchService` ruolo, puoi provare a eliminarlo. Se il servizio sta utilizzando il ruolo, l'eliminazione non riesce e si possono visualizzare le risorse che utilizzano il ruolo. Se il ruolo viene utilizzato, è necessario attendere il termine della sessione prima di poter eliminare il ruolo, and/or eliminare le risorse che lo utilizzano. Non puoi revocare la sessione per un ruolo collegato al servizio.

Eliminazione manuale di un ruolo collegato ai servizi

Elimina i ruoli collegati ai servizi dalla console IAM, dall'API o dalla CLI AWS . Per le istruzioni, consultare Eliminazione di un ruolo collegato ai servizi nella Guida per l'utente di IAM.

Utilizzo di ruoli collegati ai servizi per creare raccolte Serverless OpenSearch

OpenSearch Serverless utilizza ruoli collegati ai AWS Identity and Access Management servizi (IAM). Un ruolo collegato al servizio è un tipo unico di ruolo IAM collegato direttamente al servizio. OpenSearch I ruoli collegati ai servizi sono predefiniti da OpenSearch Service e includono tutte le autorizzazioni richieste dal servizio per chiamare altri servizi per tuo conto. AWS

OpenSearch Serverless utilizza il ruolo collegato al servizio denominato AWSServiceRoleForAmazonOpenSearchServerless, che fornisce le autorizzazioni necessarie al ruolo per pubblicare metriche relative al server sull'account dell'utente. CloudWatch La politica di autorizzazione dei ruoli associata a è denominata. AWSService RoleForAmazonOpenSearchServerless `AmazonOpenSearchServerlessServiceRolePolicy` Per ulteriori informazioni sulla policy, consulta AmazonOpenSearchServerlessServiceRolePolicyla AWS Managed Policy Reference Guide.

Autorizzazioni di ruolo collegate ai servizi per Serverless OpenSearch

OpenSearch Serverless utilizza il ruolo collegato al servizio denominato AWSServiceRoleForAmazonOpenSearchServerless, che consente a OpenSearch Serverless di chiamare i servizi per conto dell'utente. AWS

Il ruolo AWSService RoleForAmazonOpenSearchServerless collegato al servizio prevede che i seguenti servizi assumano il ruolo:

- `observability.aoss.amazonaws.com`

La politica di autorizzazione dei ruoli denominata `AmazonOpenSearchServerlessServiceRolePolicy` consente a OpenSearch Serverless di completare le seguenti azioni sulle risorse specificate:

- Azione: `cloudwatch:PutMetricData` su tutte le risorse di AWS

> **Note**
>
> La policy include la chiave di condizione`{"StringEquals": {"cloudwatch:namespace": "AWS/AOSS"}}`, il che significa che il ruolo collegato al servizio può inviare solo dati metrici allo spazio dei nomi. `AWS/AOSS` CloudWatch

Per consentire a un'entità IAM (come un utente, un gruppo o un ruolo) di creare, modificare o eliminare un ruolo collegato al servizio è necessario configurare le relative autorizzazioni. Per ulteriori informazioni, consulta Autorizzazioni del ruolo collegato ai servizi nella Guida per l'utente IAM.

Creazione del ruolo collegato al servizio per Serverless OpenSearch

Non hai bisogno di creare manualmente un ruolo collegato ai servizi. Quando crei una raccolta OpenSearch Serverless nella Console di gestione AWS, nella o nell' AWS API AWS CLI, OpenSearch Serverless crea automaticamente il ruolo collegato al servizio.

> **Note**
>
> La prima volta che si crea una raccolta, è necessario ottenere l'assegnazione del ruolo `iam:CreateServiceLinkedRole` in una policy basata sull'identità.

Se elimini questo ruolo collegato al servizio, è possibile ricrearlo seguendo lo stesso processo utilizzato per ricreare il ruolo nell'account. Quando crei una raccolta Serverless, OpenSearch OpenSearch Serverless crea nuovamente il ruolo collegato al servizio per te.

Puoi anche utilizzare la console IAM per creare un ruolo collegato ai servizi con lo use case Amazon OpenSearch Serverless. Nella AWS CLI o nell' AWS API, crea un ruolo collegato al servizio con il nome del servizio: `observability.aoss.amazonaws.com`

```
aws iam create-service-linked-role --aws-service-name
 "observability.aoss.amazonaws.com"
```

Per ulteriori informazioni, consulta Creazione di un ruolo collegato ai servizi nella Guida per l'utente IAM. Se elimini il ruolo collegato ai servizi, è possibile utilizzare lo stesso processo per crearlo nuovamente.

Modifica del ruolo collegato al servizio per Serverless OpenSearch

OpenSearch Serverless non consente di modificare il ruolo collegato al servizio. AWSService RoleForAmazonOpenSearchServerless Dopo aver creato un ruolo collegato al servizio, non è possibile modificarne il nome, perché potrebbero farvi riferimento diverse entità. È possibile tuttavia modificarne la descrizione utilizzando IAM. Per ulteriori informazioni, consulta Modifica di un ruolo collegato al servizio nella Guida per l'utente di IAM.

Eliminazione del ruolo collegato al servizio per Serverless OpenSearch

Se non è più necessario utilizzare una funzionalità o un servizio che richiede un ruolo collegato al servizio, consigliamo di eliminare il ruolo. Ciò impedisce di avere un'entità inutilizzata che non viene monitorata o gestita attivamente. Tuttavia, è necessario effettuare la pulizia delle risorse associate al ruolo collegato al servizio prima di poterlo eliminare manualmente.

Per eliminare il AWSServiceRoleForAmazonOpenSearchServerless, devi prima eliminare tutte le raccolte OpenSearch Serverless presenti nel tuo. Account AWS

> **Note**
>
> Se OpenSearch Serverless utilizza il ruolo quando si tenta di eliminare le risorse, l'eliminazione potrebbe non riuscire. In questo caso, attendi alcuni minuti e quindi ripeti l'operazione.

Per eliminare manualmente il ruolo collegato ai servizi mediante IAM

Utilizza la console IAM AWS CLI, o l' AWS API per eliminare il ruolo collegato al AWSService RoleForAmazonOpenSearchServerless servizio. Per ulteriori informazioni, consulta Eliminazione del ruolo collegato ai servizi nella Guida per l'utente IAM.

Regioni supportate per ruoli OpenSearch Serverless collegati ai servizi

OpenSearch Serverless supporta l'utilizzo del ruolo AWSService RoleForAmazonOpenSearchServerless collegato al servizio in tutte le regioni in cui è disponibile Serverless. OpenSearch Per un elenco delle regioni supportate, consulta gli endpoint e le quote Amazon OpenSearch Serverless nel. Riferimenti generali di AWS

Utilizzo di ruoli collegati ai servizi per creare pipeline di ingestione OpenSearch

Amazon OpenSearch Ingestion utilizza ruoli collegati ai servizi AWS Identity and Access Management (IAM). Un ruolo collegato ai servizi è un tipo unico di ruolo IAM collegato direttamente a Ingestion. OpenSearch I ruoli collegati ai servizi sono predefiniti da OpenSearch Ingestion e includono tutte le autorizzazioni richieste dal servizio per chiamare altri servizi per conto dell'utente. AWS

OpenSearch Ingestion utilizza il ruolo collegato al servizio denominato AWSServiceRoleForAmazonOpenSearchIngestionService, tranne quando si utilizza un VPC autogestito, nel qual caso utilizza il ruolo collegato al servizio denominato. AWSServiceRoleForOpensearchIngestionSelfManagedVpce La policy allegata fornisce le autorizzazioni necessarie al ruolo per creare un cloud privato virtuale (VPC) tra il tuo account OpenSearch e Ingestion e per CloudWatch pubblicare le metriche sul tuo account.

Permissions

Ai fini dell'assunzione del ruolo, il ruolo collegato al servizio `AWSServiceRoleForAmazonOpenSearchIngestionService` considera attendibili i seguenti servizi:

- `osis.amazonaws.com`

La politica di autorizzazione dei ruoli denominata `AmazonOpenSearchIngestionServiceRolePolicy` consente a OpenSearch Ingestion di completare le seguenti azioni sulle risorse specificate:

- Operazione: `cloudwatch:PutMetricData` su `cloudwatch:namespace": "AWS/OSIS"`
- Operazione: `ec2:CreateTags` su `arn:aws:ec2:*:*:network-interface/*`
- Operazione: `ec2:CreateVpcEndpoint` su `*`
- Operazione: `ec2:DeleteVpcEndpoints` su `*`
- Operazione: `ec2:DescribeSecurityGroups` su `*`
- Operazione: `ec2:DescribeSubnets` su `*`
- Operazione: `ec2:DescribeVpcEndpoints` su `*`
- Operazione: `ec2:ModifyVpcEndpoint` su `*`

Ai fini dell'assunzione del ruolo, il ruolo collegato al servizio `AWSServiceRoleForOpensearchIngestionSelfManagedVpce` considera attendibili i seguenti servizi:

- `self-managed-vpce.osis.amazonaws.com`

La politica di autorizzazione dei ruoli denominata `OpenSearchIngestionSelfManagedVpcePolicy` consente a OpenSearch Ingestion di completare le seguenti azioni sulle risorse specificate:

- Operazione: `ec2:DescribeSubnets` su `*`
- Operazione: `ec2:DescribeSecurityGroups` su `*`
- Operazione: `ec2:DescribeVpcEndpoints` su `*`
- Operazione: `cloudwatch:PutMetricData` su `cloudwatch:namespace": "AWS/OSIS"`

Per consentire a un'entità IAM (come un utente, un gruppo o un ruolo) di creare, modificare o eliminare un ruolo collegato al servizio è necessario configurare le relative autorizzazioni. Per ulteriori informazioni, consulta Autorizzazioni del ruolo collegato ai servizi nella Guida per l'utente IAM.

Creazione del ruolo collegato al servizio per Ingestion OpenSearch

Non hai bisogno di creare manualmente un ruolo collegato ai servizi. Quando create una pipeline di OpenSearch Ingestion nell'API, nella o nell' AWS API Console di gestione AWS, OpenSearch Ingestion crea automaticamente il ruolo collegato al servizio. AWS CLI

Se elimini questo ruolo collegato al servizio, è possibile ricrearlo seguendo lo stesso processo utilizzato per ricreare il ruolo nell'account. Quando crei una pipeline di Ingestion, OpenSearch Ingestion crea nuovamente il ruolo collegato al servizio per te. OpenSearch

Modifica del ruolo collegato al servizio per Ingestion OpenSearch

OpenSearch Ingestion non consente di modificare il ruolo collegato al servizio. `AWSServiceRoleForAmazonOpenSearchIngestionService` Dopo avere creato un ruolo collegato al servizio, non sarà possibile modificarne il nome perché varie entità potrebbero farvi riferimento. È possibile tuttavia modificarne la descrizione utilizzando IAM. Per ulteriori informazioni, consulta Modifica di un ruolo collegato al servizio nella Guida per l'utente di IAM.

Eliminazione del ruolo collegato al servizio per Ingestion OpenSearch

Se non è più necessario utilizzare una caratteristica o un servizio che richiede un ruolo collegato ai servizi, ti consigliamo di eliminare quel ruolo. In questo modo non sarà più presente un'entità non utilizzata che non viene monitorata e gestita attivamente. Tuttavia, è necessario effettuare la pulizia delle risorse associate al ruolo collegato ai servizi prima di poterlo eliminare manualmente.

Pulizia di un ruolo collegato ai servizi

Prima di utilizzare IAM; per eliminare un ruolo collegato al servizio, è necessario prima rimuovere qualsiasi risorsa utilizzata dal ruolo.

> **Note**
>
> Se OpenSearch Ingestion utilizza il ruolo quando si tenta di eliminare le risorse, l'eliminazione potrebbe non riuscire. In questo caso, attendi alcuni minuti e quindi ripeti l'operazione.

Per eliminare le risorse OpenSearch di Ingestion utilizzate dal ruolo or **`AWSServiceRoleForAmazonOpenSearchIngestionService`** **`AWSServiceRoleForOpensearchIngestionSelfManagedVpce`**

1. Vai alla console di Amazon OpenSearch Service e scegli Ingestion.
2. Elimina tutte le pipeline. Per istruzioni, consulta the section called "Eliminazione delle pipeline".

Eliminare il ruolo collegato al servizio per Ingestion OpenSearch

È possibile utilizzare la console OpenSearch Ingestion per eliminare un ruolo collegato al servizio.

Per eliminare un ruolo collegato ai servizi (console)

1. Passare alla IAM console (Console IAM).
2. Scegli Ruoli e cerca il ruolo o. AWSServiceRoleForAmazonOpenSearchIngestionServiceAWSServiceRoleForOpensearchIngestionSel
3. Seleziona il ruolo e scegli Elimina.

Codice di esempio per Amazon OpenSearch Service

Questo capitolo contiene un codice di esempio comune per lavorare con Amazon OpenSearch Service: firma delle richieste HTTP in diversi linguaggi di programmazione, compressione dei corpi delle richieste HTTP e utilizzo di AWS SDKs per creare domini.

Argomenti

- Compatibilità con i client Elasticsearch
- Compressione delle richieste HTTP in Amazon OpenSearch Service
- Utilizzo di AWS SDKs per interagire con Amazon OpenSearch Service

Compatibilità con i client Elasticsearch

Le versioni più recenti dei client Elasticsearch potrebbero includere controlli di licenza o versione che violano artificialmente la compatibilità. La tabella seguente include consigli sulle versioni di tali client da utilizzare per una migliore compatibilità con il OpenSearch Servizio.

> **Important**
>
> Queste versioni client non sono aggiornate e non sono aggiornate con le ultime dipendenze, incluso Log4j. Consigliamo vivamente di utilizzare le OpenSearch versioni dei client quando possibile.

Client	Versioni suggerite
Client REST Java di basso livello	7.13.4
Client REST Java di livello elevato	7,13,4
Client Elasticsearch Python	7,13,4
Client Elasticsearch Ruby	7,13,3
Client Elasticsearch Node.js	7,13,0

Compressione delle richieste HTTP in Amazon OpenSearch Service

Puoi comprimere le richieste e le risposte HTTP nei domini Amazon OpenSearch Service utilizzando la compressione gzip. La compressione gzip consente di ridurre le dimensioni dei documenti e di ridurre l'utilizzo della larghezza di banda e la latenza, migliorando così la velocità di trasferimento.

La compressione Gzip è supportata per tutti i domini che eseguono Elasticsearch 6.0 OpenSearch o versioni successive. Alcuni OpenSearch client dispongono del supporto integrato per la compressione gzip e molti linguaggi di programmazione dispongono di librerie che semplificano il processo.

Abilitazione della compressione gzip

Da non confondere con OpenSearch impostazioni simili, `http_compression.enabled` è specifico di OpenSearch Service e abilita o disabilita la compressione gzip su un dominio. Domini in esecuzione OpenSearch o Elasticsearch 7. x hanno la compressione gzip abilitata per impostazione predefinita, mentre i domini che eseguono Elasticsearch 6. x l'hanno disabilitata per impostazione predefinita.

Per abilitare la compressione gzip, inviare la seguente richiesta:

```
PUT _cluster/settings
{
  "persistent" : {
    "http_compression.enabled": true
  }
}
```

Le richieste a `_cluster/settings` devono essere decompresse, quindi potrebbe essere necessario utilizzare un client separato o una richiesta HTTP standard per aggiornare le impostazioni del cluster.

Per confermare di aver abilitato correttamente la compressione gzip, invia la seguente richiesta:

```
GET _cluster/settings?include_defaults=true
```

Assicurati di vedere la seguente impostazione nella risposta:

```
...
```

```
"http_compression": {
  "enabled": "true"
}
...
```

Intestazioni richieste

Quando si include un corpo di richiesta compresso con gzip, mantenere l'intestazione `Content-Type: application/json` standard e aggiungere l'intestazione `Content-Encoding: gzip`. Per accettare una risposta gzip compressa, aggiungere anche l'intestazione `Accept-Encoding: gzip`. Se un OpenSearch client supporta la compressione gzip, probabilmente include queste intestazioni automaticamente.

Codice di esempio (Python 3)

Nell'esempio seguente viene utilizzato opensearch-py per eseguire la compressione e inviare la richiesta. Questo codice firma la richiesta utilizzando le credenziali IAM.

```
from opensearchpy import OpenSearch, RequestsHttpConnection
from requests_aws4auth import AWS4Auth
import boto3

host = '' # e.g. my-test-domain.us-east-1.es.amazonaws.com
region = '' # e.g. us-west-1
service = 'es'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

# Create the client.
search = OpenSearch(
    hosts = [{'host': host, 'port': 443}],
    http_auth = awsauth,
    use_ssl = True,
    verify_certs = True,
    http_compress = True, # enables gzip compression for request bodies
    connection_class = RequestsHttpConnection
)

document = {
  "title": "Moneyball",
```

```
  "director": "Bennett Miller",
  "year": "2011"
}

# Send the request.
print(search.index(index='movies', id='1', body=document, refresh=True))

# print(search.index(index='movies', doc_type='_doc', id='1', body=document,
 refresh=True))
```

In alternativa, è possibile specificare le intestazioni corrette, comprimere il corpo della richiesta da soli e utilizzare una libreria HTTP standard come Requests. Questo codice firma la richiesta utilizzando le credenziali di base HTTP, che il dominio potrebbe supportare se si utilizza il controllo granulare degli accessi.

```
import requests
import gzip
import json

base_url = '' # The domain with https:// and a trailing slash. For example, https://my-
test-domain.us-east-1.es.amazonaws.com/
auth = ('master-user', 'master-user-password') # For testing only. Don't store
 credentials in code.

headers = {'Accept-Encoding': 'gzip', 'Content-Type': 'application/json',
           'Content-Encoding': 'gzip'}

document = {
  "title": "Moneyball",
  "director": "Bennett Miller",
  "year": "2011"
}

# Compress the document.
compressed_document = gzip.compress(json.dumps(document).encode())

# Send the request.
path = 'movies/_doc?refresh=true'
url = base_url + path
response = requests.post(url, auth=auth, headers=headers, data=compressed_document)
print(response.status_code)
print(response.text)
```

Utilizzo di AWS SDKs per interagire con Amazon OpenSearch Service

Questa sezione include esempi di come utilizzare l'API di configurazione di Amazon Service AWS SDKs per interagire con l'API OpenSearch di configurazione di Amazon Service. Questi esempi di codice mostrano come creare, aggiornare ed eliminare i domini OpenSearch di Service.

Java

Questa sezione include esempi per le versioni 1 e 2 di AWS SDK per Java.

Version 2

Questo esempio utilizza il OpenSearchClientBuildercostruttore della versione 2 di AWS SDK per Java per creare un OpenSearch dominio, aggiornarne la configurazione ed eliminarlo. Rimuovere il commento dalle chiamate a `waitForDomainProcessing` (e commentare la chiamata a `deleteDomain`) per fare in modo che il dominio sia online e utilizzabile.

```
package com.example.samples;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.regions.Region;
import software.amazon.awssdk.services.opensearch.OpenSearchClient;
import software.amazon.awssdk.services.opensearch.model.ClusterConfig;
import software.amazon.awssdk.services.opensearch.model.EBSOptions;
import software.amazon.awssdk.services.opensearch.model.CognitoOptions;
import software.amazon.awssdk.services.opensearch.model.NodeToNodeEncryptionOptions;
import software.amazon.awssdk.services.opensearch.model.CreateDomainRequest;
import software.amazon.awssdk.services.opensearch.model.CreateDomainResponse;
import software.amazon.awssdk.services.opensearch.model.DescribeDomainRequest;
import software.amazon.awssdk.services.opensearch.model.UpdateDomainConfigRequest;
import software.amazon.awssdk.services.opensearch.model.UpdateDomainConfigResponse;
import software.amazon.awssdk.services.opensearch.model.DescribeDomainResponse;
import software.amazon.awssdk.services.opensearch.model.DeleteDomainRequest;
import software.amazon.awssdk.services.opensearch.model.DeleteDomainResponse;
import software.amazon.awssdk.services.opensearch.model.OpenSearchException;
import software.amazon.awssdk.auth.credentials.DefaultCredentialsProvider;

/**
 * Sample class demonstrating how to use the Amazon Web Services SDK for Java to
 create, update,
```

```
 * and delete Amazon OpenSearch Service domains.
 */

public class OpenSearchSample {

    public static void main(String[] args) {

     String domainName = "my-test-domain";

     // Build the client using the default credentials chain.
        // You can use the CLI and run `aws configure` to set access key, secret
        // key, and default region.

        OpenSearchClient client = OpenSearchClient.builder()
          // Unnecessary, but lets you use a region different than your default.
          .region(Region.US_EAST_1)
          // Unnecessary, but if desired, you can use a different provider chain.
          .credentialsProvider(DefaultCredentialsProvider.create())
                     .build();

        // Create a new domain, update its configuration, and delete it.
        createDomain(client, domainName);
        //waitForDomainProcessing(client, domainName);
        updateDomain(client, domainName);
        //waitForDomainProcessing(client, domainName);
        deleteDomain(client, domainName);
    }

    /**
     * Creates an Amazon OpenSearch Service domain with the specified options.
     * Some options require other Amazon Web Services resources, such as an Amazon
 Cognito user pool
     * and identity pool, whereas others require just an instance type or instance
     * count.
     *
     * @param client
     *            The client to use for the requests to Amazon OpenSearch Service
     * @param domainName
     *            The name of the domain you want to create
     */

    public static void createDomain(OpenSearchClient client, String domainName) {

     // Create the request and set the desired configuration options
```

```
        try {

            ClusterConfig clusterConfig = ClusterConfig.builder()
                    .dedicatedMasterEnabled(true)
                    .dedicatedMasterCount(3)
                    // Small, inexpensive instance types for testing. Not
 recommended for production.
                    .dedicatedMasterType("t2.small.search")
                    .instanceType("t2.small.search")
                    .instanceCount(5)
                    .build();

            // Many instance types require EBS storage.
            EBSOptions ebsOptions = EBSOptions.builder()
                    .ebsEnabled(true)
                    .volumeSize(10)
                    .volumeType("gp2")
                    .build();

            NodeToNodeEncryptionOptions encryptionOptions =
 NodeToNodeEncryptionOptions.builder()
                    .enabled(true)
                    .build();

            CreateDomainRequest createRequest = CreateDomainRequest.builder()
                    .domainName(domainName)
                    .engineVersion("OpenSearch_1.0")
                    .clusterConfig(clusterConfig)
                    .ebsOptions(ebsOptions)
                    .nodeToNodeEncryptionOptions(encryptionOptions)
                    // You can uncomment this line and add your account ID, a
 username, and the
                    // domain name to add an access policy.
                    // .accessPolicies("{\"Version\":\"2012-10-17\",
\"Statement\":[{\"Effect\":\"Allow\",\"Principal\":{\"AWS\":
[\"arn:aws:iam::123456789012:user/user-name\"]},\"Action\":[\"es:*\"],\"Resource\":
\"arn:aws:es:region:123456789012:domain/domain-name/*\"}]}")
                    .build();

            // Make the request.
            System.out.println("Sending domain creation request...");
            CreateDomainResponse createResponse =
 client.createDomain(createRequest);
```

```
            System.out.println("Domain status:
"+createResponse.domainStatus().toString());
            System.out.println("Domain ID:
"+createResponse.domainStatus().domainId());


        } catch (OpenSearchException e) {
            System.err.println(e.awsErrorDetails().errorMessage());
            System.exit(1);
        }
    }

    /**
     * Updates the configuration of an Amazon OpenSearch Service domain with the
     * specified options. Some options require other Amazon Web Services resources,
such as an
     * Amazon Cognito user pool and identity pool, whereas others require just an
     * instance type or instance count.
     *
     * @param client
     *             The client to use for the requests to Amazon OpenSearch Service
     * @param domainName
     *             The name of the domain to update
     */

    public static void updateDomain(OpenSearchClient client, String domainName) {

     // Updates the domain to use three data instances instead of five.
        // You can uncomment the Cognito line and fill in the strings to enable
Cognito
        // authentication for OpenSearch Dashboards.

        try {

            ClusterConfig clusterConfig = ClusterConfig.builder()
                    .instanceCount(5)
                    .build();

            CognitoOptions cognitoOptions = CognitoOptions.builder()
                    .enabled(true)
                    .userPoolId("user-pool-id")
                    .identityPoolId("identity-pool-id")
                    .roleArn("role-arn")
                    .build();
```

```
            UpdateDomainConfigRequest updateRequest =
UpdateDomainConfigRequest.builder()
                    .domainName(domainName)
                    .clusterConfig(clusterConfig)
                    //.cognitoOptions(cognitoOptions)
                    .build();

            System.out.println("Sending domain update request...");
            UpdateDomainConfigResponse updateResponse =
client.updateDomainConfig(updateRequest);
            System.out.println("Domain config:
"+updateResponse.domainConfig().toString());


        } catch (OpenSearchException e) {
            System.err.println(e.awsErrorDetails().errorMessage());
            System.exit(1);
        }
    }

    /**
     * Deletes an Amazon OpenSearch Service domain. Deleting a domain can take
     * several minutes.
     *
     * @param client
     *             The client to use for the requests to Amazon OpenSearch Service
     * @param domainName
     *             The name of the domain that you want to delete
     */

    public static void deleteDomain(OpenSearchClient client, String domainName) {

        try {

            DeleteDomainRequest deleteRequest = DeleteDomainRequest.builder()
                    .domainName(domainName)
                    .build();

            System.out.println("Sending domain deletion request...");
            DeleteDomainResponse deleteResponse =
client.deleteDomain(deleteRequest);
            System.out.println("Domain status: "+deleteResponse.toString());
```

```
        } catch (OpenSearchException e) {
            System.err.println(e.awsErrorDetails().errorMessage());
            System.exit(1);
        }
    }

    /**
     * Waits for the domain to finish processing changes. New domains typically take
15-30 minutes
     * to initialize, but can take longer depending on the configuration. Most
updates to existing domains
     * take a similar amount of time. This method checks every 15 seconds and
finishes only when
     * the domain's processing status changes to false.
     *
     * @param client
     *            The client to use for the requests to Amazon OpenSearch Service
     * @param domainName
     *            The name of the domain that you want to check
     */

    public static void waitForDomainProcessing(OpenSearchClient client, String
domainName) {
        // Create a new request to check the domain status.
        DescribeDomainRequest describeRequest = DescribeDomainRequest.builder()
                .domainName(domainName)
                .build();

        // Every 15 seconds, check whether the domain is processing.
        DescribeDomainResponse describeResponse =
client.describeDomain(describeRequest);
        while (describeResponse.domainStatus().processing()) {
            try {
                System.out.println("Domain still processing...");
                TimeUnit.SECONDS.sleep(15);
                describeResponse = client.describeDomain(describeRequest);
            } catch (InterruptedException e) {
                e.printStackTrace();
            }
        }

        // Once we exit that loop, the domain is available
```

```
        System.out.println("Amazon OpenSearch Service has finished processing
 changes for your domain.");
        System.out.println("Domain description: "+describeResponse.toString());
    }
}
```

Version 1

Questo esempio utilizza il AWSElasticsearchClientBuildercostruttore della versione 1 di AWS SDK per Java per creare un dominio Elasticsearch legacy, aggiornarne la configurazione ed eliminarlo. Rimuovere il commento dalle chiamate a `waitForDomainProcessing` (e commentare la chiamata a `deleteDomain`) per fare in modo che il dominio sia online e utilizzabile.

```
package com.amazonaws.samples;

import java.util.concurrent.TimeUnit;
import com.amazonaws.auth.DefaultAWSCredentialsProviderChain;
import com.amazonaws.regions.Regions;
import com.amazonaws.services.elasticsearch.AWSElasticsearch;
import com.amazonaws.services.elasticsearch.AWSElasticsearchClientBuilder;
import com.amazonaws.services.elasticsearch.model.CreateElasticsearchDomainRequest;
import com.amazonaws.services.elasticsearch.model.CreateElasticsearchDomainResult;
import com.amazonaws.services.elasticsearch.model.DeleteElasticsearchDomainRequest;
import com.amazonaws.services.elasticsearch.model.DeleteElasticsearchDomainResult;
import
 com.amazonaws.services.elasticsearch.model.DescribeElasticsearchDomainRequest;
import com.amazonaws.services.elasticsearch.model.DescribeElasticsearchDomainResult;
import com.amazonaws.services.elasticsearch.model.EBSOptions;
import com.amazonaws.services.elasticsearch.model.ElasticsearchClusterConfig;
import com.amazonaws.services.elasticsearch.model.ResourceNotFoundException;
import
 com.amazonaws.services.elasticsearch.model.UpdateElasticsearchDomainConfigRequest;
import
 com.amazonaws.services.elasticsearch.model.UpdateElasticsearchDomainConfigResult;
import com.amazonaws.services.elasticsearch.model.VolumeType;

/**
 * Sample class demonstrating how to use the Amazon Web Services SDK for Java to
 create, update,
 * and delete Amazon OpenSearch Service domains.
 */

public class OpenSearchSample {
```

```
   public static void main(String[] args) {

       final String domainName = "my-test-domain";

       // Build the client using the default credentials chain.
       // You can use the CLI and run `aws configure` to set access key, secret
       // key, and default region.
       final AWSElasticsearch client = AWSElasticsearchClientBuilder
               .standard()
               // Unnecessary, but lets you use a region different than your
default.
               .withRegion(Regions.US_WEST_2)
               // Unnecessary, but if desired, you can use a different provider
chain.
               .withCredentials(new DefaultAWSCredentialsProviderChain())
               .build();

       // Create a new domain, update its configuration, and delete it.
       createDomain(client, domainName);
       // waitForDomainProcessing(client, domainName);
       updateDomain(client, domainName);
       // waitForDomainProcessing(client, domainName);
       deleteDomain(client, domainName);
   }

   /**
    * Creates an Amazon OpenSearch Service domain with the specified options.
    * Some options require other Amazon Web Services resources, such as an Amazon
Cognito user pool
    * and identity pool, whereas others require just an instance type or instance
    * count.
    *
    * @param client
    *            The client to use for the requests to Amazon OpenSearch Service
    * @param domainName
    *            The name of the domain you want to create
    */
   private static void createDomain(final AWSElasticsearch client, final String
domainName) {

       // Create the request and set the desired configuration options
       CreateElasticsearchDomainRequest createRequest = new
CreateElasticsearchDomainRequest()
```

```
                .withDomainName(domainName)
                .withElasticsearchVersion("7.10")
                .withElasticsearchClusterConfig(new ElasticsearchClusterConfig()
                        .withDedicatedMasterEnabled(true)
                        .withDedicatedMasterCount(3)
                        // Small, inexpensive instance types for testing. Not
 recommended for production
                        // domains.
                        .withDedicatedMasterType("t2.small.elasticsearch")
                        .withInstanceType("t2.small.elasticsearch")
                        .withInstanceCount(5))
                // Many instance types require EBS storage.
                .withEBSOptions(new EBSOptions()
                        .withEBSEnabled(true)
                        .withVolumeSize(10)
                        .withVolumeType(VolumeType.Gp2));
                // You can uncomment this line and add your account ID, a username,
 and the
                // domain name to add an access policy.
                // .withAccessPolicies("{\"Version\":\"2012-10-17\",
\"Statement\":[{\"Effect\":\"Allow\",\"Principal\":{\"AWS\":
[\"arn:aws:iam::123456789012:user/user-name\"]},\"Action\":[\"es:*\"],\"Resource\":
\"arn:aws:es:region:123456789012:domain/domain-name/*\"}]}")


        // Make the request.
        System.out.println("Sending domain creation request...");
        CreateElasticsearchDomainResult createResponse =
 client.createElasticsearchDomain(createRequest);
        System.out.println("Domain creation response from Amazon OpenSearch
 Service:");
        System.out.println(createResponse.getDomainStatus().toString());
    }

    /**
     * Updates the configuration of an Amazon OpenSearch Service domain with the
     * specified options. Some options require other Amazon Web Services resources,
 such as an
     * Amazon Cognito user pool and identity pool, whereas others require just an
     * instance type or instance count.
     *
     * @param client
     *            The client to use for the requests to Amazon OpenSearch Service
     * @param domainName
```

```
     *            The name of the domain to update
     */
    private static void updateDomain(final AWSElasticsearch client, final String
domainName) {
        try {
            // Updates the domain to use three data instances instead of five.
            // You can uncomment the Cognito lines and fill in the strings to enable
Cognito
            // authentication for OpenSearch Dashboards.
            final UpdateElasticsearchDomainConfigRequest updateRequest = new
UpdateElasticsearchDomainConfigRequest()
                    .withDomainName(domainName)
                    // .withCognitoOptions(new CognitoOptions()
                            // .withEnabled(true)
                            // .withUserPoolId("user-pool-id")
                            // .withIdentityPoolId("identity-pool-id")
                            // .withRoleArn("role-arn")
                    .withElasticsearchClusterConfig(new ElasticsearchClusterConfig()
                            .withInstanceCount(3));

            System.out.println("Sending domain update request...");
            final UpdateElasticsearchDomainConfigResult updateResponse = client
                    .updateElasticsearchDomainConfig(updateRequest);
            System.out.println("Domain update response from Amazon OpenSearch
Service:");
            System.out.println(updateResponse.toString());
        } catch (ResourceNotFoundException e) {
            System.out.println("Domain not found. Please check the domain name.");
        }
    }

    /**
     * Deletes an Amazon OpenSearch Service domain. Deleting a domain can take
     * several minutes.
     *
     * @param client
     *            The client to use for the requests to Amazon OpenSearch Service
     * @param domainName
     *            The name of the domain that you want to delete
     */
    private static void deleteDomain(final AWSElasticsearch client, final String
domainName) {
        try {
```

```
            final DeleteElasticsearchDomainRequest deleteRequest = new
DeleteElasticsearchDomainRequest()
                    .withDomainName(domainName);

            System.out.println("Sending domain deletion request...");
            final DeleteElasticsearchDomainResult deleteResponse =
client.deleteElasticsearchDomain(deleteRequest);
            System.out.println("Domain deletion response from Amazon OpenSearch
Service:");
            System.out.println(deleteResponse.toString());
        } catch (ResourceNotFoundException e) {
            System.out.println("Domain not found. Please check the domain name.");
        }
    }

    /**
     * Waits for the domain to finish processing changes. New domains typically take
15-30 minutes
     * to initialize, but can take longer depending on the configuration. Most
updates to existing domains
     * take a similar amount of time. This method checks every 15 seconds and
finishes only when
     * the domain's processing status changes to false.
     *
     * @param client
     *             The client to use for the requests to Amazon OpenSearch Service
     * @param domainName
     *             The name of the domain that you want to check
     */
    private static void waitForDomainProcessing(final AWSElasticsearch client, final
String domainName) {
        // Create a new request to check the domain status.
        final DescribeElasticsearchDomainRequest describeRequest = new
DescribeElasticsearchDomainRequest()
                .withDomainName(domainName);

        // Every 15 seconds, check whether the domain is processing.
        DescribeElasticsearchDomainResult describeResponse =
client.describeElasticsearchDomain(describeRequest);
        while (describeResponse.getDomainStatus().isProcessing()) {
            try {
                System.out.println("Domain still processing...");
                TimeUnit.SECONDS.sleep(15);
```

```
                describeResponse =
 client.describeElasticsearchDomain(describeRequest);
            } catch (InterruptedException e) {
                e.printStackTrace();
            }
        }

        // Once we exit that loop, the domain is available
        System.out.println("Amazon OpenSearch Service has finished processing
 changes for your domain.");
        System.out.println("Domain description response from Amazon OpenSearch
 Service:");
        System.out.println(describeResponse.toString());
    }
}
```

Python

Questo esempio utilizza il client Python di OpenSearchServicebasso livello di AWS SDK per Python (Boto) per creare un dominio, aggiornarne la configurazione ed eliminarlo.

```
import boto3
import botocore
from botocore.config import Config
import time

# Build the client using the default credential configuration.
# You can use the CLI and run 'aws configure' to set access key, secret
# key, and default region.

my_config = Config(
    # Optionally lets you specify a region other than your default.
    region_name='us-west-2'
)

client = boto3.client('opensearch', config=my_config)

domainName = 'my-test-domain'  # The name of the domain


def createDomain(client, domainName):
    """Creates an Amazon OpenSearch Service domain with the specified options."""
```

```
    response = client.create_domain(
        DomainName=domainName,
        EngineVersion='OpenSearch_1.0',
        ClusterConfig={
            'InstanceType': 't2.small.search',
            'InstanceCount': 5,
            'DedicatedMasterEnabled': True,
            'DedicatedMasterType': 't2.small.search',
            'DedicatedMasterCount': 3
        },
        # Many instance types require EBS storage.
        EBSOptions={
            'EBSEnabled': True,
            'VolumeType': 'gp2',
            'VolumeSize': 10
        },
        AccessPolicies="{\"Version\":\"2012-10-17\",\"Statement\":[{\"Effect\":\"Allow
\",\"Principal\":{\"AWS\":[\"arn:aws:iam::123456789012:user/user-name\"]},\"Action\":
[\"es:*\"],\"Resource\":\"arn:aws:es:us-west-2:123456789012:domain/my-test-domain/*
\"}]}",
        NodeToNodeEncryptionOptions={
            'Enabled': True
        }
    )
    print("Creating domain...")
    print(response)


def updateDomain(client, domainName):
    """Updates the domain to use three data nodes instead of five."""
    try:
        response = client.update_domain_config(
            DomainName=domainName,
            ClusterConfig={
                'InstanceCount': 3
            }
        )
        print('Sending domain update request...')
        print(response)

    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ResourceNotFoundException':
            print('Domain not found. Please check the domain name.')
        else:
```

```
            raise error


def deleteDomain(client, domainName):
    """Deletes an OpenSearch Service domain. Deleting a domain can take several
 minutes."""
    try:
        response = client.delete_domain(
            DomainName=domainName
        )
        print('Sending domain deletion request...')
        print(response)

    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ResourceNotFoundException':
            print('Domain not found. Please check the domain name.')
        else:
            raise error


def waitForDomainProcessing(client, domainName):
    """Waits for the domain to finish processing changes."""
    try:
        response = client.describe_domain(
            DomainName=domainName
        )
        # Every 15 seconds, check whether the domain is processing.
        while response["DomainStatus"]["Processing"] == True:
            print('Domain still processing...')
            time.sleep(15)
            response = client.describe_domain(
                DomainName=domainName)

        # Once we exit the loop, the domain is available.
        print('Amazon OpenSearch Service has finished processing changes for your
 domain.')
        print('Domain description:')
        print(response)

    except botocore.exceptions.ClientError as error:
        if error.response['Error']['Code'] == 'ResourceNotFoundException':
            print('Domain not found. Please check the domain name.')
        else:
            raise error
```

```
def main():
    """Create a new domain, update its configuration, and delete it."""
    createDomain(client, domainName)
    waitForDomainProcessing(client, domainName)
    updateDomain(client, domainName)
    waitForDomainProcessing(client, domainName)
    deleteDomain(client, domainName)
```

Nodo

Questo esempio utilizza la versione 3 del OpenSearch client SDK for JavaScript in Node.js per creare un dominio, aggiornarne la configurazione ed eliminarlo.

```
var {
    OpenSearchClient,
    CreateDomainCommand,
    DescribeDomainCommand,
    UpdateDomainConfigCommand,
    DeleteDomainCommand
} = require("@aws-sdk/client-opensearch");
var sleep = require('sleep');

var client = new OpenSearchClient();

var domainName = 'my-test-domain'

// Create a new domain, update its configuration, and delete it.
createDomain(client, domainName)
waitForDomainProcessing(client, domainName)
updateDomain(client, domainName)
waitForDomainProcessing(client, domainName)
deleteDomain(client, domainName)

async function createDomain(client, domainName) {
    // Creates an Amazon OpenSearch Service domain with the specified options.
    var command = new CreateDomainCommand({
        DomainName: domainName,
        EngineVersion: 'OpenSearch_1.0',
        ClusterConfig: {
        'InstanceType': 't2.small.search',
        'InstanceCount': 5,
```

```
        'DedicatedMasterEnabled': 'True',
        'DedicatedMasterType': 't2.small.search',
        'DedicatedMasterCount': 3
        },
        EBSOptions:{
            'EBSEnabled': 'True',
            'VolumeType': 'gp2',
            'VolumeSize': 10
        },
        AccessPolicies: "{\"Version\":\"2012-10-17\",\"Statement\":[{\"Effect\":\"Allow
\",\"Principal\":{\"AWS\":[\"arn:aws:iam::123456789012:user/user-name\"]},\"Action\":
[\"es:*\"],\"Resource\":\"arn:aws:es:us-east-1:123456789012:domain/my-test-domain/*
\"}]}",
         NodeToNodeEncryptionOptions:{
            'Enabled': 'True'
        }
    });
    const response = await client.send(command);
    console.log("Creating domain...");
    console.log(response);
}

async function updateDomain(client, domainName) {
    // Updates the domain to use three data nodes instead of five.
    var command = new UpdateDomainConfigCommand({
        DomainName: domainName,
        ClusterConfig: {
        'InstanceCount': 3
        }
    });
    const response = await client.send(command);
    console.log('Sending domain update request...');
    console.log(response);
}

async function deleteDomain(client, domainName) {
    // Deletes an OpenSearch Service domain. Deleting a domain can take several
 minutes.
    var command = new DeleteDomainCommand({
        DomainName: domainName
    });
    const response = await client.send(command);
    console.log('Sending domain deletion request...');
    console.log(response);
```

```
}

async function waitForDomainProcessing(client, domainName) {
    // Waits for the domain to finish processing changes.
    try {
        var command = new DescribeDomainCommand({
            DomainName: domainName
        });
        var response = await client.send(command);

        while (response.DomainStatus.Processing == true) {
            console.log('Domain still processing...')
            await sleep(15000) // Wait for 15 seconds, then check the status again
            function sleep(ms) {
                return new Promise((resolve) => {
                    setTimeout(resolve, ms);
                });
            }
            var response = await client.send(command);
        }
        // Once we exit the loop, the domain is available.
        console.log('Amazon OpenSearch Service has finished processing changes for your
 domain.');
        console.log('Domain description:');
        console.log(response);

    } catch (error) {
        if (error.name === 'ResourceNotFoundException') {
            console.log('Domain not found. Please check the domain name.');
            }
    };
}
```

Indicizzazione dei dati in Amazon Service OpenSearch

Poiché Amazon OpenSearch Service utilizza un'API REST, esistono numerosi metodi per indicizzare i documenti. Puoi utilizzare client standard come curl o qualsiasi linguaggio di programmazione in grado di inviare richieste HTTP. Per semplificare ulteriormente il processo di interazione con essa, OpenSearch Service dispone di client per molti linguaggi di programmazione. Gli utenti esperti possono passare direttamente a the section called “Caricamento di dati di streaming in OpenSearch Service”.

Ti consigliamo vivamente di utilizzare Amazon OpenSearch Ingestion per importare dati, un raccoglitore di dati completamente gestito integrato in Service. OpenSearch Per ulteriori informazioni, consulta Amazon OpenSearch Ingestion.

Per un'introduzione all'indicizzazione, consulta la documentazione. OpenSearch

Limitazioni di denominazione per gli indici

OpenSearch Gli indici dei servizi presentano le seguenti restrizioni di denominazione:

- Tutte le lettere devono essere minuscole.
- I nomi degli indici non possono iniziare con `_` o `-`.
- I nomi degli indici non possono contenere spazi, virgole, `:`, `"`, `*`, `+`, `/`, `\`, `|`, `?`, `#`, `>` o `<`.

Non includere informazioni riservate nei nomi degli indici, dei tipi o degli ID dei documenti. OpenSearch Il servizio utilizza questi nomi nei suoi Uniform Resource Identifiers (URIs). I server e le applicazioni spesso registrano le richieste HTTP, il che può portare a un'esposizione non necessaria dei dati se URIs contengono informazioni sensibili:

```
2018-10-03T23:39:43 198.51.100.14 200 "GET https://opensearch-domain/dr-jane-doe/flu-patients-2018/202-555-0100/ HTTP/1.1"
```

Anche se non si dispone delle autorizzazioni per visualizzare il documento JSON associato, è possibile dedurre da questa riga fittizia del log che uno dei pazienti del Dr. Doe con numero di telefono 202-555-0100 ha avuto l'influenza nel 2018.

Se OpenSearch Service rileva un indirizzo IP reale o percepito in un nome di indice (ad esempio,`my-index-12.34.56.78.91`), maschera l'indirizzo IP. Una chiamata a `_cat/indices` produce la risposta seguente:

```
green open my-index-x.x.x.x.91    soY19tBERoKo71WcEScidw 5 1 0 0   2kb  1kb
```

Per evitare inutili confusioni, evitare di includere indirizzi IP nei nomi di indice.

Riduzione delle dimensioni della risposta

Le risposte provenienti da `_index` e `_bulk` APIs contengono parecchie informazioni. Queste informazioni possono essere utili per le richieste di risoluzione dei problemi o per l'implementazione della logica dei tentativi, ma possono utilizzare una larghezza di banda considerevole. In questo esempio, l'indicizzazione di un documento a 32 byte genera una risposta a 339 byte (incluse le intestazioni):

```
PUT opensearch-domain/more-movies/_doc/1
{"title": "Back to the Future"}
```

Risposta

```
{
  "_index": "more-movies",
  "_type": "_doc",
  "_id": "1",
  "_version": 4,
  "result": "updated",
  "_shards": {
    "total": 2,
    "successful": 2,
    "failed": 0
  },
  "_seq_no": 3,
  "_primary_term": 1
}
```

Questa dimensione della risposta potrebbe sembrare minima, ma se si indicizza 1.000.000 di documenti al giorno, circa 11,5 documenti al secondo, 339 byte per risposta sono 10,17 GB di traffico di download al mese.

Se i costi di trasferimento dei dati sono un problema, utilizzate il `filter_path` parametro per ridurre le dimensioni della risposta del OpenSearch Servizio, ma fate attenzione a non filtrare i campi necessari per identificare o riprovare le richieste non riuscite. Questi campi possono variare a seconda del client. Il `filter_path` parametro funziona per tutti i OpenSearch Service REST APIs, ma è particolarmente utile APIs in caso di chiamate frequenti, ad esempio `_index` e `_bulk` APIs:

```
PUT opensearch-domain/more-movies/_doc/1?filter_path=result,_shards.total
{"title": "Back to the Future"}
```

Risposta

```
{
  "result": "updated",
  "_shards": {
    "total": 2
  }
}
```

Invece di includere i campi, puoi escluderli con un prefisso `-`. `filter_path` supporta anche i caratteri jolly:

```
POST opensearch-domain/_bulk?filter_path=-took,-items.index._*
{ "index": { "_index": "more-movies", "_id": "1" } }
{"title": "Back to the Future"}
{ "index": { "_index": "more-movies", "_id": "2" } }
{"title": "Spirited Away"}
```

Risposta

```
{
  "errors": false,
  "items": [
    {
      "index": {
        "result": "updated",
        "status": 200
      }
    },
    {
      "index": {
```

```
        "result": "updated",
        "status": 200
      }
    }
  ]
}
```

Codec indicizzati

I codec di indice determinano il modo in cui i campi memorizzati in un indice vengono compressi e archiviati su disco. Il codec dell'indice è controllato dall'`index.codec`impostazione statica, che specifica l'algoritmo di compressione. Questa impostazione influisce sulla dimensione del frammento di indice e sulle prestazioni operative.

Per un elenco dei codec supportati e delle relative caratteristiche prestazionali, consultate Codec supportati nella documentazione. OpenSearch

Quando scegliete un codec di indice, tenete presente quanto segue:

- Per evitare i problemi legati alla modifica dell'impostazione del codec di un indice esistente, testate un carico di lavoro rappresentativo in un ambiente non di produzione prima di utilizzare una nuova impostazione del codec. Per ulteriori informazioni, consultate Modifica di un codec di indice.
- Non puoi utilizzare i codec di compressione Zstandard (`"index.codec": "zstd"o"index.codec": "zstd_no_dict"`) per gli indici k-NN o Security Analytics.

Caricamento di dati di streaming in Amazon OpenSearch Service

Puoi utilizzare OpenSearch Ingestion per caricare direttamente i dati di streaming nel tuo dominio Amazon OpenSearch Service, senza dover utilizzare soluzioni di terze parti. Per inviare dati a OpenSearch Ingestion, configuri i produttori di dati e il servizio consegna automaticamente i dati al dominio o alla raccolta specificati. Per iniziare a usare OpenSearch Ingestion, consulta. the section called “Tutorial: inserisci dati in una raccolta”

Puoi comunque utilizzare altre fonti per caricare dati in streaming, come Amazon Data Firehose e Amazon CloudWatch Logs, che dispongono del supporto integrato per Service. OpenSearch Altre, come Amazon S3, Amazon Kinesis Data Streams e Amazon DynamoDB, utilizzano le funzioni AWS Lambda come gestori di eventi. Le funzioni Lambda rispondono ai nuovi dati elaborandoli ed eseguendone lo streaming nel dominio.

> **Note**
>
> Lambda supporta diversi linguaggi di programmazione tra i più diffusi ed è disponibile in gran parte delle Regioni AWS. Per ulteriori informazioni, consulta Getting started with Lambda nella AWS Lambda Developer Guide e AWS Service Endpoints nel. Riferimenti generali di AWS

Caricamento di dati di streaming da Ingestion OpenSearch

Puoi usare Amazon OpenSearch Ingestion per caricare dati in un dominio di OpenSearch servizio. Puoi configurare i produttori di dati per inviare dati a OpenSearch Ingestion, che li invia automaticamente alla raccolta specificata. Puoi anche configurare OpenSearch Ingestion per trasformare i dati prima di consegnarli. Per ulteriori informazioni, consulta OpenSearch Ingestione di Amazon.

Caricamento di dati in streaming da Amazon S3

Puoi usare Lambda per inviare dati al tuo dominio di OpenSearch servizio da Amazon S3. I nuovi dati che arrivano in un bucket S3 attivano una notifica eventi per Lambda, che quindi esegue il codice personalizzato per eseguire l'indicizzazione.

Questo metodo per lo streaming dei dati è estremamente flessibile. Puoi indicizzare i metadati degli oggetti oppure, se l'oggetto è un testo normale, analizzare e indicizzare alcuni elementi del corpo dell'oggetto. Questa sezione include alcuni semplici codici Python di esempio in cui sono utilizzate espressioni regolari per analizzare un file di log e indicizzare le corrispondenze.

Prerequisiti

Prima di procedere, devi disporre delle risorse indicate di seguito.

Prerequisito	Description
Bucket Amazon S3	Per ulteriori informazioni, consulta Creazione del primo bucket S3 nella Guida per l'utente di Amazon Simple Storage Service. Il bucket deve risiedere nella stessa regione del dominio di servizio. OpenSearch

Prerequisito	Description
OpenSearch Dominio di servizio	La destinazione dei dati dopo che la funzione Lambda li ha elaborati . Per ulteriori informazioni, consultare the section called “Creazione OpenSearch di domini di servizio”.

Creazione il pacchetto di implementazione Lambda

I pacchetti di distribuzione sono file ZIP o JAR che includono codice ed eventuali dipendenze. In questa sezione è incluso codice di esempio Python. Per altri linguaggi di programmazione, consultare Pacchetti di implementazione Lambda nella Guida per gli sviluppatori di AWS Lambda .

1. Crea una directory. In questo esempio utilizziamo il nome `s3-to-opensearch`.
2. Creare un file nella directory denominata `sample.py`:

```
import boto3
import re
import requests
from requests_aws4auth import AWS4Auth

region = '' # e.g. us-west-1
service = 'es'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

host = '' # the OpenSearch Service domain, e.g. https://search-mydomain.us-
west-1.es.amazonaws.com
index = 'lambda-s3-index'
datatype = '_doc'
url = host + '/' + index + '/' + datatype

headers = { "Content-Type": "application/json" }

s3 = boto3.client('s3')

# Regular expressions used to parse some simple log lines
ip_pattern = re.compile('(\d+\.\d+\.\d+\.\d+)')
time_pattern = re.compile('\[(\d+\/\w\w\w\/\d\d\d\d:\d\d:\d\d:\d\d\s-\d\d\d\d)\]')
message_pattern = re.compile('\"(.+)\"')
```

```
# Lambda execution starts here
def handler(event, context):
    for record in event['Records']:

        # Get the bucket name and key for the new file
        bucket = record['s3']['bucket']['name']
        key = record['s3']['object']['key']

        # Get, read, and split the file into lines
        obj = s3.get_object(Bucket=bucket, Key=key)
        body = obj['Body'].read()
        lines = body.splitlines()

        # Match the regular expressions to each line and index the JSON
        for line in lines:
            line = line.decode("utf-8")
            ip = ip_pattern.search(line).group(1)
            timestamp = time_pattern.search(line).group(1)
            message = message_pattern.search(line).group(1)

            document = { "ip": ip, "timestamp": timestamp, "message": message }
            r = requests.post(url, auth=awsauth, json=document, headers=headers)
```

Modifica le variabili per `region` e `host`.

3. Installare pip, se non è già stato fatto, quindi installare le dipendenze in una nuova directory `package`:

```
cd s3-to-opensearch

pip install --target ./package requests
pip install --target ./package requests_aws4auth
```

In tutti gli ambienti di esecuzione Lambda è installato Boto3, perciò non è necessario includerlo nel pacchetto di implementazione.

4. Crea un pacchetto con il codice dell'applicazione e le dipendenze:

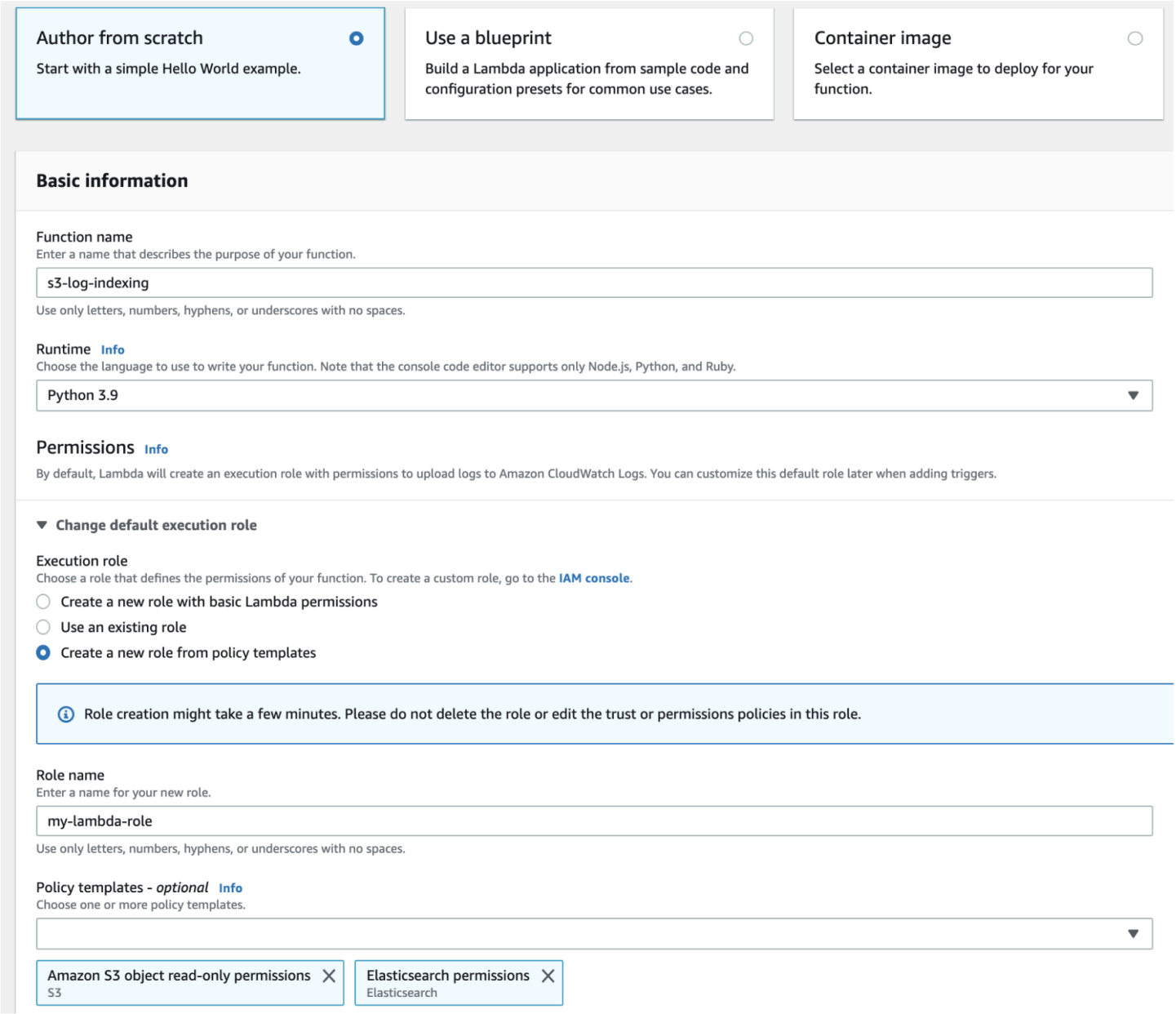
```
cd package
zip -r ../lambda.zip .

cd ..
```

```
zip -g lambda.zip sample.py
```

Creazione della funzione Lambda

Dopo aver creato il pacchetto di implementazione, è possibile creare la funzione Lambda. Quando si crea una funzione, scegliere nome, runtime (ad esempio, Python 3.8) e ruolo IAM. Il ruolo IAM definisce le autorizzazioni per la tua funzione. Per istruzioni dettagliate, consultare Creazione di una funzione Lambda con la console nella Guida per gli sviluppatori di AWS Lambda .

Questo esempio presuppone l'utilizzo della console. Scegli Python 3.9 e un ruolo con autorizzazioni di lettura S3 e autorizzazioni di scrittura del OpenSearch servizio, come mostrato nella schermata seguente:



Una volta creata la funzione, devi aggiungere un trigger. In questo esempio, vogliamo che il codice venga eseguito ogni volta che un file di log arriva in un bucket S3:

1. Scegliere Aggiungi trigger e selezionare S3.
2. Scegli il bucket.
3. Per Event type (Tipo di evento), seleziona PUT.
4. In Prefix (Prefisso), digita `logs/`.
5. Per Suffisso, digitare `.log`.
6. Confermare l'avviso di chiamata ricorsiva e scegliere Aggiungi.

Puoi infine caricare il pacchetto di implementazione:

1. Scegliere Carica da e File .zip, quindi seguire i prompt su schermo per caricare il pacchetto di implementazione.
2. Al termine del caricamento, modificare il campo Impostazioni runtime e cambiare il gestore in `sample.handler`. Questa impostazione indica a Lambda il file (`sample.py`) e il metodo (`handler`) da eseguire dopo un trigger.

A questo punto, hai un set completo di risorse: un bucket per i file di registro, una funzione che viene eseguita ogni volta che viene aggiunto un file di registro al bucket, codice che esegue l'analisi e l'indicizzazione e un dominio di servizio per la ricerca e la visualizzazione. OpenSearch

Test della funzione Lambda

Una volta creata la funzione, è possibile eseguirne il test caricando un file nel bucket Amazon S3. Crea un file denominato `sample.log` utilizzando le righe di log di esempio indicate di seguito:

```
12.345.678.90 - [10/Oct/2000:13:55:36 -0700] "PUT /some-file.jpg"
12.345.678.91 - [10/Oct/2000:14:56:14 -0700] "GET /some-file.jpg"
```

Carica il file nella cartella `logs` del bucket S3. Per le istruzioni, consulta Caricamento di un oggetto nel bucket nella Guida per l'utente di Amazon Simple Storage Service.

Utilizza quindi la console di OpenSearch servizio o le OpenSearch dashboard per verificare che l'indice contenga due documenti. `lambda-s3-index` Puoi anche effettuare una richiesta di ricerca standard:

```
GET https://domain-name/lambda-s3-index/_search?pretty
{
  "hits" : {
    "total" : 2,
    "max_score" : 1.0,
    "hits" : [
      {
        "_index" : "lambda-s3-index",
        "_type" : "_doc",
        "_id" : "vTYXaWIBJWV_TTkEuSDg",
        "_score" : 1.0,
        "_source" : {
          "ip" : "12.345.678.91",
```

```
          "message" : "GET /some-file.jpg",
          "timestamp" : "10/Oct/2000:14:56:14 -0700"
        }
      },
      {
        "_index" : "lambda-s3-index",
        "_type" : "_doc",
        "_id" : "vjYmaWIBJWV_TTkEuCAB",
        "_score" : 1.0,
        "_source" : {
          "ip" : "12.345.678.90",
          "message" : "PUT /some-file.jpg",
          "timestamp" : "10/Oct/2000:13:55:36 -0700"
        }
      }
    ]
  }
}
```

Caricamento dei dati in streaming in Amazon Kinesis Data Streams

È possibile caricare dati di streaming da Kinesis Data OpenSearch Streams to Service. I nuovi dati che arrivano nel flusso di dati attivano una notifica eventi per Lambda, che quindi esegue il codice personalizzato per eseguire l'indicizzazione. In questa sezione è incluso un semplice codice di esempio Python.

Prerequisiti

Prima di procedere, devi disporre delle risorse indicate di seguito.

Prerequisito	Description
Amazon Kinesis Data Streams	L'origine dell'evento per la funzione Lambda. Per ulteriori informazioni, consultare Kinesis Data Streams.
OpenSearch Dominio di servizio	La destinazione dei dati dopo che la funzione Lambda li ha elaborati . Per ulteriori informazioni, consultare the section called “Creazione OpenSearch di domini di servizio”
Ruolo IAM	Questo ruolo deve avere le autorizzazioni di base OpenSearch Service, Kinesis e Lambda, come le seguenti:

Prerequisito	Description

	JSON
	
	{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": [
	"es:ESHttpPost",
	"es:ESHttpPut",
	"logs:CreateLogGroup",
	"logs:CreateLogStream",
	"logs:PutLogEvents",
	"kinesis:GetShardIterator",
	"kinesis:GetRecords",
	"kinesis:DescribeStream",
	"kinesis:ListStreams"
	],
	"Resource": "*"
	}
	]
	}
	
	Il ruolo deve avere la relazione di trust seguente:
	
	JSON
	
	{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Principal": {
	"Service": "lambda.amazonaws.com"
	},
	"Action": "sts:AssumeRole"
	}
	]

Prerequisito	Description
	} Per ulteriori informazioni, consultare Creazione di ruoli IAM nella Guida per l'utente di IAM.

Creazione della funzione Lambda

Procedi come descritto in the section called “Creazione il pacchetto di implementazione Lambda”, ma crea una directory denominata `kinesis-to-opensearch` e utilizza il codice seguente per `sample.py`:

```
import base64
import boto3
import json
import requests
from requests_aws4auth import AWS4Auth

region = '' # e.g. us-west-1
service = 'es'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

host = '' # the OpenSearch Service domain, e.g. https://search-mydomain.us-
west-1.es.amazonaws.com
index = 'lambda-kine-index'
datatype = '_doc'
url = host + '/' + index + '/' + datatype + '/'

headers = { "Content-Type": "application/json" }

def handler(event, context):
    count = 0
    for record in event['Records']:
        id = record['eventID']
        timestamp = record['kinesis']['approximateArrivalTimestamp']

        # Kinesis data is base64-encoded, so decode here
        message = base64.b64decode(record['kinesis']['data'])
```

```
        # Create the JSON document
        document = { "id": id, "timestamp": timestamp, "message": message }
        # Index the document
        r = requests.put(url + id, auth=awsauth, json=document, headers=headers)
        count += 1
    return 'Processed ' + str(count) + ' items.'
```

Modifica le variabili per `region` e `host`.

Installare pip, se non è già stato fatto, quindi utilizzare i seguenti comandi per installare le dipendenze:

```
cd kinesis-to-opensearch

pip install --target ./package requests
pip install --target ./package requests_aws4auth
```

Procedi quindi come descritto in the section called “Creazione della funzione Lambda”, ma specifica il ruolo IAM dai the section called “Prerequisiti” e le impostazioni seguenti per il trigger:

- Flusso Kinesis: il flusso di Kinesis.
- Batch size (Dimensione batch): 100
- Starting position (Posizione di inizio): orizzonte di taglio

Per ulteriori informazioni, consultare Cos'è Amazon Kinesis Data Streams? nella Guida per gli sviluppatori di Amazon Kinesis Data Streams.

A questo punto, hai a disposizione un set completo di risorse: un flusso di dati Kinesis, una funzione che viene eseguita dopo che il flusso riceve nuovi dati e li indicizza e un dominio di OpenSearch servizio per la ricerca e la visualizzazione.

Test della funzione Lambda

Una volta creata la funzione, puoi provarla aggiungendo un nuovo record al flusso di dati utilizzando l' AWS CLI:

```
aws kinesis put-record --stream-name test --data "My test data." --partition-key
 partitionKey1 --region us-west-1
```

Quindi usa la console OpenSearch di servizio o OpenSearch le dashboard per verificare che `lambda-kine-index` contenga un documento. Puoi inoltre utilizzare la seguente richiesta:

```
GET https://domain-name/lambda-kine-index/_search
{
  "hits" : [
    {
      "_index": "lambda-kine-index",
      "_type": "_doc",
      "_id":
 "shardId-000000000000:495835116157626994950129608214214566865294366804960870 42",
      "_score": 1,
      "_source": {
        "timestamp": 1523648740.051,
        "message": "My test data.",
        "id":
 "shardId-000000000000:495835116157626994950129608214214566865294366804960870 42"
      }
    }
  ]
}
```

Caricamento di dati in streaming da una tabella Amazon DynamoDB

Puoi utilizzarlo AWS Lambda per inviare dati al tuo dominio di OpenSearch servizio da Amazon DynamoDB. I nuovi dati che arrivano nella tabella di database attivano una notifica eventi per Lambda, che quindi esegue il codice personalizzato per eseguire l'indicizzazione.

Prerequisiti

Prima di procedere, devi disporre delle risorse indicate di seguito.

Prerequisito	Description
DynamoDB tabella	La tabella contiene i dati di origine. Per ulteriori informazioni, consultare Operazioni di base sulle tabelle DynamoDB nella Guida per gli sviluppat ori di Amazon DynamoDB. La tabella deve risiedere nella stessa regione del dominio di OpenSearc h servizio e avere uno stream impostato su Nuova immagine. Per ulteriori informazioni, consultare Abilitazione di un flusso.

Prerequisito	Description
OpenSearch Dominio di servizio	La destinazione dei dati dopo che la funzione Lambda li ha elaborati . Per ulteriori informazioni, consultare the section called “Creazione OpenSearch di domini di servizio”.

Prerequisito	Description
Ruolo IAM	Questo ruolo deve disporre delle autorizzazioni di esecuzione di base di OpenSearch Service, DynamoDB e Lambda, come le seguenti: JSON `{ "Version":"2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["es:ESHttpPost", "es:ESHttpPut", "dynamodb:DescribeStream", "dynamodb:GetRecords", "dynamodb:GetShardIterator", "dynamodb:ListStreams", "logs:CreateLogGroup", "logs:CreateLogStream", "logs:PutLogEvents"], "Resource": "*" }] }`  Il ruolo deve avere la relazione di trust seguente:  JSON  `{ "Version":"2012-10-17", "Statement": [{ "Effect": "Allow", "Principal": { "Service": "lambda.amazonaws.com" },`

Prerequisito	Description
	"Action": "sts:AssumeRole" }] } Per ulteriori informazioni, consultare Creazione di ruoli IAM nella Guida per l'utente di IAM.

Creazione della funzione Lambda

Procedi come descritto in the section called “Creazione il pacchetto di implementazione Lambda”, ma crea una directory denominata `ddb-to-opensearch` e utilizza il codice seguente per `sample.py`:

```
import boto3
import requests
from requests_aws4auth import AWS4Auth

region = '' # e.g. us-east-1
service = 'es'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

host = '' # the OpenSearch Service domain, e.g. https://search-mydomain.us-
west-1.es.amazonaws.com
index = 'lambda-index'
datatype = '_doc'
url = host + '/' + index + '/' + datatype + '/'

headers = { "Content-Type": "application/json" }

def handler(event, context):
    count = 0
    for record in event['Records']:
        # Get the primary key for use as the OpenSearch ID
        id = record['dynamodb']['Keys']['id']['S']

        if record['eventName'] == 'REMOVE':
            r = requests.delete(url + id, auth=awsauth)
```

```
        else:
            document = record['dynamodb']['NewImage']
            r = requests.put(url + id, auth=awsauth, json=document, headers=headers)
        count += 1
    return str(count) + ' records processed.'
```

Modifica le variabili per `region` e `host`.

Installare pip, se non è già stato fatto, quindi utilizzare i seguenti comandi per installare le dipendenze:

```
cd ddb-to-opensearch

pip install --target ./package requests
pip install --target ./package requests_aws4auth
```

Procedi quindi come descritto in the section called “Creazione della funzione Lambda”, ma specifica il ruolo IAM dai the section called “Prerequisiti” e le impostazioni seguenti per il trigger:

- Tabella: la tabella DynamoDB
- Batch size (Dimensione batch): 100
- Starting position (Posizione di inizio): orizzonte di taglio

Per ulteriori informazioni, consultare Elaborazione di nuovi elementi con DynamoDB Streams e Lambdanella Guida per gli sviluppatori di Amazon DynamoDB.

A questo punto, hai a disposizione un set completo di risorse: una tabella DynamoDB per i dati di origine, un flusso DynamoDB di modifiche alla tabella, una funzione che viene eseguita dopo le modifiche dei dati di origine e indicizza tali modifiche e un dominio di servizio per la ricerca e la visualizzazione. OpenSearch

Test della funzione Lambda

Una volta creata la funzione, è possibile eseguirne il test aggiungendo un nuovo elemento alla tabella DynamoDB utilizzando la AWS CLI:

```
aws dynamodb put-item --table-name test --item '{"director": {"S": "Kevin
 Costner"},"id": {"S": "00001"},"title": {"S": "The Postman"}}' --region us-west-1
```

Utilizzate quindi la console di OpenSearch servizio o le OpenSearch dashboard per verificare che contenga un documento. `lambda-index` Puoi inoltre utilizzare la seguente richiesta:

```
GET https://domain-name/lambda-index/_doc/00001
{
    "_index": "lambda-index",
    "_type": "_doc",
    "_id": "00001",
    "_version": 1,
    "found": true,
    "_source": {
        "director": {
            "S": "Kevin Costner"
        },
        "id": {
            "S": "00001"
        },
        "title": {
            "S": "The Postman"
        }
    }
}
```

Caricamento di dati di streaming da Amazon Data Firehose

Firehose supporta il OpenSearch Servizio come destinazione di consegna. Per istruzioni su come caricare i dati di streaming in OpenSearch Service, consulta Creating a Kinesis Data Firehose Delivery Stream OpenSearch e Choose Service for Your Destination nella Amazon Data Firehose Developer Guide.

Prima di caricare i dati in OpenSearch Service, potrebbe essere necessario eseguire delle trasformazioni sui dati. Per ulteriori informazioni su come usare le funzioni Lambda per completare questa attività, consultare Trasformazione dei dati di Amazon Kinesis Data Firehose nella stessa guida.

Durante la configurazione di un flusso di distribuzione, Firehose offre un ruolo IAM «one-click» che gli fornisce l'accesso alle risorse di cui ha bisogno per inviare dati a OpenSearch Service, eseguire il backup dei dati su Amazon S3 e trasformare i dati utilizzando Lambda. Poiché creare un ruolo simile manualmente sarebbe molto complesso, è consigliabile utilizzare il ruolo fornito.

Caricamento di dati di streaming da Amazon CloudWatch

Puoi caricare dati in streaming da CloudWatch Logs al tuo dominio OpenSearch di servizio utilizzando un abbonamento CloudWatch Logs. Per informazioni sugli CloudWatch abbonamenti Amazon, consulta Elaborazione in tempo reale dei dati di registro con gli abbonamenti. Per informazioni sulla configurazione, consulta Streaming CloudWatch Logs data to Amazon OpenSearch Service nell'Amazon CloudWatch Developer Guide.

Caricamento di dati di streaming da AWS IoT

È possibile inviare dati AWS IoT utilizzando regole. Per ulteriori informazioni, consulta l'OpenSearchazione nella Guida per gli AWS IoT sviluppatori.

Caricamento di dati in Amazon OpenSearch Service con Logstash

La versione open source di Logstash (Logstash OSS) offre un modo conveniente per utilizzare l'API di massa per caricare dati nel tuo dominio Amazon Service. OpenSearch Il servizio supporta tutti i plug-in di input Logstash standard, incluso il plug-in di input Amazon S3. OpenSearch Il servizio supporta il plug-in logstash-output-opensearchdi output, che supporta sia l'autenticazione di base che le credenziali IAM. Il plug-in funziona con la versione 8.1 e precedenti di Logstash OSS.

Configurazione

La configurazione di Logstash varia in base al tipo di autenticazione utilizzata dal dominio.

Indipendentemente dal metodo di autenticazione utilizzato, è necessario impostare `ecs_compatibility` a `disabled` nella sezione di output del file di configurazione. Logstash 8.0 ha introdotto un cambiamento rivoluzionario in cui vengono eseguiti tutti i plug-in in Modalità di compatibilità ECS per impostazione predefinita. È necessario sostituire il valore di default per mantenere il comportamento legacy.

Configurazione dettagliata dei controlli degli accessi

Se il dominio OpenSearch del servizio utilizza un controllo granulare degli accessi con autenticazione di base HTTP, la configurazione è simile a quella di qualsiasi altro cluster. OpenSearch Questo file di configurazione di esempio prende il suo input dalla versione open source di Filebeat (Filebeat OSS).

```
input {
```

```
  beats  {
    port => 5044
  }
}

output {
  opensearch {
    hosts       => "https://domain-endpoint:443"
    user        => "my-username"
    password    => "my-password"
    index       => "logstash-logs-%{+YYYY.MM.dd}"
    ecs_compatibility => disabled
    ssl_certificate_verification => false
  }
}
```

La configurazione varia in base all'applicazione Beats e al caso d'uso, ma la configurazione di Filebeat OSS potrebbe essere simile a questa:

```
filebeat.inputs:
- type: log
  enabled: true
  paths:
    - /path/to/logs/dir/*.log
filebeat.config.modules:
  path: ${path.config}/modules.d/*.yml
  reload.enabled: false
setup.ilm.enabled: false
setup.ilm.check_exists: false
setup.template.settings:
  index.number_of_shards: 1
output.logstash:
  hosts: ["logstash-host:5044"]
```

Configurazione IAM

Se il tuo dominio utilizza una policy di accesso al dominio basata su IAM o un controllo granulare degli accessi con un utente principale, devi firmare tutte le richieste al Servizio utilizzando le credenziali IAM. OpenSearch La seguente politica basata sull'identità concede tutte le richieste HTTP alle sottorisorse del tuo dominio.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/domain-name/*"
    }
  ]
}
```

Per configurare la configurazione di Logstash, modifica il file di configurazione in modo da utilizzare il plugin come output. Questo file di configurazione di esempio prende il suo input dai file in un bucket S3:

```
input {
  s3 {
    bucket => "amzn-s3-demo-"
    region => "us-east-1"
  }
}

output {
  opensearch {
    hosts => ["domain-endpoint:443"]
    auth_type => {
      type => 'aws_iam'
      aws_access_key_id => 'your-access-key'
      aws_secret_access_key => 'your-secret-key'
      region => 'us-east-1'
      }
      index  => "logstash-logs-%{+YYYY.MM.dd}"
      ecs_compatibility => disabled
  }
}
```

Se non si desidera fornire le credenziali IAM all'interno del file di configurazione, è possibile esportarle (o eseguirle `aws configure`):

```
export AWS_ACCESS_KEY_ID="your-access-key"
export AWS_SECRET_ACCESS_KEY="your-secret-key"
export AWS_SESSION_TOKEN="your-session-token"
```

Se il dominio del OpenSearch servizio si trova in un VPC, il computer Logstash OSS deve essere in grado di connettersi al VPC e avere accesso al dominio tramite i gruppi di sicurezza VPC. Per ulteriori informazioni, consulta the section called “Informazioni sulle policy d'accesso nei domini VPC”.

Ricerca di dati in Amazon OpenSearch Service

Esistono diversi metodi comuni per la ricerca di documenti in Amazon OpenSearch Service, tra cui ricerche URI e ricerche nel corpo della richiesta. OpenSearch Il servizio offre funzionalità aggiuntive che migliorano l'esperienza di ricerca, come pacchetti personalizzati, supporto SQL e ricerca asincrona. Per un riferimento completo all'API OpenSearch di ricerca, consulta la documentazione. OpenSearch

> **Note**
>
> Le seguenti richieste di esempio funzionano con OpenSearch APIs. Alcune richieste potrebbero non funzionare con versioni precedenti di Elasticsearch.

Argomenti

- Ricerche negli URI
- Ricerche nel corpo della richiesta
- Paginazione dei risultati della ricerca
- Linguaggio query dashboard
- Importazione e gestione di pacchetti in Amazon Service OpenSearch
- Interrogazione dei dati OpenSearch di Amazon Service con SQL
- Ricerca tra cluster in Amazon Service OpenSearch
- Imparare a classificarsi per Amazon OpenSearch Service
- Ricerca asincrona in Amazon Service OpenSearch
- Ricerca puntuale in Amazon OpenSearch Service
- Ricerca agente in Amazon Service OpenSearch
- Ricerca semantica in Amazon Service OpenSearch
- Ricerca simultanea di segmenti in Amazon Service OpenSearch
- Generazione di query in linguaggio naturale in Amazon OpenSearch Service

Ricerche negli URI

Le ricerche URI (Universal Resource Identifier) sono la forma di ricerca più semplice. In una ricerca URI, si specifica la query come un parametro di richiesta HTTP:

```
GET https://search-my-domain.us-west-1.es.amazonaws.com/_search?q=house
```

L'aspetto di una risposta di esempio è simile al seguente:

```
{
  "took": 25,
  "timed_out": false,
  "_shards": {
    "total": 10,
    "successful": 10,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": {
      "value": 85,
      "relation": "eq",
    },
    "max_score": 6.6137657,
    "hits": [
      {
        "_index": "movies",
        "_type": "movie",
        "_id": "tt0077975",
        "_score": 6.6137657,
        "_source": {
          "directors": [
            "John Landis"
          ],
          "release_date": "1978-07-27T00:00:00Z",
          "rating": 7.5,
          "genres": [
            "Comedy",
            "Romance"
          ],
          "image_url": "http://ia.media-imdb.com/images/M/
MV5BMTY2OTQxNTc1OF5BMl5BanBnXkFtZTYwNjA3NjI5._V1_SX400_.jpg",
```

```
          "plot": "At a 1962 College, Dean Vernon Wormer is determined to expel the
 entire Delta Tau Chi Fraternity, but those troublemakers have other plans for him.",
          "title": "Animal House",
          "rank": 527,
          "running_time_secs": 6540,
          "actors": [
            "John Belushi",
            "Karen Allen",
            "Tom Hulce"
          ],
          "year": 1978,
          "id": "tt0077975"
        }
      },
      ...
    ]
  }
}
```

Per impostazione predefinita, questa query ricerca il termine house in tutti i campi di tutti gli indici. Per restringere la ricerca, specifica un indice (`movies`) e un campo di documento (`title`) nell'URI:

```
GET https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search?q=title:house
```

È possibile includere parametri aggiuntivi nella richiesta, ma i parametri supportati forniscono solo un piccolo sottoinsieme delle opzioni di OpenSearch ricerca. La seguente richiesta restituisce 20 risultati (anziché il numero predefinito di 10) ordinati per anno (anziché per `_score`):

```
GET https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search?
q=title:house&size=20&sort=year:desc
```

Ricerche nel corpo della richiesta

Per eseguire ricerche più complesse, utilizza il corpo della richiesta HTTP e il linguaggio specifico di dominio (DSL) di OpenSearch per query. Il linguaggio DSL di query ti consente di specificare l'intera gamma di valori delle opzioni di ricerca OpenSearch .

> **Note**
>
> Non è possibile includere caratteri speciali Unicode nel valore di un campo di testo, altrimenti il valore verrà analizzato come valori multipli separati dal carattere speciale. Questa analisi errata può portare a un filtraggio involontario dei documenti e persino compromettere il controllo sul loro accesso. Per ulteriori informazioni, vedere Una nota sui caratteri speciali Unicode nei campi di testo nella OpenSearch documentazione.

La seguente query `match` è simile all'esempio di ricerca URI finale:

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search
{
  "size": 20,
  "sort": {
    "year": {
      "order": "desc"
    }
  },
  "query": {
    "query_string": {
      "default_field": "title",
      "query": "house"
    }
  }
}
```

> **Note**
>
> L'API `_search` accetta `GET` e `POST` HTTP per ricerche del corpo della richiesta, ma non tutti i client HTTP supportano l'aggiunta di un corpo della richiesta a una richiesta `GET`. `POST` è la scelta più universale.

In molti casi, potrebbe essere necessario eseguire la ricerca in diversi campi, ma non in tutti. Utilizza la query `multi_match`:

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search
{
  "size": 20,
```

```
  "query": {
    "multi_match": {
      "query": "house",
      "fields": ["title", "plot", "actors", "directors"]
    }
  }
}
```

Campi di boosting

Puoi migliorare la rilevanza della ricerca "potenziando" alcuni campi. I boost sono moltiplicatori che valutano le corrispondenze in un campo più pesantemente rispetto alle corrispondenze in altri campi. Nel seguente esempio, una corrispondenza per john nel campo `title` influenza `_score` il doppio di una corrispondenza nel campo `plot` e quattro volte di più di una corrispondenza nei campi `actors` o `directors`. Ne risulta che film come John Wick e John Carter sono vicini alla parte alta dei risultati della ricerca e film con protagonista John Travolta sono nella parte bassa.

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search
{
  "size": 20,
  "query": {
    "multi_match": {
      "query": "john",
      "fields": ["title^4", "plot^2", "actors", "directors"]
    }
  }
}
```

Evidenziazione dei risultati della ricerca

L'`highlight`opzione indica OpenSearch di restituire un oggetto aggiuntivo all'interno dell'`hits`array se la query corrisponde a uno o più campi:

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search
{
  "size": 20,
  "query": {
    "multi_match": {
      "query": "house",
      "fields": ["title^4", "plot^2", "actors", "directors"]
    }
```

```
  },
  "highlight": {
    "fields": {
      "plot": {}
    }
  }
}
```

Se la query corrisponde al contenuto del campo `plot`, l'aspetto di un'occorrenza è simile al seguente:

```
{
  "_index": "movies",
  "_type": "movie",
  "_id": "tt0091541",
  "_score": 11.276199,
  "_source": {
    "directors": [
      "Richard Benjamin"
    ],
    "release_date": "1986-03-26T00:00:00Z",
    "rating": 6,
    "genres": [
      "Comedy",
      "Music"
    ],
    "image_url": "http://ia.media-imdb.com/images/M/
MV5BMTIzODEzODE2OF5BMl5BanBnXkFtZTcwNjQ3ODcyMQ@@._V1_SX400_.jpg",
    "plot": "A young couple struggles to repair a hopelessly dilapidated house.",
    "title": "The Money Pit",
    "rank": 4095,
    "running_time_secs": 5460,
    "actors": [
      "Tom Hanks",
      "Shelley Long",
      "Alexander Godunov"
    ],
    "year": 1986,
    "id": "tt0091541"
  },
  "highlight": {
    "plot": [
      "A young couple struggles to repair a hopelessly dilapidated <em>house</em>."
```

```
    ]
  }
}
```

Per impostazione predefinita, OpenSearch racchiude la stringa corrispondente nei tag, fornisce fino a 100 caratteri di contesto attorno alla corrispondenza e suddivide il contenuto in frasi identificando segni di punteggiatura, spazi, tabulazioni e interruzioni di riga. Tutte queste impostazioni sono personalizzabili:

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search
{
  "size": 20,
  "query": {
    "multi_match": {
      "query": "house",
      "fields": ["title^4", "plot^2", "actors", "directors"]
    }
  },
  "highlight": {
    "fields": {
      "plot": {}
    },
    "pre_tags": "<strong>",
    "post_tags": "</strong>",
    "fragment_size": 200,
    "boundary_chars": ".,!? "
  }
}
```

API conteggio

Se non sei interessato ai contenuti dei documenti e desideri solo conoscere il numero di corrispondenze, puoi utilizzare l'API `_count` anziché l'API `_search`. La seguente richiesta utilizza la query `query_string` per identificare commedie romantiche:

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/movies/_count
{
  "query": {
    "query_string": {
      "default_field": "genres",
      "query": "romance AND comedy"
```

```
    }
  }
}
```

L'aspetto di una risposta di esempio è simile al seguente:

```
{
  "count": 564,
  "_shards": {
    "total": 5,
    "successful": 5,
    "skipped": 0,
    "failed": 0
  }
}
```

Paginazione dei risultati della ricerca

Se è necessario visualizzare un gran numero di risultati di ricerca, è possibile implementare l'impaginazione utilizzando diversi metodi.

Point in time

La funzionalità point in time (PIT) è un tipo di ricerca che consente di eseguire diverse query su un set di dati fisso nel tempo. Questo è il metodo di impaginazione preferito in OpenSearch, specialmente per l'impaginazione profonda. È possibile utilizzare PIT con OpenSearch Service versione 2.5 e successive. Per ulteriori informazioni su PIT, vedere???.

I `size` parametri `from` e

Il modo più semplice per impaginare è con i `size` parametri `from` and. La richiesta seguente restituisce i risultati 20-39 dell'elenco indicizzato su zero dei risultati della ricerca:

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search
{
  "from": 20,
  "size": 20,
  "query": {
    "multi_match": {
      "query": "house",
```

```
      "fields": ["title^4", "plot^2", "actors", "directors"]
    }
  }
}
```

Per ulteriori informazioni sull'impaginazione della ricerca, consulta Impaginare i risultati nella documentazione. OpenSearch

Linguaggio query dashboard

Puoi utilizzare il Dashboards Query Language (DQL) per cercare dati e visualizzazioni nelle dashboard. OpenSearch DQL utilizza quattro tipi di query principali: termini, booleano, data e intervallo e campo nidificata.

Query di termine

Una query di termini richiede di specificare il termine che si sta cercando.

Per eseguire una query di termini, inserisci quanto segue:

```
host:www.example.com
```

Query booleano

Puoi utilizzare gli operatori booleani `AND`, `or` e `not` per combinare più query.

Per eseguire una query booleana, incollare quanto segue:

```
host.keyword:www.example.com and response.keyword:200
```

Query di data e intervallo

È possibile utilizzare una query di data e intervallo per trovare una data prima o dopo la query.

- > indica una ricerca di una data successiva alla data specificata.
- < indica una ricerca di una data precedente alla data specificata.

```
@timestamp > "2020-12-14T09:35:33"
```

Query di campi nidificati

Se hai un documento con campi nidificati, devi specificare quali parti del documento desideri recuperare. Di seguito è riportato un documento di esempio che contiene campi nidificati:

```
{"NBA players":[
    {"player-name": "Lebron James",
      "player-position": "Power forward",
      "points-per-game": "30.3"
    },
    {"player-name": "Kevin Durant",
      "player-position": "Power forward",
      "points-per-game": "27.1"
    },
    {"player-name": "Anthony Davis",
      "player-position": "Power forward",
      "points-per-game": "23.2"
    },
    {"player-name": "Giannis Antetokounmpo",
      "player-position": "Power forward",
      "points-per-game":"29.9"
    }
  ]
}
```

Per recuperare un campo specifico utilizzando DQL, incollare quanto segue:

```
NBA players: {player-name: Lebron James}
```

Per recuperare più oggetti dal documento nidificato, incollare quanto segue:

```
NBA players: {player-name: Lebron James} and NBA players: {player-name: Giannis
 Antetokounmpo}
```

Per eseguire una ricerca all'interno di un intervallo, incolla quanto segue:

```
NBA players: {player-name: Lebron James} and NBA players: {player-name: Giannis
 Antetokounmpo and < 30}
```

Se il documento ha un oggetto nidificato all'interno di un altro oggetto, è comunque possibile recuperare i dati specificando tutti i livelli. A tale scopo, copia quanto segue:

```
Top-Power-forwards.NBA players: {player-name:Lebron James}
```

Importazione e gestione di pacchetti in Amazon Service OpenSearch

Amazon OpenSearch Service ti consente di caricare file di dizionario personalizzati, come parole chiave e sinonimi, e associare plugin al tuo dominio. Questi plugin possono essere preconfezionati, personalizzati o di terze parti, il che ti offre la flessibilità necessaria per estendere le funzionalità del tuo dominio. Il termine generico per tutti questi tipi di file è pacchetti.

- I file di dizionario aiutano a perfezionare i risultati di ricerca indicando di OpenSearch ignorare le parole più comuni ad alta frequenza o di trattare termini simili, come «crema pasticcera congelata», «gelato» e «gelato», come equivalenti. Possono anche migliorare lo stemming, come si è visto con il plugin di analisi giapponese (kuromoji).
- I plug-in preconfezionati offrono funzionalità integrate, come il plug-in Amazon Personalize per risultati di ricerca personalizzati. Questi plugin utilizzano il tipo di pacchetto. `ZIP-PLUGIN` Per ulteriori informazioni, consulta the section called “Plug-in per versione motore”.
- I plug-in personalizzati e di terze parti consentono di aggiungere funzionalità personalizzate o di integrarsi con sistemi esterni, il che offre una flessibilità ancora maggiore per il dominio. Come i plugin preconfezionati, caricate i plugin personalizzati come pacchetti. `ZIP-PLUGIN` Per i plug-in di terze parti, è inoltre necessario importare la licenza del plug-in e i file di configurazione come pacchetti separati, quindi associarli tutti al dominio.

  Per ulteriori informazioni, consulta i seguenti argomenti:

  - the section called “Plugin personalizzati”
  - the section called “Plugin di terze parti”

> **Note**
>
> Puoi associare un massimo di 20 plugin a un singolo dominio. Questo limite include tutti i tipi di plug-in: plug-in opzionali, di terze parti e personalizzati.

Argomenti

- Autorizzazioni richieste

- Caricamento di pacchetti in Amazon S3
- Importazione e associazione di pacchetti
- Utilizzo di pacchetti con OpenSearch
- Aggiornamento dei pacchetti
- Aggiornamento manuale degli indici con un nuovo dizionario
- Dissociazione e rimozione dei pacchetti
- Gestione dei plugin personalizzati in Amazon Service OpenSearch
- Installazione di plug-in di terze parti in Amazon Service OpenSearch

Autorizzazioni richieste

Gli utenti senza accesso come amministratore richiedono determinate azioni AWS Identity and Access Management (IAM) per gestire i pacchetti:

- `es:CreatePackage`— Creare un pacchetto
- `es:DeletePackage`— Eliminare un pacchetto
- `es:AssociatePackage`— Associare un pacchetto a un dominio
- `es:DissociatePackage`— Dissociare un pacchetto da un dominio

Sono necessarie anche le autorizzazioni per il percorso del bucket Amazon S3 o l'oggetto in cui si trova il pacchetto personalizzato.

Concedere tutte le autorizzazioni all'interno di IAM, non nella policy di accesso al dominio. Per ulteriori informazioni, consultare the section called “Identity and Access Management”.

Caricamento di pacchetti in Amazon S3

Questa sezione spiega come caricare pacchetti di dizionari personalizzati, poiché i pacchetti di plug-in preconfezionati sono già installati. Prima di poter associare un dizionario personalizzato al tuo dominio, devi caricarlo in un bucket Amazon S3. Per le istruzioni, consulta Caricamento di oggetti nella Guida per l'utente di Amazon Simple Storage Service. I plugin supportati non devono essere caricati.

Se il dizionario contiene informazioni sensibili, specifica la crittografia lato server con chiavi gestite da S3 al momento del caricamento. OpenSearch Il servizio non può accedere ai file in S3 che proteggi utilizzando una chiave. AWS KMS

Dopo aver caricato il file, prendere nota del suo percorso S3. Il formato del percorso è `s3://amzn-s3-demo-bucket/file-path/file-name`.

È possibile utilizzare il seguente file di sinonimi per scopi di test. Salvalo come `synonyms.txt`.

```
danish, croissant, pastry
ice cream, gelato, frozen custard
sneaker, tennis shoe, running shoe
basketball shoe, hightop
```

Alcuni dizionari, come i dizionari Hunspell, utilizzano più file e richiedono le proprie directory nel file system. Al momento, OpenSearch Service supporta solo dizionari a file singolo.

Importazione e associazione di pacchetti

La console è il modo più semplice per importare un dizionario personalizzato in Service. OpenSearch Quando importi un dizionario da Amazon S3, OpenSearch Service archivia la propria copia del pacchetto e la crittografa automaticamente utilizzando AES-256 con chiavi gestite dal servizio. OpenSearch

I plug-in opzionali sono già preinstallati in OpenSearch Service, quindi non è necessario caricarli personalmente, ma è necessario associare un plug-in a un dominio. I plugin disponibili sono elencati nella schermata Pacchetti della console.

Importa e associa un pacchetto a un dominio

1. Nella console di Amazon OpenSearch Service, scegli Pacchetti.
2. Scegli Importa pacchetto.
3. Assegnare al pacchetto un nome descrittivo.
4. Fornire il percorso S3 del file e quindi scegliere Importa.
5. Tornare alla schermata Pacchetti .
6. Quando lo stato del pacchetto è disponibile, selezionarlo.
7. Scegli Associa a un dominio.
8. Seleziona un dominio, quindi scegli Avanti. Esamina i pacchetti e scegli Associa.
9. Nel pannello di navigazione, scegliere il dominio e passare alla scheda Pacchetti.
10. Se il pacchetto è un dizionario personalizzato, annota l'ID quando il pacchetto diventa Disponibile. Utilizza `analyzers/id` come percorso del file nelle richieste di OpenSearch.

Utilizzo di pacchetti con OpenSearch

Questa sezione spiega come utilizzare entrambi i tipi di pacchetti: dizionari personalizzati e plugin preconfezionati.

Usare dizionari personalizzati

Dopo aver associato un file a un dominio, è possibile utilizzarlo in parametri quali `synonyms_path` e `user_dictionary` quando si creano tokenizzatori e filtri token. `stopwords_path` Il parametro esatto varia in base all'oggetto. Diversi oggetti supportano `synonyms_path` e `stopwords_path`, ma `user_dictionary` è esclusivo per il plugin kuromoji.

Per il plug-in IK (Chinese) Analysis, è possibile caricare un file di dizionario personalizzato come pacchetto personalizzato e associarlo a un dominio e il plug-in lo raccoglie automaticamente senza richiedere un parametro `user_dictionary`. Se il file è un file di sinonimi, usa il parametro `synonyms_path`.

Nel seguente esempio viene aggiunto un file di sinonimi a un nuovo indice:

```
PUT my-index
{
  "settings": {
    "index": {
      "analysis": {
        "analyzer": {
          "my_analyzer": {
            "type": "custom",
            "tokenizer": "standard",
            "filter": ["my_filter"]
          }
        },
        "filter": {
          "my_filter": {
            "type": "synonym",
            "synonyms_path": "analyzers/F111111111",
            "updateable": true
          }
        }
      }
    }
  },
```

```
  "mappings": {
    "properties": {
      "description": {
        "type": "text",
        "analyzer": "standard",
        "search_analyzer": "my_analyzer"
      }
    }
  }
}
```

Questa richiesta crea un analizzatore personalizzato per l'indice che utilizza il tokenizer standard e un filtro token sinonimo.

- I tokenizer scompongono i flussi di caratteri in token (in genere parole) in base ad alcune regole. L'esempio più semplice è il tokenizer di whitespace, che scompone i caratteri precedenti in un token ogni volta che incontra un carattere di whitespace. Un esempio più complesso è il tokenizer standard, che utilizza un insieme di regole grammaticali per lavorare in molte lingue.
- I filtri token aggiungono, modificano o eliminano token. Ad esempio, un filtro token sinonimo aggiunge token quando trova una parola nell'elenco dei sinonimi. Il filtro token di arresto rimuove i token quando trova una parola nell'elenco di parole stop.

Questa richiesta aggiunge anche un campo di testo (`description`) alla mappatura e indica di OpenSearch utilizzare il nuovo analizzatore come analizzatore di ricerca. Si può vedere che come analizzatore di indice è utilizzato ancora l'analizzatore standard.

Infine, prendere nota della riga `"updateable": true` nel filtro token. Questo campo si applica solo agli analizzatori di ricerca non agli analizzatori di indice, ed è critico se in seguito si desidera aggiornare l'analizzatore di ricerca automaticamente.

A scopo di test, aggiungere alcuni documenti all'indice:

```
POST _bulk
{ "index": { "_index": "my-index", "_id": "1" } }
{ "description": "ice cream" }
{ "index": { "_index": "my-index", "_id": "2" } }
{ "description": "croissant" }
{ "index": { "_index": "my-index", "_id": "3" } }
{ "description": "tennis shoe" }
{ "index": { "_index": "my-index", "_id": "4" } }
```

```
{ "description": "hightop" }
```

Quindi cercarli usando un sinonimo:

```
GET my-index/_search
{
  "query": {
    "match": {
      "description": "gelato"
    }
  }
}
```

In questo caso, OpenSearch restituisce la seguente risposta:

```
{
  "hits": {
    "total": {
      "value": 1,
      "relation": "eq"
    },
    "max_score": 0.99463606,
    "hits": [{
      "_index": "my-index",
      "_type": "_doc",
      "_id": "1",
      "_score": 0.99463606,
      "_source": {
        "description": "ice cream"
      }
    }]
  }
}
```

> **Tip**
>
> I file di dizionario utilizzano lo spazio heap Java proporzionale alle loro dimensioni. Ad esempio, un file di dizionario a 2 GiB potrebbe consumare su un nodo 2 GiB di spazio heap. Se si utilizzano file di grandi dimensioni, assicurarsi che i nodi dispongano di spazio heap sufficiente per contenerli. Monitorare il parametro `JVMMemoryPressure` e dimensionare il cluster in base alle esigenze.

Utilizzo di plugin preconfezionati

OpenSearch Il servizio consente di associare OpenSearch plug-in opzionali preinstallati da utilizzare al dominio. Un pacchetto di plug-in preconfezionato è compatibile con una OpenSearch versione specifica e può essere associato solo a domini con quella versione. L'elenco dei pacchetti disponibili per il tuo dominio include tutti i plugin supportati compatibili con la versione del tuo dominio. Dopo aver associato un plug-in a un dominio, inizia un processo di installazione sul dominio. Quindi, puoi fare riferimento e utilizzare il plug-in quando effettui richieste al OpenSearch Servizio.

L'associazione e la dissociazione di un plug-in richiedono una distribuzione. blue/green Per ulteriori informazioni, consulta the section called “Modifiche che di solito causano implementazioni blue/green”.

I plugin opzionali includono analizzatori linguistici e risultati di ricerca personalizzati. Ad esempio, il plug-in Amazon Personalize Search Ranking utilizza l'apprendimento automatico per personalizzare i risultati di ricerca per i tuoi clienti. Per ulteriori informazioni su questo plugin, consulta Personalizzazione dei risultati di ricerca da. OpenSearch Per un elenco di tutti i plugin supportati, consulta. the section called “Plug-in per versione motore”

Plugin Sudachi

Per il plugin Sudachi, quando riassoci un file di dizionario, questo non si riflette immediatamente sul dominio. Il dizionario si aggiorna quando viene eseguita la blue/green distribuzione successiva sul dominio come parte di una modifica della configurazione o di un altro aggiornamento. In alternativa, è possibile creare un nuovo pacchetto con i dati aggiornati, creare un nuovo indice utilizzando questo nuovo pacchetto, reindicizzare l'indice esistente nel nuovo indice e quindi eliminare il vecchio indice. Se preferisci utilizzare l'approccio di reindicizzazione, utilizza un alias di indice in modo da evitare interruzioni del traffico.

Inoltre, il plugin Sudachi supporta solo dizionari binari Sudachi, che puoi caricare con l'operazione API. CreatePackage Per informazioni sul dizionario di sistema predefinito e sul processo di compilazione dei dizionari utente, consulta la documentazione Sudachi.

> **Note**
>
> Quando carichi file di dizionario binario su Amazon S3, devi impostare il Content-Type dell'oggetto S3 su. `binary/octet-stream` L'utilizzo `application/octet-stream` farà fallire l'importazione del pacchetto.

L'esempio seguente mostra come utilizzare i dizionari di sistema e utente con il tokenizer Sudachi. È necessario caricare questi dizionari come pacchetti personalizzati con tipo `TXT-DICTIONARY` e fornire il relativo pacchetto nelle impostazioni aggiuntive. IDs

```
PUT sudachi_sample
{
  "settings": {
    "index": {
      "analysis": {
        "tokenizer": {
          "sudachi_tokenizer": {
            "type": "sudachi_tokenizer",
            "additional_settings": "{\"systemDict\": \"<system-dictionary-package-id>\",\"userDict\": [\"<user-dictionary-package-id>\"]}"
        }
        },
        "analyzer": {
          "sudachi_analyzer": {
            "filter": ["my_searchfilter"],
            "tokenizer": "sudachi_tokenizer",
            "type": "custom"
          }
        },
        "filter":{
          "my_searchfilter": {
            "type": "sudachi_split",
            "mode": "search"
          }
        }
      }
    }
  }
}
```

Aggiornamento dei pacchetti

Questa sezione spiega solo come aggiornare un pacchetto di dizionario personalizzato, poiché i pacchetti di plugin preconfezionati sono già aggiornati automaticamente. Il caricamento di una nuova versione di un dizionario su Amazon S3 non aggiorna automaticamente il pacchetto su Amazon OpenSearch Service. OpenSearch Il servizio archivia la propria copia del file, quindi se carichi una nuova versione su S3, devi aggiornarla manualmente.

Ciascuno dei domini associati archivia anche la propria copia del file. Per mantenere prevedibile il comportamento di ricerca, i domini continueranno a utilizzare la versione corrente del pacchetto fino a quando non vengono aggiornati esplicitamente. Per aggiornare un pacchetto personalizzato, modifica il file in Amazon S3 Control, aggiorna il pacchetto in OpenSearch Service, quindi applica l'aggiornamento.

Console

1. Nella console OpenSearch di servizio, scegli Pacchetti.
2. Scegli un pacchetto, quindi seleziona Aggiorna.
3. Fornisci un nuovo percorso S3 al file, quindi scegli Aggiorna pacchetto.
4. Tornare alla schermata Pacchetti .
5. Quando lo stato del pacchetto è Disponibile, selezionarlo. Scegliere uno o più domini associati, selezionare Applica aggiornamento e confermare. Attendere che lo stato dell'associazione cambi in Attivo.
6. I passaggi successivi variano a seconda di come hai configurato gli indici:

   - Se il tuo dominio è in esecuzione OpenSearch con Elasticsearch 7.8 o versione successiva e utilizza solo analizzatori di ricerca con il campo aggiornabile impostato su true, non devi intraprendere ulteriori azioni. OpenSearch Il servizio aggiorna automaticamente gli indici utilizzando l'API _plugins/_refresh_search_analyzers.
   - Se il tuo dominio esegue Elasticsearch 7.7 o versioni precedenti, utilizza analizzatori di indici o non utilizza il campo, vedi. `updateable` the section called “Aggiornamento manuale degli indici con un nuovo dizionario”

Sebbene la console sia il metodo più semplice, puoi anche utilizzare l'API SDKs, o di configurazione per AWS CLI aggiornare i pacchetti di servizi. OpenSearch Per ulteriori informazioni, consulta AWS CLI Command Reference e Amazon OpenSearch Service API Reference.

AWS SDK

Invece di aggiornare manualmente un pacchetto nella console, puoi utilizzare il SDKs per automatizzare il processo di aggiornamento. Il seguente script Python di esempio carica un nuovo file di pacchetto su Amazon S3, aggiorna il pacchetto in OpenSearch Service e applica il nuovo pacchetto al dominio specificato. Dopo aver verificato l'avvenuto aggiornamento, esegue una chiamata di esempio per OpenSearch dimostrare che i nuovi sinonimi sono stati applicati.

È necessario fornire valori per `host`, `region`, `file_name`, `bucket_name`, `s3_key`, `package_id`, `domain_name` e `query`.

```
from requests_aws4auth import AWS4Auth
import boto3
import requests
import time
import json
import sys

host = ''  # The OpenSearch domain endpoint with https:// and a trailing slash. For
 example, https://my-test-domain.us-east-1.es.amazonaws.com/
region = ''  # For example, us-east-1
file_name = ''  # The path to the file to upload
bucket_name = ''  # The name of the S3 bucket to upload to
s3_key = ''  # The name of the S3 key (file name) to upload to
package_id = ''  # The unique identifier of the OpenSearch package to update
domain_name = ''  # The domain to associate the package with
query = ''  # A test query to confirm the package has been successfully updated

service = 'es'
credentials = boto3.Session().get_credentials()
client = boto3.client('opensearch')
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key,
                   region, service, session_token=credentials.token)


def upload_to_s3(file_name, bucket_name, s3_key):
    """Uploads file to S3"""
    s3 = boto3.client('s3')
    try:
        s3.upload_file(file_name, bucket_name, s3_key)
        print('Upload successful')
        return True
    except FileNotFoundError:
        sys.exit('File not found. Make sure you specified the correct file path.')


def update_package(package_id, bucket_name, s3_key):
    """Updates the package in OpenSearch Service"""
    print(package_id, bucket_name, s3_key)
    response = client.update_package(
        PackageID=package_id,
```

```
        PackageSource={
            'S3BucketName': bucket_name,
            'S3Key': s3_key
        }
    )
    print(response)


def associate_package(package_id, domain_name):
    """Associates the package to the domain"""
    response = client.associate_package(
        PackageID=package_id, DomainName=domain_name)
    print(response)
    print('Associating...')


def wait_for_update(domain_name, package_id):
    """Waits for the package to be updated"""
    response = client.list_packages_for_domain(DomainName=domain_name)
    package_details = response['DomainPackageDetailsList']
    for package in package_details:
        if package['PackageID'] == package_id:
            status = package['DomainPackageStatus']
            if status == 'ACTIVE':
                print('Association successful.')
                return
            elif status == 'ASSOCIATION_FAILED':
                sys.exit('Association failed. Please try again.')
            else:
                time.sleep(10)  # Wait 10 seconds before rechecking the status
                wait_for_update(domain_name, package_id)


def sample_search(query):
    """Makes a sample search call to OpenSearch"""
    path = '_search'
    params = {'q': query}
    url = host + path
    response = requests.get(url, params=params, auth=awsauth)
    print('Searching for ' + '"' + query + '"')
    print(response.text)
```

> **Note**
>
> Se ricevi un errore «pacchetto non trovato» quando esegui lo script utilizzando il AWS CLI, probabilmente significa che Boto3 sta utilizzando la regione specificata in ~/.aws/config, che non è la regione in cui si trova il bucket S3. Esegui `aws configure` e specifica la regione corretta oppure aggiungi esplicitamente la regione al client:
>
> ```
> client = boto3.client('opensearch', region_name='us-east-1')
> ```

Aggiornamento manuale degli indici con un nuovo dizionario

Gli aggiornamenti manuali degli indici si applicano solo ai dizionari personalizzati, non ai plugin preconfezionati. Per utilizzare un dizionario aggiornato, è necessario aggiornare manualmente gli indici se si soddisfa una delle seguenti condizioni:

- Il dominio esegue Elasticsearch 7.7 o versioni precedenti.
- I pacchetti personalizzati vengono utilizzati come analizzatori di indice.
- I pacchetti personalizzati vengono utilizzati come analizzatori di ricerca, ma non includono il campo aggiornabile.

Per aggiornare gli analizzatori con i file del nuovo pacchetto, sono disponibili due opzioni:

- Chiudi e apri gli indici da aggiornare:

```
POST my-index/_close
POST my-index/_open
```

- Reindicizzare gli indici. Innanzitutto, create un indice che utilizzi il file dei sinonimi aggiornato (o un file completamente nuovo). Nota che è supportato solo UTF-8.

```
PUT my-new-index
{
  "settings": {
    "index": {
      "analysis": {
        "analyzer": {
          "synonym_analyzer": {
```

```
            "type": "custom",
            "tokenizer": "standard",
            "filter": ["synonym_filter"]
          }
        },
        "filter": {
          "synonym_filter": {
            "type": "synonym",
            "synonyms_path": "analyzers/F222222222"
          }
        }
      }
    }
  },
  "mappings": {
    "properties": {
      "description": {
        "type": "text",
        "analyzer": "synonym_analyzer"
      }
    }
  }
}
```

Quindi reindicizzare il vecchio indice nel nuovo indice:

```
POST _reindex
{
  "source": {
    "index": "my-index"
  },
  "dest": {
    "index": "my-new-index"
  }
}
```

Se gli analizzatori di indice vengono aggiornati con una certa frequenza, utilizzare gli alias di indice per mantenere un percorso coerente per l'indice più recente:

```
POST _aliases
{
  "actions": [
```

```
    {
      "remove": {
        "index": "my-index",
        "alias": "latest-index"
      }
    },
    {
      "add": {
        "index": "my-new-index",
        "alias": "latest-index"
      }
    }
  ]
}
```

Se il vecchio indice non è necessario, eliminarlo.

```
DELETE my-index
```

Dissociazione e rimozione dei pacchetti

La dissociazione di un pacchetto, che si tratti di un dizionario personalizzato o di un plug-in preconfezionato, da un dominio significa che non è più possibile utilizzare quel pacchetto quando si creano nuovi indici. Dopo la dissociazione di un pacchetto, gli indici esistenti che lo utilizzavano non possono più utilizzarlo. È necessario rimuovere il pacchetto da qualsiasi indice prima di poterlo dissociare, altrimenti la dissociazione fallisce.

La console è il modo più semplice per dissociare un pacchetto da un dominio e rimuoverlo dal servizio. OpenSearch La rimozione di un pacchetto dal OpenSearch servizio non lo rimuove dalla sua posizione originale su Amazon S3.

Dissocia un pacchetto da un dominio

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione, scegli Domini.
3. Scegli il dominio, quindi vai alla scheda Pacchetti.
4. Scegliere un pacchetto, selezionare Operazioni, quindi scegliere Dissocia. Conferma la tua scelta.

5. Attendi che il pacchetto scompaia dall'elenco. Potrebbe essere necessario aggiornare il browser.
6. Se desideri utilizzare il pacchetto con altri domini, interrompi questa operazione a questo punto. Per continuare a rimuovere il pacchetto (se si tratta di un dizionario personalizzato), scegli Pacchetti nel riquadro di navigazione.
7. Seleziona il pacchetto e scegli Elimina.

In alternativa, usa l'API AWS CLI SDKs, o configuration per dissociare e rimuovere i pacchetti. Per ulteriori informazioni, consulta AWS CLI Command Reference e Amazon OpenSearch Service API Reference.

Gestione dei plugin personalizzati in Amazon Service OpenSearch

Utilizzando i plug-in personalizzati per OpenSearch Service, puoi estendere le OpenSearch funzionalità in aree come l'analisi del linguaggio, il filtraggio personalizzato, il posizionamento e altro ancora, consentendoti di creare esperienze di ricerca personalizzate. OpenSearch È possibile sviluppare plugin personalizzati per estendendo la `org.opensearch.plugins.Plugin` classe e quindi impacchettandola in un file. `.zip`

Le seguenti estensioni dei plugin sono attualmente supportate da Amazon OpenSearch Service:

- AnalysisPlugin— Estende la funzionalità di analisi aggiungendo, ad esempio, analizzatori personalizzati, tokenizzatori di caratteri o filtri per l'elaborazione del testo.
- SearchPlugin— Migliora le funzionalità di ricerca con tipi di query personalizzati, algoritmi di somiglianza, opzioni di suggerimento e aggregazioni.
- MapperPlugin— Consente di creare tipi di campo personalizzati e le relative configurazioni di mappatura OpenSearch, consentendo di definire in che modo i diversi tipi di dati devono essere archiviati e indicizzati.
- ScriptPlugin— Consente di aggiungere funzionalità di script personalizzate, ad esempio OpenSearch, agli script personalizzati per operazioni come l'assegnazione di punteggi, l'ordinamento e le trasformazioni dei valori dei campi durante la ricerca o l'indicizzazione.

Puoi utilizzare la console OpenSearch di servizio o i comandi API esistenti per pacchetti personalizzati per caricare e associare il plug-in al cluster Amazon OpenSearch Service. Puoi anche usare il DescribePackagescomando per descrivere tutti i pacchetti del tuo account e per visualizzare dettagli come la OpenSearch versione e i dettagli degli errori. OpenSearch Il servizio convalida il pacchetto di plug-in per verificare la compatibilità delle versioni, le vulnerabilità di sicurezza e le

operazioni consentite del plug-in. Per ulteriori informazioni sui pacchetti personalizzati, consulta. the section called “Pacchetti”

OpenSearch versione e Regione AWS supporto

I plugin personalizzati sono supportati nei domini OpenSearch di servizio che eseguono la OpenSearch versione 2.15 nei seguenti casi: Regioni AWS

- Stati Uniti orientali (Ohio) (us-east-2)
- Stati Uniti orientali (Virginia settentrionale) (us-east-1)
- Stati Uniti occidentali (Oregon) (us-west-2)
- Asia Pacifico (Mumbai) (ap-south-1)
- Asia Pacifico (Seoul) (ap-northeast-2)
- Asia Pacifico (Singapore) (ap-southeast-1)
- Asia Pacifico (Sydney) (ap-southeast-2)
- Asia Pacifico (Tokyo) (ap-northeast-1)
- Canada (Centrale) (ca-central-1)
- Europa (Francoforte) (eu-central-1)
- Europa (Irlanda) (eu-west-1)
- Europa (Londra) (eu-west-2)
- Europa (Parigi) (eu-west-3)
- Sud America (San Paolo) (sa-east-1)

> **Note**
>
> I plugin personalizzati contengono codice sviluppato dall'utente. Eventuali problemi, comprese le violazioni degli SLA, causati dal codice sviluppato dagli utenti non danno diritto ai crediti SLA. Per ulteriori informazioni, consulta Amazon OpenSearch Service - Service Level Agreement.

Argomenti

- Quote dei plugin
- Prerequisiti

- Risoluzione dei problemi
- Installazione di un plug-in personalizzato tramite la console
- Gestire i plugin personalizzati utilizzando il AWS CLI
- AWS KMS Integrazione di pacchetti personalizzati Amazon OpenSearch Service

Quote dei plugin

- Puoi creare fino a 25 plugin personalizzati per account per regione.
- La dimensione massima non compressa per un plug-in è di 1 GB.
- Il numero massimo di plugin che possono essere associati a un singolo dominio è 20. Questa quota si applica a tutti i tipi di plugin combinati: opzionali, di terze parti e personalizzati.
- I plugin personalizzati sono supportati nei domini che eseguono la OpenSearch versione 2.15 o successiva.
- Il `descriptor.properties` file del plug-in deve supportare una versione del motore simile alla 2.15.0 o qualsiasi versione 2.x.x, in cui la versione della patch è impostata su zero.

Prerequisiti

Prima di installare un plug-in personalizzato e associarlo a un dominio, assicurati di soddisfare i seguenti requisiti:

- La versione del motore supportata per il plugin nel `descriptor.properties` file deve essere simile a `2.15.0` o`2.x.0`. Cioè, la versione della patch deve essere zero.
- Le seguenti funzionalità devono essere abilitate sul tuo dominio:
  - Node-to-node crittografia
  - Crittografia dei dati inattivi
  - `EnforceHTTPS`è impostato su «true»

    Vedi anche opensearch-https-requirednella Guida per gli AWS Config sviluppatori.
  - I client devono supportare Policy-MIN-TLS-1-2-pfs-2023-10. È possibile specificare questo supporto utilizzando il seguente comando. *`placeholder value`*Sostituiscili con le tue informazioni:

    ```
    aws opensearch update-domain-config \
        --domain-name domain-name \
    ```

```
    --domain-endpoint-options '{"TLSSecurityPolicy":"Policy-Min-TLS-1-2-
PFS-2023-10" }'
```

Per ulteriori informazioni, DomainEndpointOptionsconsulta l'Amazon OpenSearch Service API Reference.

Risoluzione dei problemi

Se il sistema restituisce l'errore`PluginValidationFailureReason : The provided plugin could not be loaded`, consulta the section called “L'installazione del plug-in personalizzato non riesce a causa della compatibilità della versione” per informazioni sulla risoluzione dei problemi.

Installazione di un plug-in personalizzato tramite la console

Per associare un plug-in di terze parti a un dominio, importate innanzitutto la licenza e la configurazione del plug-in come pacchetti.

Per installare un plugin personalizzato

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli Pacchetti.
3. Scegli Importa pacchetto.
4. Per Nome, inserisci un nome univoco e facilmente identificabile per il plugin.
5. (Facoltativo) In Descrizione, fornite tutti i dettagli utili sul pacchetto o sul suo scopo.
6. Per Tipo di pacchetto, scegli Plugin.
7. Per Package source, inserisci il percorso o cerca il file ZIP del plugin in Amazon S3.
8. Per la versione OpenSearch del motore, scegli la versione OpenSearch supportata dal plugin.
9. Per Package encryption, scegli se personalizzare la chiave di crittografia per il pacchetto. Per impostazione predefinita, OpenSearch Service crittografa il pacchetto del plug-in con un Chiave di proprietà di AWS. Puoi invece utilizzare una chiave gestita dal cliente.
10. Scegli Importa.

Dopo aver importato il pacchetto del plugin, associalo a un dominio. Per istruzioni, consulta the section called “Importa e associa un pacchetto a un dominio”.

Gestire i plugin personalizzati utilizzando il AWS CLI

È possibile utilizzare il AWS CLI per gestire una serie di attività personalizzate relative ai plug-in.

Processi

- Installazione di un plug-in personalizzato utilizzando il AWS CLI
- Aggiornamento di un plug-in personalizzato utilizzando il AWS CLI
- Crea o aggiorna un plug-in personalizzato con una AWS KMS chiave di sicurezza
- Aggiornamento di un dominio OpenSearch di servizio con plug-in personalizzati a una versione successiva di utilizzo di OpenSearch AWS CLI
- Disinstallazione e visualizzazione dello stato di dissociazione di un plug-in personalizzato

Installazione di un plug-in personalizzato utilizzando il AWS CLI

Prima di iniziare

Prima di poter associare un plug-in personalizzato al tuo dominio, devi caricarlo in un bucket Amazon Simple Storage Service (Amazon S3). Il bucket deve trovarsi nello stesso Regione AWS punto in cui intendi utilizzare il plug-in. Per informazioni sull'aggiunta di un oggetto a un bucket S3, consulta Caricamento di oggetti nella Amazon Simple Storage Service User Guide.

Se il plug-in contiene informazioni sensibili, specifica la crittografia lato server con chiavi gestite da S3 al momento del caricamento. Dopo aver caricato il file, prendere nota del suo percorso S3. Il formato del percorso è `s3://amzn-s3-demo-bucket/file-path/file-name`.

> **Note**
>
> Facoltativamente, puoi proteggere un plug-in personalizzato quando crei il plug-in specificando una chiave (). AWS Key Management Service AWS KMS Per informazioni, consulta the section called “Crea o aggiorna un plug-in personalizzato con una AWS KMS chiave di sicurezza”.

Per installare un plugin personalizzato utilizzando il AWS CLI

1. Crea un nuovo pacchetto per il tuo plugin personalizzato eseguendo il seguente comando create-package, assicurandoti che siano soddisfatti i seguenti requisiti:

- La posizione del bucket e della chiave deve puntare al `.zip` file del plug-in in un bucket S3 dell'account in cui stai eseguendo i comandi.
- Il bucket S3 deve trovarsi nella stessa regione in cui viene creato il pacchetto.
- Per `.zip` `ZIP-PLUGIN` i pacchetti sono supportati solo i file.
- Il contenuto del `.zip` file deve seguire la struttura delle cartelle prevista dal plugin.
- Il valore di `--engine-version` deve essere nel formato`OpenSearch_`*`{MAJOR}`*`.`*`{MINOR}`*. Ad esempio: **`OpenSearch_2.17`**.

*`placeholder values`*Sostituiscili con le tue informazioni:

```
aws opensearch create-package \
    --package-name package-name \
    --region region \
    --package-type ZIP-PLUGIN \
    --package-source S3BucketName=amzn-s3-demo-bucket,S3Key=s3-key \
    --engine-version opensearch-version
```

2. (Facoltativo) Visualizza lo stato dell'`create-package`operazione, inclusi eventuali risultati di convalida e vulnerabilità di sicurezza, utilizzando il comando describe-packages. Sostituiscili con le tue informazioni*`placeholder values`*:

```
aws opensearch describe-packages \
    --region region  \
    --filters '[{"Name": "PackageType","Value": ["ZIP-PLUGIN"]}, {"Name":
 "PackageName","Value": ["package-name"]}]'
```

Questo comando restituisce informazioni simili alle seguenti:

```
{
    "PackageDetailsList": [{
        "PackageID": "pkg-identifier",
        "PackageName": "package-name",
        "PackageType": "ZIP-PLUGIN",
        "PackageStatus": "VALIDATION_FAILED",
        "CreatedAt": "2024-11-11T13:07:18.297000-08:00",
        "LastUpdatedAt": "2024-11-11T13:10:13.843000-08:00",
        "ErrorDetails": {
            "ErrorType": "",
```

```
            "ErrorMessage": "PluginValidationFailureReason : Dependency Scan
 reported 3 vulnerabilities for the plugin: CVE-2022-23307, CVE-2019-17571,
 CVE-2022-23305"
        },
        "EngineVersion": "OpenSearch_2.15",
        "AllowListedUserList": [],
        "PackageOwner": "OWNER-XXXX"
    }]
}
```

> **Note**
>
> Durante l'`create-package`operazione, Amazon OpenSearch Service verifica il `ZIP-PLUGIN` valore per la compatibilità delle versioni, le estensioni dei plugin supportate e le vulnerabilità di sicurezza. Le vulnerabilità di sicurezza vengono analizzate utilizzando il servizio Amazon Inspector. I risultati di questi controlli vengono visualizzati nel `ErrorDetails` campo della risposta dell'API.

3. Utilizzate il comando associate-package per associare il plugin al dominio di OpenSearch servizio di vostra scelta utilizzando l'ID del pacchetto creato nel passaggio precedente.

> **Tip**
>
> Se disponi di più plugin, puoi invece utilizzare il comando associate-packages per associare più pacchetti a un dominio con un'unica operazione.

Sostituiscili *`placeholder values`* con le tue informazioni:

```
aws opensearch associate-package \
    --domain-name domain-name \
    --region region \
    --package-id package-id
```

> **Note**
>
> Il plugin viene installato e disinstallato utilizzando un processo di distribuzione blu/verde.

4. (Facoltativo) Utilizzate il list-packages-for-domaincomando per visualizzare lo stato dell'associazione. Lo stato dell'associazione cambia man mano che il flusso di lavoro avanza da `ASSOCIATING` a`ACTIVE`. Lo stato dell'associazione diventa ATTIVO dopo il completamento dell'installazione del plug-in e il plug-in è pronto per l'uso.

   Sostituisci *`placeholder values`* con le informazioni appropriate.

   ```
   aws opensearch list-packages-for-domain \
       --region region \
       --domain-name domain-name
   ```

Aggiornamento di un plug-in personalizzato utilizzando il AWS CLI

Utilizzate il comando update-package per apportare modifiche a un plugin.

> **Note**
>
> Facoltativamente, puoi proteggere un plug-in personalizzato quando aggiorni il plug-in specificando una AWS Key Management Service chiave ().AWS KMS Per informazioni, consulta the section called “Crea o aggiorna un plug-in personalizzato con una AWS KMS chiave di sicurezza”.

Per aggiornare un plug-in personalizzato utilizzando il AWS CLI

- Eseguire il seguente comando seguente. Sostituisci *`placeholder values`* con le informazioni appropriate.

  ```
  aws opensearch update-package \
      --region region \
      --package-id package-id \
      --package-source S3BucketName=amzn-s3-demo-bucket,S3Key=s3-key \
      --package-description description
  ```

Dopo aver aggiornato un pacchetto, è possibile utilizzare il comando associate-package o associate-packages per applicare gli aggiornamenti del pacchetto a un dominio.

> **Note**
>
> È possibile controllare, creare, aggiornare, associare e dissociare le operazioni sul plug-in utilizzando. AWS CloudTrail Per ulteriori informazioni, consulta Monitoraggio delle chiamate API di Amazon OpenSearch Service con AWS CloudTrail.

Crea o aggiorna un plug-in personalizzato con una AWS KMS chiave di sicurezza

È possibile proteggere un plug-in personalizzato quando si crea o si aggiorna il plug-in specificando una AWS KMS chiave. A tale scopo, imposta `PackageEncryptionOptions` `true` e specifica l'Amazon Resource Name (ARN) della chiave, come mostrato negli esempi seguenti.

Esempio: crea un plug-in personalizzato con AWS KMS chiave di sicurezza

```
aws opensearch create-package \
    --region us-east-2  --package-name my-custom-package \
    --package-type ZIP-PLUGIN \
    --package-source S3BucketName=amzn-s3-demo-bucket,S3Key=my-s3-key
    --engine-version OpenSearch_2.15
"PackageConfigOptions": {
     ...
  }
  "PackageEncryptionOptions": {
    "Enabled": true,
    "KmsKeyId":"arn:aws:kms:us-east-2:111222333444:key/2ba228d5-1d09-456c-ash9-
daf42EXAMPLE"
  }
```

Esempio: aggiorna un plug-in personalizzato con AWS KMS chiave di sicurezza

```
aws opensearch update-package \
    --region us-east-2  --package-name my-custom-package \
    --package-type ZIP-PLUGIN \
    --package-source S3BucketName=amzn-s3-demo-bucket,S3Key=my-s3-key
    --engine-version OpenSearch_2.15
"PackageConfigOptions": {
     ...
  }
  "PackageEncryptionOptions": {
    "Enabled": true,
```

```
    "KmsKeyId":"arn:aws:kms:us-east-2:111222333444:key/2ba228d5-1d09-456c-ash9-
daf42EXAMPLE"
  }
```

> **Important**
>
> Se la AWS KMS chiave specificata è disabilitata o eliminata, può lasciare il cluster associato inoperativo.

Per ulteriori informazioni sull' AWS KMS integrazione con pacchetti personalizzati,the section called “AWS KMS Integrazione di pacchetti personalizzati Amazon OpenSearch Service”.

Aggiornamento di un dominio OpenSearch di servizio con plug-in personalizzati a una versione successiva di utilizzo di OpenSearch AWS CLI

Quando devi aggiornare un dominio di OpenSearch servizio che utilizza plug-in personalizzati a una versione successiva di OpenSearch, completa i seguenti processi.

Per aggiornare un dominio OpenSearch di servizio con plug-in personalizzati a una versione successiva utilizzando il OpenSearch AWS CLI

1. Usa il comando create-package per creare un nuovo pacchetto per il tuo plugin specificando la nuova versione. OpenSearch

   Assicurati che il nome del pacchetto sia lo stesso per il plugin per tutte le versioni del motore. La modifica del nome del pacchetto causa il fallimento del processo di aggiornamento del dominio durante la blue/green distribuzione.
2. Aggiorna il tuo dominio alla versione superiore seguendo la procedura riportata di seguitothe section called “Aggiornamento dei domini”.

   Durante questo processo, Amazon OpenSearch Service dissocia la versione precedente del pacchetto di plug-in e installa la nuova versione utilizzando una distribuzione. blue/green

Disinstallazione e visualizzazione dello stato di dissociazione di un plug-in personalizzato

Per disinstallare il plugin da qualsiasi dominio, puoi usare il comando dissociate-package.
L'esecuzione di questo comando rimuove anche qualsiasi pacchetto di configurazione o licenza

correlato. È quindi possibile utilizzare il list-packages-for-domaincomando per visualizzare lo stato della dissociazione.

> **Tip**
>
> È inoltre possibile utilizzare il comando dissociate-packages per disinstallare più plugin da un dominio in un'unica operazione.

Per disinstallare e visualizzare lo stato di dissociazione di un plug-in personalizzato

1. Disabilita il plugin in ogni indice. Questa operazione deve essere eseguita prima di dissociare il pacchetto del plugin.

   Se provate a disinstallare un plugin prima di disabilitarlo da tutti gli indici, il processo di blue/green distribuzione rimane bloccato nello stato. `Processing`
2. Eseguite il comando seguente per disinstallare il plugin. Sostituisci *`placeholder values`* con le informazioni appropriate.

   ```
   aws opensearch dissociate-package \
       --region region \
       --package-id plugin-package-id \
       --domain-name domain name
   ```
3. (Facoltativo) Eseguite il list-packages-for-domaincomando per visualizzare lo stato della dissociazione.

AWS KMS Integrazione di pacchetti personalizzati Amazon OpenSearch Service

I pacchetti personalizzati di Amazon OpenSearch Service forniscono la crittografia di default per proteggere `ZIP-PLUGIN` i pacchetti inutilizzati Chiavi gestite da AWS.

- Chiavi di proprietà di AWS— I pacchetti personalizzati di Amazon OpenSearch Service utilizzano queste chiavi per impostazione predefinita per crittografare automaticamente i `ZIP-PLUGIN` pacchetti. Non puoi visualizzare, gestire, utilizzare Chiavi di proprietà di AWS o controllare il loro utilizzo. Tuttavia, non è necessario effettuare alcuna operazione o modificare programmi per proteggere le chiavi che eseguono la crittografia dei dati. Per ulteriori informazioni, consulta la sezione Chiavi di proprietà di AWS nella Guida per gli sviluppatori di AWS Key Management Service .

- Chiavi gestite dal cliente: puoi aggiungere un secondo livello di crittografia rispetto a quello esistente Chiavi di proprietà di AWS scegliendo una chiave gestita dal cliente al momento della creazione del pacchetto `ZIP-PLUGIN` personalizzato.

  I pacchetti personalizzati di Amazon OpenSearch Service supportano l'utilizzo di una chiave simmetrica gestita dal cliente che crei, possiedi e gestisci per aggiungere un secondo livello di crittografia alla crittografia di AWS proprietà esistente. Poiché hai il pieno controllo di questo livello di crittografia, puoi eseguire le seguenti attività:

  - Stabilire e mantenere le policy della chiave
  - Stabilire e mantenere politiche e sovvenzioni AWS Identity and Access Management (IAM)
  - Abilitare e disabilitare le policy della chiave
  - Ruotare i materiali crittografici delle chiavi
  - Aggiungere tag
  - Creare alias delle chiavi
  - Pianificare l'eliminazione di chiavi

Per ulteriori informazioni, consulta Customer managed keys nella Guida per sviluppatori AWS Key Management Service .

> **Note**
>
> I pacchetti personalizzati di Amazon OpenSearch Service abilitano automaticamente e gratuitamente la crittografia Chiavi di proprietà di AWS a riposo. Tuttavia, quando si utilizza una chiave gestita dal cliente, vengono applicati dei AWS KMS costi. Per ulteriori informazioni sui prezzi, consulta Prezzi di AWS Key Management Service.

In che modo il OpenSearch servizio di pacchetti personalizzati di Amazon Service utilizza le sovvenzioni in AWS KMS

OpenSearch I pacchetti personalizzati di assistenza richiedono una concessione per utilizzare la chiave gestita dal cliente.

Quando crei un `ZIP-PLUGIN` pacchetto crittografato con una chiave gestita dal cliente, il OpenSearch servizio di pacchetti personalizzati di Amazon Service crea una sovvenzione per tuo conto inviando una CreateGrantrichiesta a AWS KMS. Concede a AWS KMS Ingive OpenSearch Service l'accesso a una AWS KMS chiave del tuo account. Le sovvenzioni create dai pacchetti

personalizzati del OpenSearch Servizio hanno un vincolo che consente le operazioni solo quando la richiesta include un contesto di crittografia con l'ID del pacchetto personalizzato.

I pacchetti personalizzati di Amazon OpenSearch Service richiedono la concessione dell'utilizzo della chiave gestita dal cliente per le seguenti operazioni interne:

Operation	Description
`DescribeKey`	Invia `DescribeKey` richieste per AWS KMS verificare che l'ID della chiave simmetrica gestita dal cliente inserito durante la creazione del pacchetto di plug-in sia valido.
`GenerateDataKeyWithoutPlaintext`	Invia `GenerateDataKeyWithoutPlaintext` richieste per AWS KMS generare chiavi dati crittografate dalla chiave gestita dal cliente.
`GenerateDataKey`	Invia `GenerateDataKey` richieste per AWS KMS generare chiavi di dati per crittografare il pacchetto durante la copia interna.
`Decrypt`	Invia `Decrypt` richieste per AWS KMS decrittografare le chiavi di dati crittografate in modo che possano essere utilizzate per decrittografare i dati.

Puoi revocare l'accesso alla concessione o rimuovere l'accesso del servizio alla chiave gestita dal cliente in qualsiasi momento. In tal caso, il OpenSearch Servizio non sarà in grado di accedere ai dati crittografati dalla chiave gestita dal cliente, il che influirà sulle operazioni che dipendono da tali dati. Ad esempio, se tenti di associare un pacchetto di plug-in a cui OpenSearch Service non può accedere, l'operazione restituisce un `AccessDeniedException` errore.

Creazione di una chiave gestita dal cliente

È possibile creare una chiave simmetrica gestita dal cliente utilizzando Console di gestione AWS o il. AWS KMS APIs

Per creare una chiave simmetrica gestita dal cliente

- Segui i passaggi descritti nella sezione Creazione di una chiave KMS nella Guida per gli AWS Key Management Service sviluppatori.

Policy della chiave

Le policy della chiave controllano l'accesso alla chiave gestita dal cliente. Ogni chiave gestita dal cliente deve avere esattamente una policy della chiave, che contiene istruzioni che determinano chi può usare la chiave e come la possono usare. Quando crei la chiave gestita dal cliente, è possibile specificare una policy della chiave. Per ulteriori informazioni, consulta Policy delle chiavi in AWS KMS nella Guida per gli sviluppatori di AWS Key Management Service .

Per utilizzare la chiave gestita dal cliente con le risorse del plug-in, è necessario consentire le seguenti operazioni API nella politica chiave:

- `kms:CreateGrant`: aggiunge una concessione a una chiave gestita dal cliente. Concede l'accesso di controllo a una AWS KMS chiave specificata, consentendo l'accesso alle operazioni di concessione richieste dai pacchetti personalizzati del OpenSearch servizio. Per ulteriori informazioni sull'utilizzo delle sovvenzioni, consulta la Guida per gli AWS KMS sviluppatori.

  Ciò consente a OpenSearch Service di eseguire le seguenti operazioni:

  - Chiama `GenerateDataKeyWithoutPlainText` per generare una chiave dati crittografata e archiviarla per ulteriori convalide.
  - Chiama `GenerateDataKey` per copiare internamente il pacchetto del plugin.
  - Chiama `Decrypt` per accedere al pacchetto del plugin internamente.
  - Configurare un principale ritirato per consentire al servizio di `RetireGrant`.
- `kms:DescribeKey`— Fornisce i dettagli chiave gestiti dal cliente per consentire al OpenSearch Servizio di convalidare la chiave.
- `kms:GenerateDataKey`,`kms:GenerateDataKeyWithoutPlaintext`, `kms:Decrypt` — Fornisce ai pacchetti personalizzati di OpenSearch Service l'accesso per utilizzare queste operazioni nella concessione.

Di seguito sono riportati alcuni esempi di dichiarazioni politiche che è possibile aggiungere per i pacchetti personalizzati di OpenSearch Service:

```
"Statement" : [
  {
    "Sid" : "Allow access to principals authorized to use OpenSearch Service custom
 packages",
    "Effect" : "Allow",
    "Principal" : {
```

```
      "AWS" : "*"
    },
    "Action" : [
      "kms:CreateGrant",
      "kms:GenerateDataKey",
      "kms:GenerateDataKeyWithoutPlaintext",
      "kms:Decrypt"
    ],
    "Resource" : "*",
    "Condition" : {
      "StringEquals" : {
        "kms:ViaService" : "custom-packages.region.amazonaws.com"
      },
      "StringEquals" : {
        "kms:EncryptionContext:packageId": "Id of the package"
      }
    }
  },
  {
    "Sid" : "Allow access to principals authorized to use Amazon OpenSearch Service
 custom packages",
    "Effect" : "Allow",
    "Principal" : {
      "AWS" : "*"
    },
    "Action" : [
      "kms:DescribeKey"
    ],
    "Resource" : "*",
    "Condition" : {
      "StringEquals" : {
        "kms:ViaService" : "custom-packages.region.amazonaws.com"
      }
    }
  }
]
```

Per ulteriori informazioni sulla specificazione delle autorizzazioni in una politica, consulta le politiche chiave AWS KMS nella Guida per gli AWS Key Management Service sviluppatori.

Per ulteriori informazioni sulla risoluzione dei problemi di accesso tramite chiave, consulta Risoluzione dei problemi relativi alle AWS KMS autorizzazioni nella Guida per gli AWS Key Management Service sviluppatori.

Specificare una chiave gestita dal cliente per i pacchetti personalizzati OpenSearch di Amazon Service

Puoi specificare una chiave gestita dal cliente come secondo livello di crittografia per `ZIP-PLUGIN` i tuoi pacchetti.

Quando si crea un pacchetto di plug-in, è possibile specificare la chiave dati inserendo un ID di AWS KMS chiave, utilizzato dai pacchetti personalizzati di OpenSearch Service per crittografare il pacchetto di plug-in.

AWS KMS key ID: un identificatore chiave per una chiave gestita AWS KMS dal cliente. Immetti l'ID della chiave, l'ARN della chiave, il nome dell'alias o l'ARN dell'alias.

Contesto di crittografia dei pacchetti personalizzati di Amazon OpenSearch Service

Un contesto di crittografia è un set facoltativo di coppie chiave-valore che contengono ulteriori informazioni contestuali sui dati.

AWS KMS utilizza il contesto di crittografia come dati autenticati aggiuntivi per supportare la crittografia autenticata. Quando includi un contesto di crittografia in una richiesta di crittografia dei dati, AWS KMS associa il contesto di crittografia ai dati crittografati. Per decrittografare i dati, nella richiesta deve essere incluso lo stesso contesto di crittografia.

Contesto di crittografia dei pacchetti personalizzati di Amazon OpenSearch Service

I pacchetti personalizzati di Amazon OpenSearch Service utilizzano lo stesso contesto di crittografia in tutte le operazioni AWS KMS crittografiche, in cui la chiave è `packageId` e il valore è il pacchetto `package-id` di plug-in.

Usa il contesto di crittografia per il monitoraggio

Quando si utilizza una chiave simmetrica gestita dal cliente per crittografare il pacchetto di plug-in, è possibile utilizzare il contesto di crittografia nei record e nei registri di controllo per identificare come viene utilizzata la chiave gestita dal cliente. Il contesto di crittografia appare anche nei log generati da AWS CloudTrail o Amazon CloudWatch Logs.

Utilizzo del contesto di crittografia per controllare l'accesso alla chiave gestita dal cliente

È possibile utilizzare il contesto di crittografia nelle policy delle chiavi e nelle policy IAM come condizioni per controllare l'accesso alla chiave simmetrica gestita dal cliente. È possibile utilizzare i vincoli del contesto di crittografia in una concessione.

OpenSearch I pacchetti personalizzati di servizio utilizzano un vincolo di contesto di crittografia nelle concessioni per controllare l'accesso alla chiave gestita dal cliente nell'account o nella regione. Il vincolo della concessione richiede che le operazioni consentite dalla concessione utilizzino il contesto di crittografia specificato.

Di seguito sono riportati alcuni esempi di istruzioni delle policy della chiave per concedere l'accesso a una chiave gestita dal cliente per un contesto di crittografia specifico. Questa istruzione della policy impone come condizione che le concessioni abbiano un vincolo che specifica il contesto di crittografia.

```
{
    "Sid": "Enable DescribeKey",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/ExampleReadOnlyRole"
    },
    "Action": "kms:DescribeKey",
    "Resource": "*"
},
{
    "Sid": "Enable OpenSearch Service custom packages to use the key",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/ExampleReadOnlyRole"
    },
    "Action" : [
         "kms:CreateGrant",
        "kms:GenerateDataKey",
        "kms:GenerateDataKeyWithoutPlaintext",
        "kms:Decrypt"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals" : {
            "kms:EncryptionContext:packageId": "ID of the package"
```

```
            }
        }
}
```

Monitoraggio delle chiavi di crittografia per OpenSearch il servizio di pacchetti personalizzati

Quando utilizzi una chiave gestita AWS KMS dal cliente con le risorse del OpenSearch servizio Service Custom Packages, puoi utilizzare CloudTrail o CloudWatch Logs per tenere traccia delle richieste a cui vengono inviati i pacchetti OpenSearch personalizzati. AWS KMS

Ulteriori informazioni

Le seguenti risorse forniscono ulteriori informazioni sulla crittografia dei dati a riposo.

- Per ulteriori informazioni sui concetti AWS KMS di base, consulta AWS KMS keysla Guida per gli AWS Key Management Service sviluppatori.
- Per ulteriori informazioni sulle migliori pratiche di sicurezza per AWS KMS, consulta la guida AWS Prescriptive Guidance per le AWS Key Management Service migliori pratiche.

Installazione di plug-in di terze parti in Amazon Service OpenSearch

Amazon OpenSearch Service supporta plug-in di terze parti di partner selezionati. Questi plugin possono migliorare la OpenSearch configurazione con funzionalità aggiuntive come analizzatori personalizzati, tokenizzatori o funzionalità di crittografia. Segui le istruzioni specifiche di installazione e configurazione fornite dagli sviluppatori di terze parti per garantire la corretta integrazione con il tuo dominio di servizio. OpenSearch

> **Note**
>
> È necessario ottenere e mantenere licenze valide direttamente dagli sviluppatori di terze parti. Alcuni provider potrebbero non abilitare completamente i propri plug-in Regioni AWS, quindi verifica la disponibilità con il fornitore del plug-in.

I seguenti plugin di terze parti sono disponibili per l'uso con OpenSearch Service:

- Plugin di crittografia Portal26 (Titanium-lockbox): utilizza la crittografia certificata NIST FIPS 140-2 per crittografare i dati durante l'indicizzazione. Include il supporto Bring Your Own Key (BYOK),

che consente di gestire le chiavi di crittografia per una maggiore sicurezza. Il plugin è fornito da Portal26 e richiede la OpenSearch versione 2.15 o successiva.

- Name Match (RNI): abbina nomi, organizzazioni, indirizzi e date in oltre 24 lingue, il che migliora la sicurezza e la conformità. Il plugin è fornito da Babel Street e richiede la OpenSearch versione 2.15 o successiva.

Argomenti

- Prerequisiti
- Installazione di plugin di terze parti
- Fasi successive

Prerequisiti

Prima di installare un plug-in di terze parti, effettuate le seguenti operazioni:

- Ha ottenuto i file di configurazione e licenza del plug-in e li ha caricati in un bucket Amazon S3. Il bucket deve appartenere allo stesso Regione AWS dominio.
- Un plug-in di terze parti è un tipo di plug-in personalizzato. Assicurati che il dominio soddisfi i prerequisiti per i plugin personalizzati.

Installazione di plugin di terze parti

Per associare un plug-in di terze parti a un dominio di OpenSearch servizio, devi prima caricare tre pacchetti separati: il pacchetto di licenza, il pacchetto di configurazione e il pacchetto del plug-in.

- Il pacchetto di licenza include le informazioni sulla licenza o i metadati associati al plug-in, in formato.json o.xml.
- Il pacchetto di configurazione contiene i file di configurazione del plug-in e le risorse e le impostazioni di supporto. Questi file definiscono il comportamento o l'integrazione del plugin con. OpenSearch
- Il pacchetto del plugin contiene il file binario del plugin compilato, che è il codice eseguibile che OpenSearch viene eseguito. Questo è il fulcro della funzionalità del plugin.

Dopo aver caricato entrambi i pacchetti, puoi associare il plugin e la licenza a un dominio compatibile.

Console

Per associare un plug-in di terze parti a un dominio, importate innanzitutto la licenza e la configurazione del plug-in come pacchetti.

Per installare un plug-in di terze parti

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli Pacchetti.
3. Innanzitutto, importa il pacchetto di licenza. Scegli Importa pacchetto.
4. Per Tipo di pacchetto, scegli Licenza.
5. Per Package source, inserisci il percorso del file JSON o XML della licenza in Amazon S3.
6. Scegli Importa. Il pacchetto viene visualizzato nella scheda Licenze della pagina Pacchetti.
7. Ora, importa la configurazione del plugin. Scegli nuovamente Importa pacchetto.
8. Per Tipo di Package, scegli Configuration.
9. Per Package source, inserisci il percorso del file ZIP di configurazione del plug-in in Amazon S3.
10. Scegli Importa.
11. Infine, importa il plugin stesso. Scegli Importa pacchetto.
12. Per Tipo di pacchetto, scegli Plugin.
13. Per Package source, inserisci il percorso del file ZIP del plug-in in Amazon S3.
14. Seleziona la versione OpenSearch del motore supportata dal plug-in.
15. Scegli Importa.

Per associare un plug-in di terze parti a un dominio

1. Ora associa la licenza e la configurazione del plugin al dominio. Nel riquadro di navigazione a sinistra, scegli Domains (Domini).
2. Scegli il nome del dominio per aprire la configurazione del cluster.
3. Vai alla scheda Plugin.
4. Scegli Associa pacchetti e seleziona il plugin, la licenza e i pacchetti di configurazione che hai appena importato.
5. Scegli Seleziona.
6. Scegli Next (Successivo). Esamina i pacchetti da associare e scegli Associa.

CLI

Innanzitutto, utilizzate il comando create-package per creare un nuovo pacchetto che contenga la licenza del plugin. S3KeyDeve puntare a un file.json o.xml in Amazon S3 che includa il testo o i metadati della licenza.

```
aws opensearch create-package \
  --package-name plugin-license-package \
  --package-type PACKAGE-LICENSE \
  --package-source S3BucketName=my-bucket,S3Key=licenses/my-plugin-license.json
```

Usa nuovamente il comando create-package per creare un pacchetto che contenga la configurazione del plugin. S3KeyDeve puntare a un file.zip in Amazon S3 che aderisca alla struttura di directory prevista dal plug-in.

```
aws opensearch create-package \
  --package-name plugin-config-package \
  --package-type PACKAGE-CONFIG \
  --package-source S3BucketName=my-bucket,S3Key=path/to/package.zip
```

Usa nuovamente il comando create-package per creare un pacchetto che contenga il plugin stesso. S3KeyDeve puntare al file.zip del plugin in Amazon S3.

```
aws opensearch create-package \
  --package-name plugin-package \
  --package-type ZIP-PLUGIN \
  --package-source S3BucketName=my-bucket,S3Key=path/to/package.zip
```

Infine, usa il comando associate-package per collegare il plug-in partner, la licenza e la configurazione a un dominio compatibile specificando il pacchetto per ciascuno. IDs Specificate l'ID del plugin come prerequisito per gli altri pacchetti, il che significa che deve essere associato al dominio prima degli altri pacchetti.

```
aws opensearch associate-packages \
  --domain-name my-domain \
  --package-list '[{"PackageID": "plugin-package-id"},{"PackageID": "license-package-
id","PrerequisitePackageIDList":["plugin-package-id"]},{"PackageID":"config-package-
id","PrerequisitePackageIDList":["plugin-package-id"]}]'
```

Fasi successive

Una volta completata l'associazione, potete abilitare il plugin su indici specifici o configurarlo secondo necessità in base alle vostre esigenze. Per applicare la funzionalità dei plug-in di terze parti a indici specifici, modifica le impostazioni dell'indice durante la creazione dell'indice o aggiorna gli indici esistenti. Ad esempio, se il tuo plug-in di terze parti include un analizzatore personalizzato, fai riferimento a esso nelle impostazioni dell'indice.

Per applicare le funzionalità del plug-in in modo uniforme su più indici, utilizzate modelli di indice che includono le configurazioni del plug-in. Consultate sempre la documentazione del plugin per capire come configurarne le funzionalità per la vostra configurazione. OpenSearch

Interrogazione dei dati OpenSearch di Amazon Service con SQL

Puoi usare SQL per interrogare il tuo Amazon OpenSearch Service, anziché usare la OpenSearch query DSL basata su JSON. Le query con SQL sono utili se si ha già familiarità con SQL o si desidera integrare il dominio con un'applicazione che utilizza SQL. Il supporto SQL è disponibile sui domini che eseguono Elasticsearch 6.5 OpenSearch o versioni successive.

> **Note**
>
> Questa documentazione descrive la compatibilità delle versioni tra OpenSearch Service e varie versioni del plug-in SQL, nonché il driver JDBC e ODBC. Consulta la OpenSearchdocumentazione open source per informazioni sulla sintassi di query, funzioni, query di metadati e funzioni aggregate di base e complesse.

Usa la tabella seguente per trovare la versione del plugin SQL supportata da ciascuna versione e da Elasticsearch. OpenSearch

OpenSearch

OpenSearch versione	Versione del plug-in SQL	Caratteristiche da tenere in considerazione
2.19.0	2.19.0.0	
2.18.0	2,180,0	

OpenSearch versione	Versione del plug-in SQL	Caratteristiche da tenere in considerazione
2.17.0	2,170,0	
2.15.0	2,150,0	
2.13.0	2,130,0	
2.11.0	211,0,0	Aggiunge il supporto per il linguaggio e le query PPL
2.9.0	2.9.0.0	Aggiungi il connettore Spark e supporta le funzioni tabellari e PromQL
2.7.0	2.7.0.0	Aggiungi API `datasource`
2.5.0	2.5.0.0	
2.3.0	2,30,0	Aggiungi funzioni di datetime `maketime` e `makedate`
1.3.0	1,30.0	Supporta la dimensione limite predefinita della query e la clausola IN per eseguire la selezione all'interno di un elenco di valori
1.2.0	1,20.0	Aggiungi un nuovo protocollo per il formato della risposta di visualizzazione
1.1.0	1.1.0.0	Supporto per la funzione di corrispondenza come filtro in SQL e PPL
1.0.0	1.0.0.0	Supporto per l'interrogazione di un flusso dei dati

Open Distro per Elasticsearch

Versione di Elasticsearch	Versione del plug-in SQL	Caratteristiche da tenere in considerazione
7,10	1,13,0	NULL FIRST e LAST per le funzioni finestra, la funzione CAST(), i comandi SHOW e DESCRIBE

Versione di Elasticsearch	Versione del plug-in SQL	Caratteristiche da tenere in considerazione
7.9	1.11.0	Aggiungi date/time funzioni aggiuntive, ORDER BY keyword
7.8	1.9.0	
7.7	18.0	
7.3	1.3.0	Operatori multipli di stringhe e numeri
7.1	1.1.0	

Chiamata di esempio

Per eseguire query sui dati con SQL, inviare le richieste HTTP a `_sql` utilizzando il seguente formato:

```
POST domain-endpoint/_plugins/_sql
{
  "query": "SELECT * FROM my-index LIMIT 50"
}
```

> **Note**
>
> Se il tuo dominio utilizza Elasticsearch anziché Elasticsearch OpenSearch, il formato è. `_opendistro/_sql`

Note e differenze

Le chiamate a `_plugins/_sql` includono i nomi degli indici nel corpo della richiesta e quindi hanno le stesse considerazioni delle policy di accesso delle operazioni bulk, mget e msearch. Come sempre, seguire il principio del privilegio minimo quando si concedono autorizzazioni alle operazioni API.

Per le considerazioni sulla sicurezza relativa all'utilizzo di SQL con il controllo granulare degli accessi, consultare the section called “Controllo granulare degli accessi”.

Il plugin OpenSearch SQL include molte impostazioni regolabili. In OpenSearch Service, utilizzate il `_cluster/settings` percorso, non il percorso delle impostazioni del plugin path (`_plugins/_query/settings`):

```
PUT _cluster/settings
{
  "transient" : {
    "plugins.sql.enabled" : true
  }
}
```

Per i domini Elasticsearch legacy, sostituisci `plugins` con `opendistro`:

```
PUT _cluster/settings
{
  "transient" : {
    "opendistro.sql.enabled" : true
  }
}
```

SQL Workbench

SQL Workbench è un'interfaccia utente di OpenSearch Dashboards che consente di eseguire query SQL su richiesta, tradurre SQL nel suo equivalente REST e visualizzare e salvare i risultati come testo, JSON, JDBC o CSV. Per ulteriori informazioni, consultare Query Workbench.

SQL CLI

SQL CLI è un'applicazione Python autonoma che è possibile avviare con il comando `opensearchsql`. Per la procedura di installazione, configurazione e utilizzo, consulta SQL CLI.

Driver JDBC

Il driver Java Database Connectivity (JDBC) consente di integrare i domini di OpenSearch servizio con le applicazioni di business intelligence (BI) preferite. Per scaricare il driver, fai clic qui. Per ulteriori informazioni, consulta il repository. GitHub

Driver ODBC

Il driver Open Database Connectivity (ODBC) è un driver ODBC di sola lettura per Windows e macOS che consente di collegare applicazioni di business intelligence e visualizzazione dei dati come Microsoft Excel al plug-in SQL.

Per informazioni sull'installazione del driver, consulta l'archivio SQL su. GitHub

Ricerca tra cluster in Amazon Service OpenSearch

La ricerca tra cluster in Amazon OpenSearch Service ti consente di eseguire query e aggregazioni su più domini connessi. Spesso ha più senso utilizzare più domini più piccoli invece di un singolo dominio di grandi dimensioni, soprattutto quando si eseguono diversi tipi di carichi di lavoro.

I domini specifici del carico di lavoro consentono di completare le seguenti attività:

- Ottimizzare ogni dominio scegliendo i tipi di istanza per carichi di lavoro specifici.
- Stabilire i limiti di isolamento del guasto tra i carichi di lavoro. Ciò significa che se uno dei carichi di lavoro genera errori, il guasto è contenuto all'interno di quel dominio specifico e non influisce sugli altri carichi di lavoro.
- Scalabilità più semplice tra domini.

La ricerca tra cluster supporta OpenSearch le dashboard, quindi puoi creare visualizzazioni e dashboard in tutti i tuoi domini. Paghi le tariffe standard per il trasferimento AWS dei dati per i risultati di ricerca trasferiti tra domini.

> **Note**
>
> L'open source dispone OpenSearch anche di documentazione per la ricerca tra cluster. La configurazione differisce notevolmente per i cluster open source rispetto ai domini Amazon OpenSearch Service gestiti. In particolare, in OpenSearch Service, si configurano le connessioni tra cluster utilizzando cURL Console di gestione AWS anziché tramite cURL. Inoltre, il servizio gestito utilizza AWS Identity and Access Management (IAM) per l'autenticazione tra cluster oltre al controllo granulare degli accessi. Pertanto, consigliamo di utilizzare questa documentazione, anziché la documentazione open source, per configurare la OpenSearch ricerca tra cluster per i domini.

Argomenti

- Limitazioni
- Prerequisiti di ricerca tra cluster
- Prezzi della funzionalità di ricerca tra cluster
- Configurazione di una connessione
- Rimozione di una connessione
- Configurazione della sicurezza e spiegazione passo per passo di esempio
- OpenSearch Pannelli di controllo

Limitazioni

La ricerca tra cluster presenta diverse limitazioni importanti:

- Non puoi connettere un dominio Elasticsearch a un dominio. OpenSearch
- Non puoi connetterti a cluster OpenSearch/Elasticsearch autogestiti.
- Per connettere domini tra regioni, entrambi i domini devono essere su Elasticsearch 7.10 o versione successiva oppure. OpenSearch
- Un dominio può avere un massimo di 20 connessioni in uscita. Analogamente, un dominio può avere un massimo di 20 connessioni in ingresso. In altre parole, un dominio può connettersi a un massimo di 20 altri domini.
- Il dominio di origine deve avere la stessa versione o una versione successiva del dominio di destinazione. Se configuri una connessione bidirezionale tra due domini e desideri aggiornarne uno o entrambi, devi prima eliminare una delle connessioni.
- Non è possibile utilizzare dizionari personalizzati o SQL con la ricerca tra cluster.
- Non puoi utilizzare CloudFormation per connettere domini.
- Non è possibile utilizzare la ricerca tra cluster su istanze M3 o istanze espandibili (T2 e T3).

Prerequisiti di ricerca tra cluster

Prima di configurare la ricerca tra cluster, assicurarsi che i domini soddisfino i seguenti requisiti:

- Due OpenSearch domini o domini Elasticsearch nella versione 6.7 o successiva
- Controllo granulare degli accessi abilitato

- Node-to-node crittografia abilitata
- Se uno dei domini si trova in un VPC, i domini devono essere connessi tramite VPC Peering o Transit Gateway e i gruppi di sicurezza devono consentire il traffico tra di loro.

Prezzi della funzionalità di ricerca tra cluster

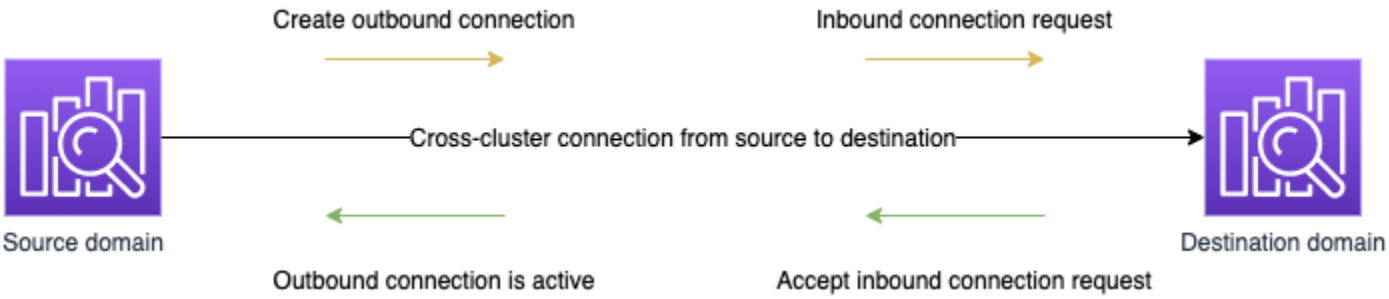
Non è previsto alcun costo aggiuntivo per la ricerca tra domini.

Configurazione di una connessione

Il dominio "di origine" si riferisce al dominio da cui proviene una richiesta di ricerca tra cluster. In altre parole, il dominio di origine è quello a cui si invia la richiesta iniziale di ricerca.

Il dominio "di destinazione" è il dominio sul quale il dominio di origine esegue le query.

Una connessione tra cluster è unidirezionale dall'origine al dominio di destinazione. Ciò significa che il dominio di destinazione non può eseguire query sul dominio di origine. Tuttavia, è possibile impostare un'altra connessione nella direzione opposta.



Il dominio di origine crea una connessione "in uscita" al dominio di destinazione. Il dominio di destinazione riceve una richiesta di connessione "in ingresso" dal dominio di origine.

Per impostare una connessione

1. Nel pannello di controllo del dominio scegliere un dominio e selezionare la scheda Connessioni.
2. Nella sezione Connessioni in uscita scegli Richiesta.
3. Per Alias di connessione, specifica un nome per la connessione.
4. Scegli se connetterti a un dominio nella tua Account AWS regione o in un altro account o regione.

    - Per connetterti a un cluster nella tua regione Account AWS e nella tua regione, seleziona il dominio dal menu a discesa e scegli Richiedi.

- Per connetterti a un cluster in un'altra regione Account AWS o in un'altra regione, seleziona l'ARN del dominio remoto e scegli Richiesta. Per connettere domini tra regioni, entrambi i domini devono eseguire Elasticsearch versione 7.10 o successiva oppure. OpenSearch

5. Per ignorare i cluster non disponibili per le query sui cluster, seleziona Ignora non disponibile. Questa impostazione garantisce che le query tra cluster restituiscano risultati parziali nonostante gli errori su uno o più cluster remoti.
6. La ricerca tra cluster convalida innanzitutto la richiesta di connessione per assicurare che i prerequisiti siano soddisfatti. Se i domini risultano incompatibili, la richiesta di connessione entra nello stato `Validation failed`.
7. Dopo che la richiesta di connessione è stata convalidata correttamente, viene inviata al dominio di destinazione, dove deve essere approvata. Fino a quando non si avviene questa approvazione, la connessione rimane in uno stato `Pending acceptance`. Quando la richiesta di connessione viene accettata nel dominio di destinazione, lo stato cambia in `Active` e il dominio di destinazione diventa disponibile per le query.

   - La pagina del dominio mostra i dettagli generali dello stato del dominio e dello stato dell'istanza del dominio di destinazione. Solo i proprietari dei domini dispongono della flessibilità necessaria per creare, visualizzare, rimuovere e monitorare le connessioni da o verso i propri domini.

Dopo aver stabilito la connessione, tutto il traffico che scorre tra i nodi dei domini connessi viene crittografato. Se si connette un dominio VPC a un dominio non VPC e il dominio non VPC è un endpoint pubblico in grado di ricevere traffico da Internet, il traffico tra cluster tra i domini è ancora crittografato e protetto.

Rimozione di una connessione

La rimozione di una connessione interrompe qualsiasi operazione tra cluster sui relativi indici.

1. Nel pannello di controllo del dominio, passa alla scheda Connessioni.
2. Selezionare le connessioni di dominio che si desidera rimuovere e scegliere Elimina, quindi confermare l'eliminazione.

È possibile eseguire queste operazioni sul dominio di origine o di destinazione per rimuovere la connessione. Dopo aver rimosso la connessione, questa sarà ancora visibile con lo stato `Deleted` per altri 15 giorni.

Non è possibile eliminare un dominio con connessioni tra cluster attive. Per eliminare un dominio, rimuovere innanzitutto tutte le connessioni in ingresso e in uscita da tale dominio. Questo per essere certi di considerare gli utenti del dominio tra cluster prima di eliminare il dominio.

Configurazione della sicurezza e spiegazione passo per passo di esempio

1. Si invia una richiesta di ricerca tra cluster al dominio di origine.
2. Il dominio di origine valuta tale richiesta in base alla policy di accesso al dominio. Poiché la ricerca tra cluster richiede un controllo di accesso granulare, è consigliabile una policy di accesso aperta nel dominio di origine.

   JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "*"
        ]
      },
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/src-domain/*"
    }
  ]
}
```

> **Note**
>
> Se includi indici remoti nel percorso, devi codificare in formato URL l'URI nell'ARN del dominio. Ad esempio, utilizzare `arn:aws:es:us-east-1:123456789012:domain/my-domain/local_index,dst%3Aremote_index` anziché `arn:aws:es:us-east-1:123456789012:domain/my-domain/local_index,dst:remote_index`.

Se si sceglie di utilizzare una policy di accesso restrittiva oltre al controllo di accesso granulare, la policy deve consentire l'accesso almeno a `es:ESHttpGet`.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "arn:aws:iam::111122223333:user/test-user"
        ]
      },
      "Action": "es:ESHttpGet",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/src-domain/*"
    }
  ]
}
```

3. Il controllo di accesso granulare nel dominio di origine valuta la richiesta:

   - La richiesta è firmata con credenziali di base IAM o HTTP valide?
   - In tal caso, l'utente dispone dell'autorizzazione per eseguire la ricerca e accedere ai dati?

   Se la richiesta ricerca solo i dati nel dominio di destinazione, ad esempio `dest-alias:dest-index/_search`, sono necessarie solo le autorizzazioni per il dominio di destinazione.

   Se la richiesta cerca dati su entrambi i domini, ad esempio `source-index,dest-alias:dest-index/_search`, sono necessarie le autorizzazioni per entrambi i domini.

   Nel controllo granulare degli accessi, gli utenti devono disporre dell'`indices:admin/shards/search_shards`autorizzazione oltre alle autorizzazioni standard `read` o `search` alle autorizzazioni per gli indici pertinenti.

4. Il dominio di origine invia la richiesta al dominio di destinazione. Il dominio di destinazione valuta la richiesta in base alla policy di accesso al dominio. È necessario includere l'autorizzazione `es:ESCrossClusterGet` per il dominio di destinazione:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
      "Action": "es:ESCrossClusterGet",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/dst-domain"
    }
  ]
}
```

Assicurarsi che l'autorizzazione `es:ESCrossClusterGet` sia applicata per `/dst-domain` e non per `/dst-domain/*`.

Tuttavia, questo policy minima consente solo ricerche tra cluster. Per eseguire altre operazioni, ad esempio l'indicizzazione dei documenti e l'esecuzione di ricerche standard, sono necessarie ulteriori autorizzazioni. Si consiglia la seguente policy per il dominio di destinazione:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "*"
        ]
      },
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/dst-domain/*"
```

```
    },
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
      "Action": "es:ESCrossClusterGet",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/dst-domain"
    }
  ]
}
```

> **Note**
>
> Per impostazione predefinita, tutte le richieste di ricerca intercluster tra domini vengono crittografate in transito come parte della crittografia. node-to-node

5. Il dominio di destinazione esegue la ricerca e restituisce i risultati al dominio di origine.
6. Il dominio di origine combina i propri risultati (se presenti) con i risultati del dominio di destinazione e li restituisce all'utente.
7. Consigliamo Postman per le richieste di test:

   - Nel dominio di destinazione, indicizzare un documento:

   ```
   POST https://dst-domain.us-east-1.es.amazonaws.com/books/_doc/1

   {
     "Dracula": "Bram Stoker"
   }
   ```

   - Per eseguire una query su questo indice dal dominio di origine, includere l'alias di connessione del dominio di destinazione all'interno della query.

   ```
   GET https://src-domain.us-east-1.es.amazonaws.com/<connection_alias>:books/
   _search

   {
       ...
     "hits": [
       {
   ```

```
      "_index": "source-destination:books",
      "_type": "_doc",
      "_id": "1",
      "_score": 1,
      "_source": {
        "Dracula": "Bram Stoker"
      }
    }
  ]
}
```

È possibile trovare l'alias di connessione nella scheda Connessioni nel pannello di controllo del dominio.

- Se si imposta una connessione tra `domain-a -> domain-b` con l'alias di connessione `cluster_b` e `domain-a -> domain-c` con l'alias di connessione `cluster_c`, ricercare `domain-a`, `domain-b` e `domain-c` come segue:

```
GET https://src-domain.us-east-1.es.amazonaws.com/
local_index,cluster_b:b_index,cluster_c:c_index/_search
{
  "query": {
    "match": {
      "user": "domino"
    }
  }
}
```

Risposta

```
{
  "took": 150,
  "timed_out": false,
  "_shards": {
    "total": 3,
    "successful": 3,
    "failed": 0,
    "skipped": 0
  },
  "_clusters": {
    "total": 3,
    "successful": 3,
```

```
    "skipped": 0
  },
  "hits": {
    "total": 3,
    "max_score": 1,
    "hits": [
      {
        "_index": "local_index",
        "_type": "_doc",
        "_id": "0",
        "_score": 1,
        "_source": {
          "user": "domino",
          "message": "This is message 1",
          "likes": 0
        }
      },
      {
        "_index": "cluster_b:b_index",
        "_type": "_doc",
        "_id": "0",
        "_score": 2,
        "_source": {
          "user": "domino",
          "message": "This is message 2",
          "likes": 0
        }
      },
      {
        "_index": "cluster_c:c_index",
        "_type": "_doc",
        "_id": "0",
        "_score": 3,
        "_source": {
          "user": "domino",
          "message": "This is message 3",
          "likes": 0
        }
      }
    ]
  }
}
```

Se non hai scelto di ignorare i cluster non disponibili nella configurazione della connessione, tutti i cluster di destinazione che cerchi devono essere disponibili affinché la richiesta di ricerca venga eseguita correttamente. In caso contrario, l'intera richiesta avrà esito negativo: anche se uno dei domini non è disponibile non verrà restituito alcun risultato della ricerca.

OpenSearch Pannelli di controllo

È possibile visualizzare i dati di più domini connessi nello stesso modo di un singolo dominio, tranne per il fatto che è necessario accedere agli indici remoti utilizzando `connection-alias:index`. Quindi, il modello di indice deve corrispondere a `connection-alias:index`.

Imparare a classificarsi per Amazon OpenSearch Service

OpenSearch utilizza un framework di classificazione probabilistico chiamato BM-25 per calcolare i punteggi di rilevanza. Se una parola chiave distintiva appare più frequentemente in un documento, BM-25 assegna un punteggio di pertinenza più elevato a tale documento. Questo framework, tuttavia, non tiene conto del comportamento degli utenti come i dati click-through, che possono aumentare ulteriormente la rilevanza.

Learning to Rank è un plug-in OpenSearch open source che consente di utilizzare il machine learning e i dati comportamentali per ottimizzare la pertinenza dei documenti. Utilizza modelli delle librerie XGBoost e Ranklib per riscrivere i risultati della ricerca. Il plugin Elasticsearch LTR è stato inizialmente sviluppato da OpenSource Connections, con contributi significativi da Wikimedia Foundation, Snagajob Engineering, Bonsai e Yelp Engineering. La OpenSearch versione del plugin è derivata dal plugin Elasticsearch LTR.

Learning to Rank richiede Elasticsearch OpenSearch 7.7 o versione successiva. Per utilizzare il plug-in Learning to Rank, è necessario disporre delle autorizzazioni complete di amministratore. Per ulteriori informazioni, consulta the section called “Modifica dell'utente principale”.

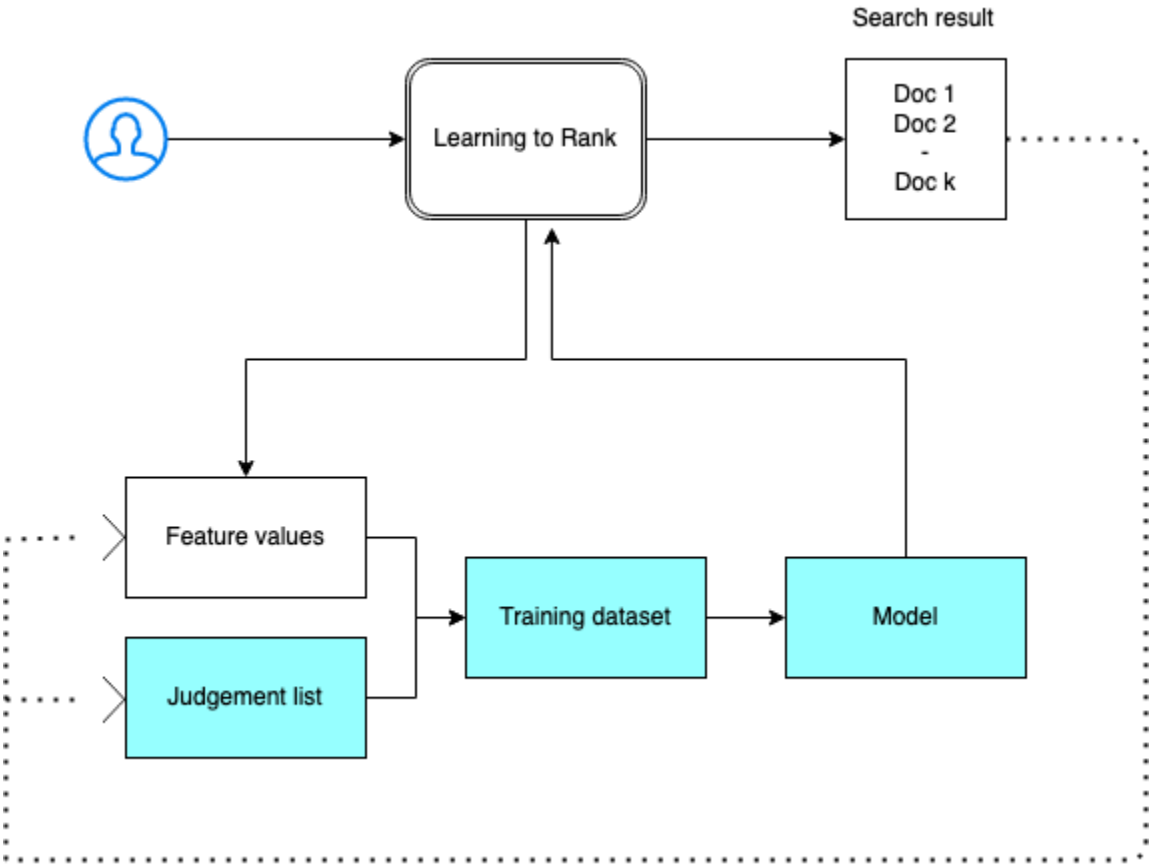
> **Note**
>
> Questa documentazione fornisce una panoramica generale del plugin Learning to Rank e ti aiuta a iniziare a usarlo. La documentazione completa, comprese le descrizioni delle fasi e dell'API, è disponibile nella documentazione di Learning to Rank.

Argomenti

- Nozioni di base su Learning to Rank
- API Learning to Rank

Nozioni di base su Learning to Rank

Devi fornire un elenco di giudizi, preparare un set di dati di addestramento e addestrare il modello al di fuori di Amazon OpenSearch Service. Le parti in blu si trovano al di fuori del OpenSearch servizio:



Fase 1: Inizializzazione del plug-in

Per inizializzare il plugin Learning to Rank, invia la seguente richiesta al tuo dominio OpenSearch di servizio:

```
PUT _ltr
```

```
{
  "acknowledged" : true,
```

```
  "shards_acknowledged" : true,
  "index" : ".ltrstore"
}
```

Questo comando crea un indice `.ltrstore` nascosto che memorizza informazioni sui metadati, ad esempio set di funzioni e modelli.

Fase 2: Creazione di un elenco dei giudizi

> **Note**
>
> È necessario eseguire questo passaggio al di fuori del OpenSearch Servizio.

Un elenco di giudizi è una raccolta di esempi da cui apprende un modello di machine learning. L'elenco dei giudizi deve includere parole chiave importanti e un insieme di documenti classificati per ogni parola chiave.

In questo esempio, abbiamo un elenco di giudizi per un set di dati di film. Un grado pari a 4 indica una corrispondenza perfetta. Un grado pari a 0 indica la corrispondenza peggiore.

Grado	Parola chiave	ID documento	Nome del film
4	rambo	7555	Rambo
3	rambo	1370	Rambo III
3	rambo	1369	Rambo 2 - La vendetta
3	rambo	1368	La vendetta

Preparare l'elenco dei giudizi nel formato seguente:

```
4 qid:1 # 7555 Rambo
3 qid:1 # 1370 Rambo III
3 qid:1 # 1369 Rambo: First Blood Part II
3 qid:1 # 1368 First Blood
```

```
where qid:1 represents "rambo"
```

Per un esempio più completo di un elenco dei giudizi, consultare Giudizi di film.

È possibile creare questo elenco di giudizi manualmente con l'aiuto di annotatori umani o dedurlo a livello di programmazione dai dati analitici.

Fase 3: Creazione di un set di funzionalità

Una funzionalità è un campo che corrisponde alla pertinenza di un documento, ad esempio `title`, `overview`, `popularity score`(numero di visualizzazioni) e così via.

Creare un set di funzionalità con un modello Mustache per ogni funzionalità. Per ulteriori informazioni sulle funzionalità, consultare Utilizzo delle funzionalità.

In questo esempio viene creato un set di funzionalità `movie_features` con i campi `title` e `overview`:

```
POST _ltr/_featureset/movie_features
{
  "featureset" : {
      "name" : "movie_features",
      "features" : [
        {
          "name" : "1",
          "params" : [
            "keywords"
          ],
          "template_language" : "mustache",
          "template" : {
            "match" : {
              "title" : "{{keywords}}"
            }
          }
        },
        {
          "name" : "2",
          "params" : [
            "keywords"
          ],
          "template_language" : "mustache",
          "template" : {
            "match" : {
```

```
              "overview" : "{{keywords}}"
            }
          }
        }
      ]
    }
}
```

Se si esegue una query sull'indice `.ltrstore` originale, viene recuperato il set di funzionalità:

```
GET _ltr/_featureset
```

Fase 4: Registrazione dei valori delle funzionalità

I valori delle funzionalità sono i punteggi di pertinenza calcolati da BM-25 per ogni funzionalità.

Combinare il set di funzionalità e l'elenco dei giudizi per registrare i valori delle funzionalità. Per ulteriori informazioni sulle funzionalità di registrazione, consultare Punteggi della funzionalità di registrazione.

In questo esempio, `bool` recupera i documenti classificati con il filtro e quindi seleziona il set di funzionalità con la query `sltr`. La query `ltr_log` combina i documenti e le funzionalità in modo da registrare i valori delle funzionalità corrispondenti:

```
POST tmdb/_search
{
  "_source": {
    "includes": [
      "title",
      "overview"
    ]
  },
  "query": {
    "bool": {
      "filter": [
        {
          "terms": {
            "_id": [
              "7555",
              "1370",
              "1369",
              "1368"
```

```
            ]
          }
        },
        {
          "sltr": {
            "_name": "logged_featureset",
            "featureset": "movie_features",
            "params": {
              "keywords": "rambo"
            }
          }
        }
      ]
    }
  },
  "ext": {
    "ltr_log": {
      "log_specs": {
        "name": "log_entry1",
        "named_query": "logged_featureset"
      }
    }
  }
}
```

L'aspetto di una risposta di esempio è simile al seguente:

```
{
  "took" : 7,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 4,
      "relation" : "eq"
    },
    "max_score" : 0.0,
    "hits" : [
```

```
    {
      "_index" : "tmdb",
      "_type" : "movie",
      "_id" : "1368",
      "_score" : 0.0,
      "_source" : {
        "overview" : "When former Green Beret John Rambo is harassed by local law
enforcement and arrested for vagrancy, the Vietnam vet snaps, runs for the hills and
rat-a-tat-tats his way into the action-movie hall of fame. Hounded by a relentless
sheriff, Rambo employs heavy-handed guerilla tactics to shake the cops off his tail.",
        "title" : "First Blood"
      },
      "fields" : {
        "_ltrlog" : [
          {
            "log_entry1" : [
              {
                "name" : "1"
              },
              {
                "name" : "2",
                "value" : 10.558305
              }
            ]
          }
        ]
      },
      "matched_queries" : [
        "logged_featureset"
      ]
    },
    {
      "_index" : "tmdb",
      "_type" : "movie",
      "_id" : "7555",
      "_score" : 0.0,
      "_source" : {
        "overview" : "When governments fail to act on behalf of captive missionaries,
ex-Green Beret John James Rambo sets aside his peaceful existence along the Salween
River in a war-torn region of Thailand to take action.  Although he's still haunted
by violent memories of his time as a U.S. soldier during the Vietnam War, Rambo can
hardly turn his back on the aid workers who so desperately need his help.",
        "title" : "Rambo"
      },
```

```
        "fields" : {
          "_ltrlog" : [
            {
              "log_entry1" : [
                {
                  "name" : "1",
                  "value" : 11.2569065
                },
                {
                  "name" : "2",
                  "value" : 9.936821
                }
              ]
            }
          ]
        },
        "matched_queries" : [
          "logged_featureset"
        ]
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "1369",
        "_score" : 0.0,
        "_source" : {
          "overview" : "Col. Troutman recruits ex-Green Beret John Rambo for a highly
secret and dangerous mission. Teamed with Co Bao, Rambo goes deep into Vietnam to
rescue POWs. Deserted by his own team, he's left in a hostile jungle to fight for his
life, avenge the death of a woman and bring corrupt officials to justice.",
          "title" : "Rambo: First Blood Part II"
        },
        "fields" : {
          "_ltrlog" : [
            {
              "log_entry1" : [
                {
                  "name" : "1",
                  "value" : 6.334839
                },
                {
                  "name" : "2",
                  "value" : 10.558305
                }
```

```
              ]
            }
          ]
        },
        "matched_queries" : [
          "logged_featureset"
        ]
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "1370",
        "_score" : 0.0,
        "_source" : {
          "overview" : "Combat has taken its toll on Rambo, but he's finally begun to
 find inner peace in a monastery. When Rambo's friend and mentor Col. Trautman asks for
 his help on a top secret mission to Afghanistan, Rambo declines but must reconsider
 when Trautman is captured.",
          "title" : "Rambo III"
        },
        "fields" : {
          "_ltrlog" : [
            {
              "log_entry1" : [
                {
                  "name" : "1",
                  "value" : 9.425955
                },
                {
                  "name" : "2",
                  "value" : 11.262714
                }
              ]
            }
          ]
        },
        "matched_queries" : [
          "logged_featureset"
        ]
      }
    ]
  }
}
```

Nell'esempio precedente, la prima funzionalità non ha un valore di funzionalità perché la parola chiave "rambo" non viene visualizzata nel campo del titolo del documento con un ID uguale a 1368. Questo è un valore di funzionalità mancante nei dati di addestramento.

Fase 5: Creazione di un set di dati di addestramento

> **Note**
>
> È necessario eseguire questa operazione al di fuori del OpenSearch Servizio.

Il passo successivo consiste nel combinare l'elenco dei giudizi e i valori delle funzionalità per creare un set di dati di addestramento. Se l'elenco dei giudizi originale ha la seguente struttura:

```
4 qid:1 # 7555 Rambo
3 qid:1 # 1370 Rambo III
3 qid:1 # 1369 Rambo: First Blood Part II
3 qid:1 # 1368 First Blood
```

Convertirlo nel set di dati di addestramento finale, che assomiglia a questo:

```
4 qid:1 1:12.318474 2:10.573917 # 7555 rambo
3 qid:1 1:10.357875 2:11.950391 # 1370 rambo
3 qid:1 1:7.010513 2:11.220095 # 1369 rambo
3 qid:1 1:0.0 2:11.220095 # 1368 rambo
```

È possibile eseguire questo passaggio manualmente o scrivere un programma per automatizzarlo.

Fase 6: Scelta di un algoritmo e costruzione del modello

> **Note**
>
> È necessario eseguire questo passaggio al di fuori del OpenSearch Servizio.

Una volta installato il set di dati di addestramento, il passaggio successivo consiste nell'utilizzare XGBoost le nostre librerie Ranklib per creare un modello. XGBoost e le librerie Ranklib consentono di creare modelli popolari come LambdaMart, Random Forests e così via.

Per i passaggi da utilizzare XGBoost e Ranklib per creare il modello, consultate rispettivamente la documentazione e. XGBoostRankLib Per utilizzare Amazon SageMaker per creare il XGBoost modello, consulta XGBoostAlgorithm.

Fase 7: Implementazione del modello

Dopo aver creato il modello, implementarlo nel plug-in Learning to Rank. Per ulteriori informazioni sull'implementazione di un modello, consultare Caricamento di un modello addestrato.

In questo esempio viene creato un modello `my_ranklib_model` utilizzando la libreria Ranklib:

```
POST _ltr/_featureset/movie_features/_createmodel?pretty
{
  "model": {
    "name": "my_ranklib_model",
    "model": {
      "type": "model/ranklib",
      "definition": """## LambdaMART
## No. of trees = 10
## No. of leaves = 10
## No. of threshold candidates = 256
## Learning rate = 0.1
## Stop early = 100

<ensemble>
   <tree id="1" weight="0.1">
      <split>
         <feature>1</feature>
         <threshold>10.357875</threshold>
         <split pos="left">
            <feature>1</feature>
            <threshold>0.0</threshold>
            <split pos="left">
               <output>-2.0</output>
            </split>
            <split pos="right">
               <feature>1</feature>
               <threshold>7.010513</threshold>
               <split pos="left">
                  <output>-2.0</output>
               </split>
               <split pos="right">
                  <output>-2.0</output>
```

```
            </split>
         </split>
      </split>
      <split pos="right">
         <output>2.0</output>
      </split>
   </split>
</tree>
<tree id="2" weight="0.1">
   <split>
      <feature>1</feature>
      <threshold>10.357875</threshold>
      <split pos="left">
         <feature>1</feature>
         <threshold>0.0</threshold>
         <split pos="left">
            <output>-1.67031991481781</output>
         </split>
         <split pos="right">
            <feature>1</feature>
            <threshold>7.010513</threshold>
            <split pos="left">
               <output>-1.67031991481781</output>
            </split>
            <split pos="right">
               <output>-1.6703200340270996</output>
            </split>
         </split>
      </split>
      <split pos="right">
         <output>1.6703201532363892</output>
      </split>
   </split>
</tree>
<tree id="3" weight="0.1">
   <split>
      <feature>2</feature>
      <threshold>10.573917</threshold>
      <split pos="left">
         <output>1.479954481124878</output>
      </split>
      <split pos="right">
         <feature>1</feature>
         <threshold>7.010513</threshold>
```

```
            <split pos="left">
               <feature>1</feature>
               <threshold>0.0</threshold>
               <split pos="left">
                  <output>-1.4799546003341675</output>
               </split>
               <split pos="right">
                  <output>-1.479954481124878</output>
               </split>
            </split>
            <split pos="right">
               <output>-1.479954481124878</output>
            </split>
         </split>
      </split>
   </tree>
   <tree id="4" weight="0.1">
      <split>
         <feature>1</feature>
         <threshold>10.357875</threshold>
         <split pos="left">
            <feature>1</feature>
            <threshold>0.0</threshold>
            <split pos="left">
               <output>-1.3569872379302979</output>
            </split>
            <split pos="right">
               <feature>1</feature>
               <threshold>7.010513</threshold>
               <split pos="left">
                  <output>-1.3569872379302979</output>
               </split>
               <split pos="right">
                  <output>-1.3569872379302979</output>
               </split>
            </split>
         </split>
         <split pos="right">
            <output>1.3569873571395874</output>
         </split>
      </split>
   </tree>
   <tree id="5" weight="0.1">
      <split>
```

```
         <feature>1</feature>
         <threshold>10.357875</threshold>
         <split pos="left">
            <feature>1</feature>
            <threshold>0.0</threshold>
            <split pos="left">
               <output>-1.2721362113952637</output>
            </split>
            <split pos="right">
               <feature>1</feature>
               <threshold>7.010513</threshold>
               <split pos="left">
                  <output>-1.2721363306045532</output>
               </split>
               <split pos="right">
                  <output>-1.2721363306045532</output>
               </split>
            </split>
         </split>
         <split pos="right">
            <output>1.2721362113952637</output>
         </split>
      </split>
   </tree>
   <tree id="6" weight="0.1">
      <split>
         <feature>1</feature>
         <threshold>10.357875</threshold>
         <split pos="left">
            <feature>1</feature>
            <threshold>7.010513</threshold>
            <split pos="left">
               <feature>1</feature>
               <threshold>0.0</threshold>
               <split pos="left">
                  <output>-1.2110036611557007</output>
               </split>
               <split pos="right">
                  <output>-1.2110036611557007</output>
               </split>
            </split>
            <split pos="right">
               <output>-1.2110037803649902</output>
            </split>
```

```
         </split>
         <split pos="right">
            <output>1.2110037803649902</output>
         </split>
      </split>
   </tree>
   <tree id="7" weight="0.1">
      <split>
         <feature>1</feature>
         <threshold>10.357875</threshold>
         <split pos="left">
            <feature>1</feature>
            <threshold>7.010513</threshold>
            <split pos="left">
               <feature>1</feature>
               <threshold>0.0</threshold>
               <split pos="left">
                  <output>-1.165616512298584</output>
               </split>
               <split pos="right">
                  <output>-1.165616512298584</output>
               </split>
            </split>
            <split pos="right">
               <output>-1.165616512298584</output>
            </split>
         </split>
         <split pos="right">
            <output>1.165616512298584</output>
         </split>
      </split>
   </tree>
   <tree id="8" weight="0.1">
      <split>
         <feature>1</feature>
         <threshold>10.357875</threshold>
         <split pos="left">
            <feature>1</feature>
            <threshold>7.010513</threshold>
            <split pos="left">
               <feature>1</feature>
               <threshold>0.0</threshold>
               <split pos="left">
                  <output>-1.131177544593811</output>
```

```
            </split>
            <split pos="right">
               <output>-1.131177544593811</output>
            </split>
         </split>
         <split pos="right">
            <output>-1.131177544593811</output>
         </split>
      </split>
      <split pos="right">
         <output>1.131177544593811</output>
      </split>
   </split>
</tree>
<tree id="9" weight="0.1">
   <split>
      <feature>2</feature>
      <threshold>10.573917</threshold>
      <split pos="left">
         <output>1.1046180725097656</output>
      </split>
      <split pos="right">
         <feature>1</feature>
         <threshold>7.010513</threshold>
         <split pos="left">
            <feature>1</feature>
            <threshold>0.0</threshold>
            <split pos="left">
               <output>-1.1046180725097656</output>
            </split>
            <split pos="right">
               <output>-1.1046180725097656</output>
            </split>
         </split>
         <split pos="right">
            <output>-1.1046180725097656</output>
         </split>
      </split>
   </split>
</tree>
<tree id="10" weight="0.1">
   <split>
      <feature>1</feature>
      <threshold>10.357875</threshold>
```

```
         <split pos="left">
            <feature>1</feature>
            <threshold>7.010513</threshold>
            <split pos="left">
               <feature>1</feature>
               <threshold>0.0</threshold>
               <split pos="left">
                  <output>-1.0838804244995117</output>
               </split>
               <split pos="right">
                  <output>-1.0838804244995117</output>
               </split>
            </split>
            <split pos="right">
               <output>-1.0838804244995117</output>
            </split>
         </split>
         <split pos="right">
            <output>1.0838804244995117</output>
         </split>
      </split>
   </tree>
</ensemble>
"""
    }
  }
}
```

Per visualizzare il modello, inviare la seguente richiesta:

```
GET _ltr/_model/my_ranklib_model
```

Passaggio 8: Ricerca con Learning to Rank

Dopo aver implementato il modello, è possibile eseguire la ricerca.

Eseguire la query `sltr` con le funzionalità di cui si sta utilizzando e il nome del modello da eseguire:

```
POST tmdb/_search
{
  "_source": {
    "includes": ["title", "overview"]
  },
```

```
  "query": {
    "multi_match": {
      "query": "rambo",
      "fields": ["title", "overview"]
    }
  },
  "rescore": {
    "query": {
      "rescore_query": {
        "sltr": {
          "params": {
            "keywords": "rambo"
          },
          "model": "my_ranklib_model"
        }
      }
    }
  }
}
```

Con Learning to Rank, "Rambo" viene visualizzato come primo risultato perché gli è stato assegnato il voto più alto nell'elenco dei giudizi:

```
{
  "took" : 12,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 7,
      "relation" : "eq"
    },
    "max_score" : 13.096414,
    "hits" : [
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "7555",
```

```
        "_score" : 13.096414,
        "_source" : {
          "overview" : "When governments fail to act on behalf of captive missionaries,
ex-Green Beret John James Rambo sets aside his peaceful existence along the Salween
River in a war-torn region of Thailand to take action.  Although he's still haunted
by violent memories of his time as a U.S. soldier during the Vietnam War, Rambo can
hardly turn his back on the aid workers who so desperately need his help.",
          "title" : "Rambo"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "1370",
        "_score" : 11.17245,
        "_source" : {
          "overview" : "Combat has taken its toll on Rambo, but he's finally begun to
find inner peace in a monastery. When Rambo's friend and mentor Col. Trautman asks for
his help on a top secret mission to Afghanistan, Rambo declines but must reconsider
when Trautman is captured.",
          "title" : "Rambo III"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "1368",
        "_score" : 10.442155,
        "_source" : {
          "overview" : "When former Green Beret John Rambo is harassed by local law
enforcement and arrested for vagrancy, the Vietnam vet snaps, runs for the hills and
rat-a-tat-tats his way into the action-movie hall of fame. Hounded by a relentless
sheriff, Rambo employs heavy-handed guerilla tactics to shake the cops off his tail.",
          "title" : "First Blood"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "1369",
        "_score" : 10.442155,
        "_source" : {
          "overview" : "Col. Troutman recruits ex-Green Beret John Rambo for a highly
secret and dangerous mission. Teamed with Co Bao, Rambo goes deep into Vietnam to
```

```
rescue POWs. Deserted by his own team, he's left in a hostile jungle to fight for his
life, avenge the death of a woman and bring corrupt officials to justice.",
          "title" : "Rambo: First Blood Part II"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "31362",
        "_score" : 7.424202,
        "_source" : {
          "overview" : "It is 1985, and a small, tranquil Florida town is being rocked
by a wave of vicious serial murders and bank robberies. Particularly sickening to the
authorities is the gratuitous use of violence by two “Rambo” like killers who dress
themselves in military garb. Based on actual events taken from FBI files, the movie
depicts the Bureau’s efforts to track down these renegades.",
          "title" : "In the Line of Duty: The F.B.I. Murders"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "13258",
        "_score" : 6.43182,
        "_source" : {
          "overview" : """Will Proudfoot (Bill Milner) is looking for an escape from
his family's stifling home life when he encounters Lee Carter (Will Poulter), the
school bully. Armed with a video camera and a copy of "Rambo: First Blood", Lee plans
to make cinematic history by filming his own action-packed video epic. Together, these
two newfound friends-turned-budding-filmmakers quickly discover that their imaginative
– and sometimes mishap-filled – cinematic adventure has begun to take on a life of its
own!""",
          "title" : "Son of Rambow"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "61410",
        "_score" : 3.9719706,
        "_source" : {
          "overview" : "It's South Africa 1990. Two major events are about to happen:
The release of Nelson Mandela and, more importantly, it's Spud Milton's first year
at an elite boys only private boarding school. John Milton is a boy from an ordinary
```

```
 background who wins a scholarship to a private school in Kwazulu-Natal, South Africa.
 Surrounded by boys with nicknames like Gecko, Rambo, Rain Man and Mad Dog, Spud has
 his hands full trying to adapt to his new home. Along the way Spud takes his first
 tentative steps along the path to manhood. (The path it seems could be a rather long
 road). Spud is an only child. He is cursed with parents from well beyond the lunatic
 fringe and a senile granny. His dad is a fervent anti-communist who is paranoid that
 the family domestic worker is running a shebeen from her room at the back of the
 family home. His mom is a free spirit and a teenager's worst nightmare, whether it's
 shopping for Spud's underwear in the local supermarket",
          "title" : "Spud"
        }
      }
    ]
  }
}
```

Se esegui una ricerca senza utilizzare il plug-in Learning to Rank, OpenSearch restituisce risultati diversi:

```
POST tmdb/_search
{
  "_source": {
    "includes": ["title", "overview"]
  },
  "query": {
    "multi_match": {
      "query": "Rambo",
      "fields": ["title", "overview"]
    }
  }
}
```

```
{
  "took" : 5,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
```

```
      "value" : 5,
      "relation" : "eq"
    },
    "max_score" : 11.262714,
    "hits" : [
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "1370",
        "_score" : 11.262714,
        "_source" : {
          "overview" : "Combat has taken its toll on Rambo, but he's finally begun to
find inner peace in a monastery. When Rambo's friend and mentor Col. Trautman asks for
his help on a top secret mission to Afghanistan, Rambo declines but must reconsider
when Trautman is captured.",
          "title" : "Rambo III"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "7555",
        "_score" : 11.2569065,
        "_source" : {
          "overview" : "When governments fail to act on behalf of captive missionaries,
ex-Green Beret John James Rambo sets aside his peaceful existence along the Salween
River in a war-torn region of Thailand to take action.  Although he's still haunted
by violent memories of his time as a U.S. soldier during the Vietnam War, Rambo can
hardly turn his back on the aid workers who so desperately need his help.",
          "title" : "Rambo"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "1368",
        "_score" : 10.558305,
        "_source" : {
          "overview" : "When former Green Beret John Rambo is harassed by local law
enforcement and arrested for vagrancy, the Vietnam vet snaps, runs for the hills and
rat-a-tat-tats his way into the action-movie hall of fame. Hounded by a relentless
sheriff, Rambo employs heavy-handed guerilla tactics to shake the cops off his tail.",
          "title" : "First Blood"
        }
```

```
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "1369",
        "_score" : 10.558305,
        "_source" : {
          "overview" : "Col. Troutman recruits ex-Green Beret John Rambo for a highly
 secret and dangerous mission. Teamed with Co Bao, Rambo goes deep into Vietnam to
 rescue POWs. Deserted by his own team, he's left in a hostile jungle to fight for his
 life, avenge the death of a woman and bring corrupt officials to justice.",
          "title" : "Rambo: First Blood Part II"
        }
      },
      {
        "_index" : "tmdb",
        "_type" : "movie",
        "_id" : "13258",
        "_score" : 6.4600153,
        "_source" : {
          "overview" : """Will Proudfoot (Bill Milner) is looking for an escape from
 his family's stifling home life when he encounters Lee Carter (Will Poulter), the
 school bully. Armed with a video camera and a copy of "Rambo: First Blood", Lee plans
 to make cinematic history by filming his own action-packed video epic. Together, these
 two newfound friends-turned-budding-filmmakers quickly discover that their imaginative
 – and sometimes mishap-filled – cinematic adventure has begun to take on a life of its
 own!""",
          "title" : "Son of Rambow"
        }
      }
    ]
  }
}
```

In base a quanto si pensa che il modello stia funzionando, regolare l'elenco dei giudizi e le funzionalità. Ripetere quindi i passaggi da 2 a 8 per migliorare i risultati della classifica nel tempo.

API Learning to Rank

Utilizzare le operazioni di Learning to Rank per lavorare a livello di programmazione con set di funzionalità e modelli.

Creazione dell'archiviazione

Crea un indice `.ltrstore` nascosto che memorizza informazioni sui metadati, ad esempio set di funzionalità e modelli.

```
PUT _ltr
```

Eliminazione dell'archiviazione

Elimina l'indice `.ltrstore` nascosto e reimposta il plug-in.

```
DELETE _ltr
```

Creazione di un set di funzionalità

Crea un set di funzionalità.

```
POST _ltr/_featureset/<name_of_features>
```

Eliminazione di un set di funzionalità

Elimina un set di funzionalità.

```
DELETE _ltr/_featureset/<name_of_feature_set>
```

Ottenimento di un set di funzionalità

Recupera un set di funzionalità.

```
GET _ltr/_featureset/<name_of_feature_set>
```

Crea modello

Crea un modello.

```
POST _ltr/_featureset/<name_of_feature_set>/_createmodel
```

Eliminazione di un modello

Elimina un modello.

```
DELETE _ltr/_model/<name_of_model>
```

Ottenimento del modello

Recupera un modello.

```
GET _ltr/_model/<name_of_model>
```

Ottenimento delle statistiche

Fornisce informazioni sul comportamento del plug-in.

```
GET _ltr/_stats
```

Puoi anche usare i filtri per recuperare una singola statistica:

```
GET _ltr/_stats/<stat>
```

Inoltre, puoi limitare le informazioni a un singolo nodo del cluster:

```
GET _ltr/_stats/<stat>/nodes/<nodeId>

{
  "_nodes" : {
    "total" : 1,
    "successful" : 1,
    "failed" : 0
  },
  "cluster_name" : "873043598401:ltr-77",
  "stores" : {
    ".ltrstore" : {
      "model_count" : 1,
      "featureset_count" : 1,
      "feature_count" : 2,
      "status" : "green"
    }
  },
  "status" : "green",
  "nodes" : {
    "DjelK-_ZSfyzstO5dhGGQA" : {
```

```
      "cache" : {
        "feature" : {
          "eviction_count" : 0,
          "miss_count" : 0,
          "entry_count" : 0,
          "memory_usage_in_bytes" : 0,
          "hit_count" : 0
        },
        "featureset" : {
          "eviction_count" : 2,
          "miss_count" : 2,
          "entry_count" : 0,
          "memory_usage_in_bytes" : 0,
          "hit_count" : 0
        },
        "model" : {
          "eviction_count" : 2,
          "miss_count" : 3,
          "entry_count" : 1,
          "memory_usage_in_bytes" : 3204,
          "hit_count" : 1
        }
      },
      "request_total_count" : 6,
      "request_error_count" : 0
    }
  }
}
```

Le statistiche sono fornite a due livelli, nodo e cluster, come specificato nelle tabelle seguenti:

Statistiche a livello di nodo

Nome campo	Description
request_total_count	Conteggio totale delle richieste di classific azione.
request_error_count	Conteggio totale delle richieste non riuscite.
cache	Statistiche su tutte le cache (funzionalità, set di funzionalità, modelli). Una occorrenza della cache si verifica quando un utente esegue una

Nome campo	Description
	query sul plug-in e il modello è già caricato in memoria.
cache.eviction_count	Numero di rimozioni della cache.
cache.hit_count	Numero di occorrenze della cache.
cache.miss_count	Numero di mancati riscontri nella cache. Un mancato riscontro nella cache si verifica quando un utente esegue una query sul plug-in e il modello non è ancora caricato in memoria.
cache.entry_count	Numero di voci nella cache.
cache.memory_usage_in_bytes	Memoria totale utilizzata in byte.
cache.cache_capacity_reached	Indica se viene raggiunto il limite della cache.

Statistiche a livello di cluster

Nome campo	Description
archiviazioni	Indica dove vengono archiviati i set di funzionalità e i metadati del modello. (Il valore di default è “.ltrstore”. Altrimenti, viene aggiunto il prefisso “.ltrstore_” a un nome fornito dall'utente).
stores.status	Stato dell'indice.
stores.feature_sets	Numero di set di funzionalità.
stores.features_count	Numero di funzionalità.
stores.model_count	Numero di modelli.
status	Lo stato del plug-in in base allo stato degli indici del feature store (rosso, giallo o verde) e

Nome campo	Description
	allo stato dell'interruttore automatico (aperto o chiuso).
cache.cache_capacity_reached	Indica se viene raggiunto il limite della cache.

Ottenimento delle statistiche della cache

Restituisce le statistiche sull'utilizzo della cache e della memoria.

```
GET _ltr/_cachestats

{
    "_nodes": {
        "total": 2,
        "successful": 2,
        "failed": 0
    },
    "cluster_name": "opensearch-cluster",
    "all": {
        "total": {
            "ram": 612,
            "count": 1
        },
        "features": {
            "ram": 0,
            "count": 0
        },
        "featuresets": {
            "ram": 612,
            "count": 1
        },
        "models": {
            "ram": 0,
            "count": 0
        }
    },
    "stores": {
        ".ltrstore": {
            "total": {
                "ram": 612,
```

```
                "count": 1
            },
            "features": {
                "ram": 0,
                "count": 0
            },
            "featuresets": {
                "ram": 612,
                "count": 1
            },
            "models": {
                "ram": 0,
                "count": 0
            }
        }
    },
    "nodes": {
        "ejF6uutERF20wOFNOXB61A": {
            "name": "opensearch1",
            "hostname": "172.18.0.4",
            "stats": {
                "total": {
                    "ram": 612,
                    "count": 1
                },
                "features": {
                    "ram": 0,
                    "count": 0
                },
                "featuresets": {
                    "ram": 612,
                    "count": 1
                },
                "models": {
                    "ram": 0,
                    "count": 0
                }
            }
        },
        "Z2RZNWRLSveVcz2c6lHf5A": {
            "name": "opensearch2",
            "hostname": "172.18.0.2",
            "stats": {
                ...
```

```
            }
        }
    }
}
```

Cancellazione della cache

Cancella la cache del plug-in. Utilizzare questa opzione per aggiornare il modello.

```
POST _ltr/_clearcache
```

Ricerca asincrona in Amazon Service OpenSearch

Con la ricerca asincrona per Amazon OpenSearch Service puoi inviare una query di ricerca che viene eseguita in background, monitorare l'avanzamento della richiesta e recuperare i risultati in una fase successiva. È possibile recuperare i risultati parziali man mano che diventano disponibili prima del completamento della ricerca. Al termine della ricerca, salvare i risultati per il recupero e l'analisi successivi.

La ricerca asincrona richiede OpenSearch 1.0 o versione successiva oppure Elasticsearch 7.10 o versione successiva.

Questa documentazione fornisce una breve panoramica della ricerca asincrona. Descrive inoltre i limiti dell'uso della ricerca asincrona con un dominio Amazon OpenSearch Service gestito anziché con un cluster open source. OpenSearch Per la documentazione completa sulla ricerca asincrona, comprese le impostazioni disponibili, le autorizzazioni e un riferimento completo all'API, consulta Ricerca asincrona nella documentazione. OpenSearch

Chiamata di ricerca di esempio

Per eseguire una ricerca asincrona, inviare le richieste HTTP a `_plugins/_asynchronous_search` utilizzando il seguente formato:

```
POST opensearch-domain/_plugins/_asynchronous_search
```

> **Note**
>
> Se utilizzi Elasticsearch 7.10 invece di una versione, sostituiscilo con in tutte le richieste di ricerca asincrone OpenSearch . `_plugins` `_opendistro`

È possibile specificare le seguenti opzioni di ricerca asincrona:

Opzioni	Description	Valore predefinito	Richiesto
`wait_for_completion_timeout`	Specifica il tempo che si prevede di attendere per i risultati. È possibile visualizzare tutti i risultati ottenuti in questo tempo, proprio come in una normale ricerca. È possibile eseguire il polling dei risultati rimanenti in base a un ID. Il valore massimo è 300 secondi.	1 secondo	No
`keep_on_completion`	Specifica se si desidera salvare i risultati nel cluster al termine della ricerca. È possibile esaminare i risultati memorizzati in un secondo momento.	`false`	No
`keep_alive`	Specifica il tempo in cui il risultato viene salvato nel cluster. Ad esempio, 2d significa che i risultati vengono memorizzati nel cluster per 48 ore. I risultati della ricerca salvati vengono eliminati dopo questo periodo o se la ricerca viene annullata. Si noti che questo tempo include il runtime della query. Se la query impiega più di questo periodo di tempo, il processo la annulla automaticamente.	12 ore	No

Richiesta di esempio

```
POST _plugins/_asynchronous_search/?
pretty&size=10&wait_for_completion_timeout=1ms&keep_on_completion=true&request_cache=false
{
  "aggs": {
    "city": {
      "terms": {
        "field": "city",
        "size": 10
      }
```

```
    }
  }
}
```

> **Note**
>
> Sono supportati tutti i parametri della richiesta che si applicano a una query `_search` standard. Se utilizzi Elasticsearch 7.10 anziché una versione, sostituiscilo con. OpenSearch `_plugins` `_opendistro`

Autorizzazioni di ricerca asincrona

La ricerca asincrona supporta il controllo granulare degli accessi. Per informazioni dettagliate sull'uso delle autorizzazioni per adattarle al proprio caso d'uso, consultare Sicurezza asincrona della ricerca.

Per i domini con il controllo granulare degli accessi abilitato, sono necessarie le seguenti autorizzazioni minime per un ruolo:

```
# Allows users to use all asynchronous search functionality
asynchronous_search_full_access:
  reserved: true
  cluster_permissions:
    - 'cluster:admin/opensearch/asynchronous-search/*'
  index_permissions:
    - index_patterns:
        - '*'
      allowed_actions:
        - 'indices:data/read/search*'

# Allows users to read stored asynchronous search results
asynchronous_search_read_access:
  reserved: true
  cluster_permissions:
    - 'cluster:admin/opensearch/asynchronous-search/get'
```

Per i domini con il controllo granulare degli accessi disabilitato, utilizzare l'accesso IAM e la chiave segreta per firmare tutte le richieste. È possibile accedere ai risultati con l'ID di ricerca asincrona.

Impostazioni della ricerca asincrona

OpenSearch consente di modificare tutte le impostazioni di ricerca asincrona disponibili utilizzando l'API. `_cluster/settings` In OpenSearch Service, puoi modificare solo le seguenti impostazioni:

- `plugins.asynchronous_search.node_concurrent_running_searches`
- `plugins.asynchronous_search.persist_search_failures`

Funzionalità di ricerca tra cluster

È possibile eseguire una ricerca asincrona tra cluster con le seguenti limitazioni minori:

- È possibile eseguire una ricerca asincrona solo nel dominio di origine.
- Non è possibile ridurre a icona i round trip di rete come parte di una query di ricerca tra cluster.

Se si imposta una connessione tra `domain-a -> domain-b` con l'alias di connessione `cluster_b` e `domain-a -> domain-c` con l'alias di connessione `cluster_c`, ricercare in maniera asincrona `domain-a`, `domain-b` e `domain-c` come segue:

```
POST https://src-domain.us-east-1.es.amazonaws.com/
local_index,cluster_b:b_index,cluster_c:c_index/_plugins/_asynchronous_search/?
pretty&size=10&wait_for_completion_timeout=500ms&keep_on_completion=true&request_cache=false
{
  "size": 0,
  "_source": {
    "excludes": []
  },
  "aggs": {
    "2": {
      "terms": {
        "field": "clientip",
        "size": 50,
        "order": {
          "_count": "desc"
        }
      }
    }
  },
  "stored_fields": [
    "*"
```

```
  ],
  "script_fields": {},
  "docvalue_fields": [
    "@timestamp"
  ],
  "query": {
    "bool": {
      "must": [
        {
          "query_string": {
            "query": "status:404",
            "analyze_wildcard": true,
            "default_field": "*"
          }
        },
        {
          "range": {
            "@timestamp": {
              "gte": 1483747200000,
              "lte": 1488326400000,
              "format": "epoch_millis"
            }
          }
        }
      ],
      "filter": [],
      "should": [],
      "must_not": []
    }
  }
}
```

Risposta

```
{
  "id" :
 "Fm9pYzJyVG91U19xb0hIQUJnMHJfRFEAAAAAAAknghQ1OWVBczNZQjVEa2dMYTBXaTdEagAAAAAAAAAB",
  "state" : "RUNNING",
  "start_time_in_millis" : 1609329314796,
  "expiration_time_in_millis" : 1609761314796
}
```

Per ulteriori informazioni, consulta the section called “Funzionalità di ricerca tra cluster”.

UltraWarm

Le ricerche asincrone con UltraWarm indici continuano a funzionare. Per ulteriori informazioni, consulta the section called “UltraWarm archiviazione”.

> **Note**
>
> È possibile monitorare le statistiche di ricerca asincrone in. CloudWatch Per un elenco completo di parametri, consulta the section called “Parametri di ricerca asincrona”.

Ricerca puntuale in Amazon OpenSearch Service

Point in Time (PIT) è un tipo di ricerca che consente di eseguire diverse query su un set di dati fisso nel tempo. In genere, quando si esegue la stessa query sullo stesso indice in momenti diversi, si ottengono risultati diversi perché i documenti vengono costantemente indicizzati, aggiornati ed eliminati. Con PIT, è possibile eseguire query su uno stato costante del set di dati.

L'uso principale della ricerca PIT è abbinarla alla `search_after` funzionalità. Questo è il metodo di impaginazione preferito OpenSearch, specialmente per l'impaginazione profonda, perché opera su un set di dati congelato nel tempo, non è vincolato a una query e supporta un'impaginazione coerente in avanti e indietro. È possibile utilizzare PIT con un dominio che esegue la versione 2.5. OpenSearch

> **Note**
>
> Questo argomento fornisce una panoramica di PIT e alcuni aspetti da considerare quando lo si utilizza su un dominio Amazon OpenSearch Service gestito anziché su un OpenSearch cluster autogestito. Per la documentazione completa di PIT, incluso un riferimento completo sulle API, consulta Point in Time nella OpenSearch documentazione open source.

Considerazioni

Considerate quanto segue quando configurate le vostre ricerche PIT:

- Se state eseguendo l'aggiornamento da un dominio che esegue la OpenSearch versione 2.3 e avete bisogno di un controllo preciso degli accessi sulle azioni PIT, dovete aggiungere manualmente tali azioni e ruoli.

- Non c'è resilienza per PIT. Il riavvio del nodo, la chiusura del nodo, le implementazioni blu/verdi e il riavvio del OpenSearch processo causano la perdita di tutti i dati PIT.
- Se uno shard si riposiziona durante l' blue/green implementazione, solo i segmenti di dati in tempo reale vengono trasferiti al nuovo nodo. I segmenti di shard detenuti da PIT (sia esclusivamente che quelli condivisi con i dati in tempo reale) rimangono sul vecchio nodo.
- Le ricerche PIT attualmente non funzionano con la ricerca asincrona.

Crea un PIT

Per eseguire una query PIT, inviate le richieste HTTP a `_search/point_in_time` utilizzando il seguente formato:

```
POST opensearch-domain/my-index/_search/point_in_time?keep_alive=time
```

È possibile specificare le seguenti opzioni PIT:

Opzioni	Description	Valore predefinito	Richiesto
`keep_alive`	Il periodo di tempo necessario per conservare il PIT. Ogni volta che si accede a un PIT con una richiesta di ricerca, la durata del PIT viene prolungata del periodo di tempo pari al `keep_alive` parametro. Questo parametro di interrogazione è obbligatorio quando si crea un PIT, ma facoltativo in una richiesta di ricerca.		Sì
`preference`	Una stringa che specifica il nodo o lo shard utilizzato per eseguire la ricerca.	Casuale	No
`routing`	Una stringa che specifica di indirizzare le richieste di ricerca a uno shard specifico.	Il documento `_id`	No
`expand_wildcards`	Una stringa che specifica il tipo di indice che può corrispondere al modello dei caratteri jolly. Supporta valori separati da virgole. I valori validi sono i seguenti.	`open`	No

Opzioni	Description	Valore predefinito	Richiesto
	• `all`: corrisponde a qualsiasi indice o flusso di dati, compresi quelli nascosti. • `open`: abbina indici aperti e non nascosti o flussi di dati non nascosti. • `closed`: abbina indici chiusi e non nascosti o flussi di dati non nascosti. • `hidden`: abbina indici o flussi di dati nascosti. Deve essere combinato con aperto, chiuso o sia aperto che chiuso. • `none`: Non sono accettati modelli jolly.		
`allow_partial_pit_creation`	Un valore booleano che specifica se creare un PIT con errori parziali.	`true`	No

Risposta di esempio

```
{
    "pit_id":
 "o463QQEPbXktaW5kZXgtMDAwMDAxFnNOWU43ckt3U3IyaFVpbGE1UWEtMncAFjFyeXBsRGJmVFM2RTB6eVg1aVVqQncAA
    "_shards": {
        "total": 1,
        "successful": 1,
        "skipped": 0,
        "failed": 0
    },
    "creation_time": 1658146050064
}
```

Quando si crea un PIT, si riceve un ID PIT nella risposta. Questo è l'ID che utilizzate per eseguire ricerche con il PIT.

Autorizzazioni temporanee

PIT supporta un controllo granulare degli accessi. Se stai eseguendo l'aggiornamento a un dominio OpenSearch versione 2.5 e hai bisogno di un controllo degli accessi dettagliato, devi creare manualmente ruoli con le seguenti autorizzazioni:

```
# Allows users to use all point in time search search functionality
point_in_time_full_access:
  reserved: true
  index_permissions:
    - index_patterns:
        - '*'
      allowed_actions:
        - "indices:data/read/point_in_time/create"
        - "indices:data/read/point_in_time/delete"
        - "indices:data/read/point_in_time/readall"
        - "indices:data/read/search"
        - "indices:monitor/point_in_time/segments"


# Allows users to use point in time search search functionality for specific index
# All type operations like list all PITs, delete all PITs are not supported in this
 case

point_in_time_index_access:
  reserved: true
  index_permissions:
    - index_patterns:
        - 'my-index-1'
      allowed_actions:
        - "indices:data/read/point_in_time/create"
        - "indices:data/read/point_in_time/delete"
        - "indices:data/read/search"
        - "indices:monitor/point_in_time/segments"
```

Per i domini con OpenSearch versione 2.5 e successive, puoi utilizzare il ruolo integrato. `point_in_time_full_access` Per ulteriori informazioni, consulta il modello di sicurezza nella OpenSearch documentazione.

Impostazioni PIT

OpenSearch consente di modificare tutte le impostazioni PIT disponibili utilizzando l'`_cluster/settings`API. In OpenSearch Service, al momento non è possibile modificare le impostazioni.

Funzionalità di ricerca tra cluster

È possibile creare PITs, cercare con PIT IDs PITs, elencare ed eliminare PITs tra cluster con le seguenti limitazioni minori:

- È possibile elencare tutto ed eliminare tutto PITs solo nel dominio di origine.
- Non è possibile ridurre a icona i round trip di rete come parte di una query di ricerca tra cluster.

Per ulteriori informazioni, consulta the section called “Funzionalità di ricerca tra cluster”.

UltraWarm

Le ricerche PIT con UltraWarm indici continuano a funzionare. Per ulteriori informazioni, consulta the section called “UltraWarm archiviazione”.

> **Note**
>
> È possibile monitorare le statistiche di ricerca PIT in. CloudWatch Per un elenco completo di parametri, consulta the section called “Metriche puntuali”.

Ricerca agente in Amazon Service OpenSearch

A partire dalla OpenSearch versione 3.3, la ricerca agentica abilita un processo basato sull'intelligenza artificiale che utilizza un agente autonomo per eseguire ricerche complesse per conto dell'utente.

Agentic search introduce un sistema di agenti intelligenti che comprende le intenzioni dell'utente, orchestra gli strumenti giusti e genera query ottimizzate. Fornisce riepiloghi trasparenti delle sue decisioni attraverso un'interfaccia in linguaggio naturale.

Con OpenSearch Service, puoi configurare connettori AI per Servizi AWS servizi esterni. Utilizzando la console, puoi anche creare un modello ML con un CloudFormation modello che può essere

utilizzato per creare il tuo agente. Per ulteriori informazioni, vedere Configurazione di Agentic Search con Bedrock Claude.

Per la documentazione e l' step-by-stepimplementazione complete, consulta Agentic Search nella documentazione. OpenSearch

Ricerca semantica in Amazon Service OpenSearch

A partire dalla versione 2.19, puoi utilizzare l'arricchimento semantico automatico per implementare la ricerca semantica con uno sforzo di configurazione minimo. Questa funzionalità genera automaticamente una codifica sparsa per i campi di testo, eliminando la necessità di configurare manualmente la pipeline di inserimento. Per i dettagli, consulta Automatic Semantic Enrichment nella documentazione. OpenSearch

A partire dalla OpenSearch versione 2.9, è possibile utilizzare la ricerca semantica per comprendere le query di ricerca e migliorare la pertinenza della ricerca. Puoi utilizzare la ricerca semantica in due modi: con la ricerca neurale e con la ricerca K-Nearest Neighbor (k-NN).

Con OpenSearch Service, puoi configurare connettori AI e servizi esterni. Servizi AWS Utilizzando la console, puoi anche creare un modello ML con un CloudFormation modello. Per ulteriori informazioni, consulta the section called “CloudFormation integrazioni di modelli”.

Per la documentazione completa sulla ricerca semantica, inclusa una step-by-step guida all'uso della ricerca semantica, vedi Ricerca semantica nella documentazione open source. OpenSearch

Ricerca simultanea di segmenti in Amazon Service OpenSearch

A partire dalla OpenSearch versione 2.17, la ricerca simultanea per segmenti utilizza una nuova impostazione per controllare il comportamento di ricerca simultanea.

- I nuovi domini creati con la versione 2.17 hanno la ricerca di segmenti simultanea predefinita impostata sulla modalità auto per impostazione predefinita sui nodi 2xl o superiori.
- Nei domini esistenti che eseguono l'aggiornamento alla versione 2.17 la ricerca di segmenti simultanea predefinita è impostata su auto in base al tipo di istanza per tutti i nodi di dimensioni pari o superiori a 2xl e se l'utilizzo complessivo della CPU del cluster è inferiore al 45% nell'ultima settimana.
- Per ulteriori informazioni, consulta Concurrent Segment Search versione 2.17.

A partire dalla OpenSearch versione 2.13, puoi utilizzare la ricerca simultanea per segmenti per aiutarti a cercare segmenti in parallelo durante la fase di interrogazione. Per la documentazione completa sulla ricerca simultanea di segmenti, consulta Ricerca simultanea di segmenti nella documentazione open source. OpenSearch Per informazioni sui CloudWatch parametri di Amazon relativi alla ricerca simultanea di segmenti, consulta Metriche e metriche delle istanze. UltraWarm

Esistono alcune limitazioni aggiuntive che si applicano quando utilizzi la ricerca di segmenti corrente con Amazon OpenSearch Service:

- Per impostazione predefinita, OpenSearch Service utilizza un conteggio di 2 sezioni con il meccanismo del numero massimo di sezioni.

Generazione di query in linguaggio naturale in Amazon OpenSearch Service

La funzionalità di generazione di query in linguaggio naturale di Amazon OpenSearch Service ti consente di interrogare i dati dei log operativi e di sicurezza tramite linguaggio naturale. OpenSearch è un'opzione ideale per esplorare i dati di log perché è un motore di analisi e ricerca dei log altamente scalabile e performante, e ora puoi usare il linguaggio naturale per esplorare questi log. Questa funzionalità consente di identificare i problemi senza fare affidamento sul OpenSearch Piped Processing Language (PPL) o dover cercare le definizioni dei dati durante la creazione delle query. È possibile utilizzare la funzionalità di generazione di query in linguaggio naturale sui domini di OpenSearch servizio con la versione 2.13 e successive. È necessario che il controllo granulare degli accessi sia abilitato.

Questa funzionalità è stata creata con l'Assistant ToolkitOpenSearch . Se desideri creare funzionalità simili che si colleghino ai tuoi modelli linguistici di grandi dimensioni, puoi utilizzare il toolkit per configurare agenti e strumenti personalizzati.

Prerequisiti

Prima di poter utilizzare la funzionalità di generazione di query in linguaggio naturale, il dominio deve avere quanto segue:

- Versione 2.13 o successiva.
- Software di assistenza R20240520-P4 o versione successiva.

- Controllo granulare degli accessi abilitato. Per ulteriori informazioni, consulta the section called “Abilitazione del controllo granulare degli accessi”.

Nozioni di base

La generazione di query in linguaggio naturale è abilitata per impostazione predefinita su tutti i domini creati con la versione 2.13 o successiva che hanno abilitato il controllo granulare degli accessi.

Per gli altri domini, abilitalo selezionando le funzionalità Enable Natural Language Query Generation e Amazon Q Developer.

Dopo averlo abilitato, vai alla pagina Logs in OpenSearch Dashboards. Scegli Event Explorer e fai una domanda con l'assistente alle query.

Configurazione delle autorizzazioni

Se abiliti la generazione di query in linguaggio naturale su un dominio di OpenSearch servizio preesistente, il ruolo query_assistant_access potrebbe non essere definito nel dominio. Gli utenti senza privilegi di amministratore devono essere mappati a questo ruolo in modo da gestire gli indici a caldo sui domini che utilizzano il controllo granulare degli accessi. Per creare manualmente il ruolo query_assistant_access, effettuate le seguenti operazioni:

1. In OpenSearch Dashboards, vai a Sicurezza e scegli Ruoli.
2. Scegli Crea ruolo e configura le seguenti autorizzazioni del cluster:
   - `cluster:admin/opensearch/ml/config/get`
   - `cluster:admin/opensearch/ml/execute`
   - `cluster:admin/opensearch/ml/predict`
   - `cluster:admin/opensearch/ppl`
3. Assegna un nome al ruolo query_assistant_access.
4. Scegli Crea ruolo. Il ruolo query_assistant_access è ora disponibile.

> **Note**
>
> È inoltre necessario disporre delle autorizzazioni `indices:admin/mappings/get` e `read` indicizzazione per gli indici con cui si desidera utilizzare le domande in linguaggio naturale.

Automazione della configurazione

Flow Framework è un OpenSearch plug-in che fornisce un modo per automatizzare OpenSearch le configurazioni per casi d'uso come la generazione di query e la chat conversazionale. Poiché il plug-in tiene traccia delle risorse che abilitano la funzionalità di generazione di query in linguaggio naturale, l'indice del framework flow memorizza un modello per ogni dominio che utilizza l'assistenza alle query.

Flow Framework consente di scegliere tra un set di modelli predefiniti o di creare automazioni personalizzate per connettori, strumenti, agenti e altri componenti di machine learning da preparare OpenSearch come backend per modelli generativi.

Ricerca vettoriale

La ricerca vettoriale in Amazon OpenSearch Service ti consente di cercare contenuti semanticamente simili utilizzando incorporamenti di machine learning anziché la tradizionale corrispondenza di parole chiave. La ricerca vettoriale converte i dati (testo, immagini, audio, ecc.) in vettori numerici ad alta dimensione (incorporamenti) che catturano il significato semantico del contenuto. Quando esegui una ricerca, OpenSearch confronta la rappresentazione vettoriale della tua query con i vettori memorizzati per trovare gli elementi più simili.

La ricerca vettoriale include i seguenti componenti chiave.

Campi vettoriali

OpenSearch supporta il tipo di `knn_vector` campo per memorizzare vettori densi con dimensioni configurabili (fino a 16.000).

Metodi di ricerca

- k-NN (k-nearest neighbors): trova i k vettori più simili
- K-nn approssimativo: utilizza algoritmi come HNSW (Hierarchical Navigable Small World) per ricerche più rapide su set di dati di grandi dimensioni

Metriche della distanza

Supporta vari calcoli di somiglianza, tra cui:

- Distanza Euclidea
- Similitudine coseno
- Prodotto Dot

Casi di utilizzo comune

La ricerca vettoriale supporta i seguenti casi d'uso comuni.

- Ricerca semantica: trova documenti con un significato simile, non solo con parole chiave corrispondenti
- Sistemi di raccomandazione: suggerisci prodotti, contenuti o utenti simili
- Ricerca immagini: trova immagini visivamente simili
- Rilevamento delle anomalie: identifica i valori anomali nei modelli di dati
- RAG (Retrieval Augmented Generation): migliora le risposte LLM con un contesto pertinente

Integrazione con l'apprendimento automatico

OpenSearch si integra con i seguenti servizi e modelli di machine learning:

- Amazon Bedrock: per generare incorporamenti utilizzando modelli di base
- Amazon SageMaker AI: per l'implementazione di modelli ML personalizzati
- Modelli Hugging Face: modelli di incorporamento pre-addestrati
- Modelli personalizzati: modelli di incorporamento personalizzati

Con la ricerca vettoriale, puoi creare sofisticate applicazioni basate sull'intelligenza artificiale in grado di comprendere il contesto e il significato, andando ben oltre le tradizionali funzionalità di corrispondenza del testo.

Importazione da Amazon S3 Vectors a Serverless OpenSearch

Amazon S3 Vectors offre il primo archivio di oggetti nel cloud con supporto nativo per archiviare e interrogare vettori. S3 Vectors offre uno storage vettoriale economico, elastico e durevole che può essere interrogato in base al significato e alla somiglianza semantici. Offre tempi di risposta alle query inferiori al secondo e costi fino al 90% inferiori per il caricamento, l'archiviazione e l'interrogazione dei vettori.

Amazon S3 Vectors introduce i bucket vettoriali S3, che puoi utilizzare per archiviare, accedere e interrogare dati vettoriali senza dover fornire alcuna infrastruttura. In un bucket vettoriale, puoi organizzare i dati vettoriali all'interno di indici vettoriali. Il tuo bucket vettoriale può avere più indici vettoriali e ogni indice vettoriale può contenere milioni di vettori. Per ulteriori informazioni, consulta Working with Amazon S3 Vectors and Vector bucket nella Amazon S3 User Guide.

Ogni vettore è composto da:

- Una chiave unica
- Dati vettoriali
- Metadati opzionali in formato JSON

Gli indici vettoriali supportano le funzioni di distanza euclidea e cosena per le operazioni di ricerca di similarità.

> **Note**
>
> Il vantaggio principale dei bucket vettoriali è la loro capacità di archiviare enormi set di dati a costi estremamente bassi, fornendo al contempo l'accesso diretto alle API per le operazioni vettoriali.

Per ulteriori informazioni sui bucket vettoriali Amazon S3, incluso come crearne uno, consulta Working with Amazon S3 Vectors and Vector bucket nella Amazon S3 User Guide. Per ulteriori informazioni sull'integrazione con OpenSearch Service oltre a quelle descritte in questo argomento, consulta Using S3 Vectors with Service OpenSearch

Puoi usare S3 Vectors con Amazon OpenSearch Service per ridurre il costo dello storage vettoriale quando le query sono meno frequenti e quindi spostare rapidamente quei set di dati all'aumentare delle richieste o per migliorare OpenSearch le capacità di ricerca.

OpenSearch Il servizio si integra con Amazon S3 Vectors per fornire prestazioni e funzionalità migliorate oltre a quelle offerte dai bucket vettoriali Amazon S3 da soli. Prendi in considerazione questa integrazione quando hai bisogno di:

- Maggiore velocità di trasmissione delle query
- Latenza di ricerca inferiore al secondo
- Funzionalità di analisi avanzate come le aggregazioni
- Ricerca ibrida che combina testo e dati vettoriali

Questa integrazione è particolarmente utile quando più applicazioni utilizzano gli stessi dati vettoriali con requisiti prestazionali diversi. Alcune applicazioni possono interagire direttamente con i bucket vettoriali Amazon S3 per casi d'uso sensibili ai costi, mentre altre sfruttano l'integrazione per operazioni critiche in termini di prestazioni. OpenSearch

Architettura di integrazione

L'integrazione utilizza Amazon OpenSearch Ingestion (OSI) come pipeline di dati tra gli indici vettoriali Amazon S3 e le raccolte vettoriali Amazon Serverless. OpenSearch OpenSearch Ingestion esporta automaticamente i dati vettoriali dall'indice vettoriale specificato e li inserisce in raccolte vettoriali Serverless per operazioni di ricerca ad alte prestazioni. OpenSearch

> **Note**
>
> Dopo l'esportazione, i dati sono ancora presenti nell'indice vettoriale S3. Hai due copie dei dati.

Ogni indice vettoriale viene mappato a un indice corrispondente nella raccolta OpenSearch Serverless. L'integrazione:

- Conserva le dimensioni vettoriali
- Conserva i metadati
- Ottimizza la struttura dei dati per le funzionalità OpenSearch di ricerca vettoriale di

Dopo la configurazione, OpenSearch Ingestion avvia il processo di esportazione dei dati consumando i vettori dall'indice vettoriale specificato utilizzando l'API Amazon S3. ListVectors Il servizio elabora i vettori in parallelo per ottimizzare la velocità di inserimento rispettando i limiti di scalabilità di OpenSearch Ingestion e Amazon Serverless. OpenSearch

Durante l'ingestione, il servizio:

- Trasforma i dati vettoriali in modo che corrispondano al formato previsto per il servizio OpenSearch
- Conserva le informazioni essenziali tra cui valori vettoriali, metadati e metriche di distanza
- Gestisce gli scenari di errore tramite meccanismi di ripetizione intelligenti
- Inserisce i record problematici in un bucket Amazon S3 utilizzato come coda di lettere morte per analisi successive

L'integrazione gestisce enormi set di dati in modo efficiente, con prestazioni che dipendono dalle dimensioni del vettore, dalle dimensioni del set di dati e dai limiti di scalabilità configurati. OSI può scalare fino a 16 lavoratori per pipeline, mentre OpenSearch Serverless regola automaticamente la capacità in base alle richieste di ingestione. Per impostazione predefinita, OpenSearch aumenta a 100 l'unità `maxSearch` OpenSearch computazionale (OCU) sul lato serverless. OpenSearch

> **Note**
>
> L'integrazione dà priorità all'efficienza dei costi attraverso:
>
> - Arresto automatico della pipeline dopo il completamento dell'esportazione

- OpenSearch Scalabilità della raccolta senza server
- Pay-per-use modello di risorse

Autorizzazioni IAM richieste

L'integrazione richiede un'attenta configurazione delle autorizzazioni IAM per consentire una comunicazione sicura tra i servizi. OpenSearch Ingestion richiede le autorizzazioni per leggere dagli indici vettoriali di Amazon S3, scrivere nelle raccolte vettoriali di OpenSearch Service e gestire le politiche di sicurezza associate.

Quando abiliti l'integrazione utilizzando la procedura riportata più avanti in questo argomento, puoi scegliere una delle seguenti opzioni per la gestione delle autorizzazioni:

- Consenti al sistema di creare automaticamente un ruolo di servizio con le autorizzazioni necessarie
- Fornisci un ruolo esistente che soddisfi i requisiti

Il ruolo creato automaticamente include politiche per:

- Accesso all'indice vettoriale Amazon S3 APIs
- Gestione delle operazioni di raccolta dei OpenSearch servizi
- Gestione delle operazioni di coda delle lettere morte in caso di tentativi di inserimento non riusciti

Se scegli di specificare un ruolo esistente, verifica che il ruolo disponga delle seguenti autorizzazioni IAM:

(Obbligatorio): autorizzazioni per la pipeline di dati tra OpenSearch Ingestion e Serverless OpenSearch

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "allowAPIs",
            "Effect": "Allow",
```

```
            "Action": [ "aoss:APIAccessAll", "aoss:BatchGetCollection" ],
            "Resource": [ "arn:aws:aoss:*:111122223333:collection/collection-
id" ]
        },
        {
            "Sid": "allowSecurityPolicy",
            "Effect": "Allow",
            "Action": [
                "aoss:CreateSecurityPolicy",
                "aoss:UpdateSecurityPolicy",
                "aoss:GetSecurityPolicy"
            ],
            "Resource": "*",
            "Condition":{
               "StringLike":{
                  "aoss:collection": [ "collection-name" ]
                },
               "StringEquals": {
                  "aws:ResourceAccount": [ "111122223333" ]
               }
            }
        }
    ]
}
```

(Obbligatorio): autorizzazioni per l'inserimento dei dati tra OpenSearch Ingestion e la coda di lettere morte di Amazon S3

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "s3Access",
            "Effect": "Allow",
            "Action": [
              "s3:PutObject"
            ],
            "Resource": [ "arn:aws:s3:::bucket/*" ]
        }
```

```
    ]
}
```

(Obbligatorio): autorizzazioni di inserimento dati tra OpenSearch Ingestion e Amazon S3 Vectors

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "AllowS3VectorIndexAccess",
            "Effect": "Allow",
            "Action": [
               "s3vectors:ListVectors",
               "s3vectors:GetVectors"
            ],
            "Resource": [
                "arn:aws:s3vectors:us-east-1:111122223333:bucket/bucket-name/
index/index-name"
            ]
        }
    ]
}
```

(Richiesto se AWS KMS la crittografia è abilitata): autorizzazioni di decrittografia per la comunicazione tra OpenSearch Ingestion e Amazon S3 Vectors

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "allowS3VectorDecryptionOfCustomManagedKey",
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
```

```
            "Resource": [ "arn:aws:kms:us-east-1:111122223333:key/key-id" ],
            "Condition": {
                "StringEquals": {
                    "kms:ViaService": "s3vectors.us-east-1.amazonaws.com",
                    "kms:EncryptionContext:aws:s3vectors:arn": [
                        "arn:aws:s3vectors:us-east-1:111122223333:bucket/example-
bucket",
                        "arn:aws:s3vectors:us-east-1:111122223333:bucket/example-
bucket/index/example-index"
                        ]
                 }
             }
         }
    ]
}
```

Configurazione dell'integrazione di Amazon S3 Vectors con OpenSearch

Utilizza la seguente procedura per configurare l'integrazione di Amazon S3 Vectors con Serverless. OpenSearch

> **Note**
>
> Se hai avviato il processo di configurazione dell'integrazione dalla console Amazon S3 scegliendo l'opzione Esporta OpenSearch in nella pagina Vector bucket, alcuni dei passaggi della procedura seguente non sono applicabili, come indicato nella procedura.

Per configurare l'integrazione di Amazon S3 Vectors con Serverless OpenSearch

1. Apri la pagina Importa l'indice vettoriale S3 nel motore OpenSearch vettoriale nella console di Amazon OpenSearch Service. La pagina viene visualizzata automaticamente se hai fatto clic su Esporta OpenSearch in nella console Amazon S3. Se inizi dalla OpenSearch console, scegli Integrazione nella barra di navigazione a sinistra, quindi scegli Importa indice vettoriale S3.
2. Nella sezione Source, se hai iniziato dalla console Amazon S3, verifica che il nome dell'indice vettoriale e il relativo Amazon Resource Name (ARN) siano già specificati. Se hai iniziato dalla OpenSearch console, inserisci l'indice ARN nel campo ARN dell'indice vettoriale S3.

3. Nella sezione Accesso al servizio, scegli un'opzione. Se scegli un ruolo esistente, verifica che disponga di tutte le autorizzazioni necessarie per l'integrazione, come descritto inAutorizzazioni IAM richieste.
4. (Facoltativo) Espandere Additional settings (Impostazioni aggiuntive). Per Abilita la ridondanza (repliche attive) consigliamo di lasciare selezionata questa opzione per gli ambienti di produzione. Quando create la prima raccolta, OpenSearch Serverless ne crea due istanze, una per l'indicizzazione e OCUs una per la ricerca. Inoltre, avvia anche un set di nodi in standby in un'altra Zona di disponibilità per garantire un'elevata disponibilità. Per scopi di sviluppo e test, è possibile disabilitare l'impostazione Enable redondancy per una raccolta, che elimina le due repliche in standby e ne crea solo due. OCUs Per impostazione predefinita, le repliche attive ridondanti sono abilitate, il che significa che ne vengono create un totale di quattro OCUs per la prima raccolta in un account.

   Per Aggiungi AWS KMS chiave gestita dal cliente per Amazon OpenSearch Serverless vector, scegli questa opzione per crittografare i dati nella raccolta vettoriale utilizzando una chiave gestita dal cliente. Per impostazione predefinita, utilizza un. OpenSearch Chiave gestita da AWS
5. Se hai avviato questo processo facendo clic sull' OpenSearchopzione Esporta in nella console Amazon S3, la sezione Esporta dettagli elenca i OpenSearch passaggi successivi. Quando sei pronto, scegli Esporta.

   Se hai avviato questo processo nella console OpenSearch di servizio, nella sezione Dettagli di importazione sono elencati i OpenSearch passaggi successivi. Quando sei pronto, scegli Importa.

   OpenSearch apre la pagina della cronologia per visualizzare tutti gli indici vettoriali exports/imports di Amazon S3 in indici Serverless. OpenSearch

Una volta completata con successo l'inserimento, OSI interrompe automaticamente la pipeline per evitare costi inutili e mantenere i dati esportati all'interno. OpenSearch È possibile monitorare l'avanzamento dell'integrazione tramite CloudWatch metriche e accedere a log dettagliati per la risoluzione dei problemi.

La OpenSearch raccolta rimane attiva e disponibile per le interrogazioni dopo il completamento dell'inserimento iniziale. È possibile eseguire:

- Ricerche per somiglianza
- Aggregazioni

- Operazioni di analisi

Funzionalità di ricerca avanzate con un motore vettoriale Amazon S3

Amazon OpenSearch Service offre la possibilità di utilizzare Amazon S3 come motore vettoriale per indici vettoriali. Questa funzionalità consente di scaricare dati vettoriali su Amazon S3 mantenendo funzionalità di ricerca vettoriale in meno di un secondo a costi contenuti.

Con questa funzionalità, OpenSearch archivia gli incorporamenti vettoriali in un indice vettoriale Amazon S3 mantenendo altri campi del documento nello OpenSearch storage del cluster. Questa architettura offre i seguenti vantaggi:

- Durabilità: i dati scritti su S3 Vectors vengono archiviati su S3, progettato per 11 secondi di durata dei dati.
- Scalabilità: trasferisci set di dati vettoriali di grandi dimensioni su S3 senza consumare lo storage del cluster.
- Convenienza: ottimizza i costi di archiviazione per carichi di lavoro con molti vettori.

OpenSearch presenta i seguenti requisiti per l'utilizzo degli indici vettoriali S3:

- OpenSearch versione 2.19 o successiva
- OpenSearch Istanze ottimizzate
- Versione della patch più recente per la tua OpenSearch versione

Abilitazione di S3 Vectors

Quando crei un nuovo dominio o aggiorni un dominio esistente, puoi scegliere l'opzione Abilita i vettori S3 come motore nella sezione Funzioni avanzate. Questa impostazione consente di OpenSearch creare un bucket vettoriale S3 quando utilizzi S3 Vectors come motore. Quando abiliti questa opzione, OpenSearch configura S3 Vectors per il tuo dominio tramite:

1. Creando due nuove concessioni sulla AWS KMS chiave configurata con il tuo dominio:
   - Una sovvenzione per i lavori di indicizzazione in background di S3 Vectors con privilegi di decrittografia

- Una sovvenzione per creare bucket vettoriali S3 con autorizzazioni OpenSearch `GenerateDataKey`

2. Configurazione della chiave KMS utilizzata dal OpenSearch dominio come CMK per la crittografia del resto di tutti i dati dell'indice vettoriale.

Creazione di indici con il motore vettoriale S3

Dopo aver configurato un dominio, puoi creare uno o più indici k-NN con campi da utilizzare `s3vector` come motore vettoriale di backend nelle mappature degli indici. Puoi configurare diversi campi vettoriali con diversi tipi di motore in base al tuo caso d'uso.

> **Important**
>
> È possibile utilizzare il `s3vector` motore solo per mappare una definizione di campo durante la creazione dell'indice. Non è possibile aggiungere o aggiornare la mappatura con il `s3vector` motore dopo la creazione dell'indice.

Ecco alcuni esempi che creano indici di motori vettoriali S3.

Esempio: creazione di un indice k-NN con il motore vettoriale S3

```
PUT my-first-s3vector-index
{
  "settings": {
    "index": {
      "knn": true
    }
  },
  "mappings": {
    "properties": {
        "my_vector_1": {
          "type": "knn_vector",
          "dimension": 2,
          "space_type": "l2",
          "method": {
            "engine": "s3vector"
          }
        },
        "price": {
```

```
          "type": "float"
        }
    }
  }
}
```

Esempio: creazione di un indice k-NN con motori vettoriali S3 e FAISS

Questo esempio evidenzia il fatto che è possibile utilizzare più motori vettoriali all'interno dello stesso indice.

```
PUT my-vector-index
{
  "settings": {
    "index": {
      "knn": true
    }
  },
  "mappings": {
    "properties": {
        "my_vector_1": {
          "type": "knn_vector",
          "dimension": 2,
          "space_type": "l2",
          "method": {
            "engine": "s3vector"
          }
        },
        "price": {
          "type": "float"
        },
        "my_vector_2": {
            "type": "knn_vector",
            "dimension": 2,
            "space_type": "cosine",
            "method": {
                "name": "hnsw",
                "engine": "faiss",
                "parameters": {
                    "ef_construction": 128,
                    "m": 24
                }
            }
```

```
        }
    }
  }
}
```

Esempio non supportato: aggiunta del motore vettoriale S3 dopo la creazione dell'indice

Il seguente approccio non è supportato e avrà esito negativo.

```
PUT my-first-s3vector-index
{
  "settings": {
    "index": {
      "knn": true
    }
  }
}

PUT my-first-s3vector-index/_mapping
{
  "properties": {
        "my_vector_1": {
          "type": "knn_vector",
          "dimension": 2,
          "space_type": "l2",
          "method": {
            "engine": "s3vector"
          }
        },
        "price": {
          "type": "float"
        }
    }
}
```

Limitazioni funzionali

Considerate le seguenti limitazioni prima di utilizzare `s3vector` il motore in un indice:

Caratteristiche e comportamenti non supportati dal motore s3vector

Funzionalità	Comportamento
indice Split/Shrink/Clone	Questi APIs falliscono se utilizzati con un indice configurato con `s3vector` il motore in campo. `knn_vector`
Snapshot	Gli indici che utilizzano il `s3vector` motore non supportano le istantanee. Per i domini gestiti: • Le istantanee automatizzate includono solo gli indici che non utilizzano il motore. `s3vector` • Le richieste manuali di istantanee per `s3vector` gli indici hanno esito negativo. **Note** Sebbene le istantanee non siano supportate per il point-in-time ripristino, `s3vector` Engine, insieme alle istanze OpenSearch Optimized, offre 11 livelli di durabilità.
UltraWarm livello	Gli indici configurati con il `s3vector` motore non possono migrare al UltraWarm livello.
Replica tra cluster	Gli indici configurati con il `s3vector` motore non supportano la replica tra cluster.
Protezione da eliminazione accidentale	Poiché le istantanee non sono supportate per gli indici che utilizzano il `s3vector` motore, la protezione da eliminazione accidentale non è disponibile. Puoi comunque ripristinare altri indici nel dominio.
Ricerca radiale	Le query con ricerca radiale non sono supportate nei campi che utilizzano il motore. `s3vector`

Indicizzazione dei documenti

Dopo aver creato un indice con il motore vettoriale S3, puoi importare documenti utilizzando l'API standard. `_bulk` OpenSearch scarica automaticamente i dati vettoriali dei `knn_vector` campi che utilizzano il `s3vector` motore all'indice vettoriale S3 in tempo reale. I dati che appartengono ad altri campi o ai `knn_vector` campi che utilizzano motori diversi verranno conservati OpenSearch nel relativo livello di archiviazione.

Per tutte le richieste di massa che vengono riconosciute, OpenSearch garantisce che tutti i dati (vettoriali e non vettoriali) siano durevoli. Se una richiesta riceve un riscontro negativo, non vi sono garanzie sulla durabilità dei documenti contenuti nella richiesta di massa. È consigliabile riprovare tali richieste preferibilmente dopo aver eliminato la precedente richiesta non riuscita utilizzando l'ID del documento per evitare documenti duplicati in questi rari casi.

Esempio di indicizzazione in blocco

```
POST _bulk
{ "index": { "_index": "my-first-s3vector-index", "_id": "1" } }
{ "my_vector_1": [1.5, 2.5], "price": 12.2 }
{ "index": { "_index": "my-first-s3vector-index", "_id": "2" } }
{ "my_vector_1": [2.5, 3.5], "price": 7.1 }
{ "index": { "_index": "my-first-s3vector-index", "_id": "3" } }
{ "my_vector_1": [3.5, 4.5], "price": 12.9 }
{ "index": { "_index": "my-first-s3vector-index", "_id": "4" } }
{ "my_vector_1": [5.5, 6.5], "price": 1.2 }
{ "index": { "_index": "my-first-s3vector-index", "_id": "5" } }
{ "my_vector_1": [4.5, 5.5], "price": 3.7 }
```

Ricerca di documenti

Puoi cercare nell'indice utilizzando l'`_search`API standard per eseguire query di testo, k-NN o ibride. Per le query sui `knn_vector` campi configurati con il `s3vector` motore, scarica OpenSearch automaticamente la query nell'indice vettoriale S3 corrispondente.

> **Note**
>
> Con `s3vector` engine, le query di ricerca k-NN supportano un valore massimo di 100. k Ciò significa che è possibile restituire un massimo di 100 vicini più vicini nei risultati della ricerca.

Esempio di query di ricerca

```
GET my-first-s3vector-index/_search
{
  "size": 2,
  "query": {
    "knn": {
      "my_vector_1": {
        "vector": [2.5, 3.5],
        "k": 2
      }
    }
  }
}
```

È possibile eseguire una ricerca vettoriale filtrata sull'indice OpenSearch kNN utilizzando il motore s3vector. OpenSearch applica il filtro come filtro successivo e utilizza un meccanismo di sovracampionamento basato su determinate euristiche per bilanciare il richiamo rispetto alla latenza.

Esempio di query di ricerca con filtro:

```
GET my-index/_search
{
  "size": 10,
  "query": {
    "knn": {
      "my_vector_field": {
        "vector": [2.5, 3.5, 1.2, 4.8],
        "k": 10,
        "filter": {
          "range": {
            "price": {
              "gte": 10,
              "lte": 100
            }
          }
        }
      }
    }
  }
}
```

Parametri di mappatura supportati

Con `s3vector` engine, il `knn_vector` campo supporta i seguenti parametri nelle mappature.

Parametri del campo vettoriale

Parametro	Obbligatorio	Description	Valori supportati
`type`	Sì	Il tipo di campo presente nel documento.	`knn_vector`
`dimension`	Sì	La dimensione di ogni vettore che verrà inserito nell'indice.	>0, <=4096
`space_type`	No	Lo spazio vettoriale usato per calcolare la distanza tra i vettori.	`l2`, `cosinesimil`
`method.engine`	Sì	Il motore k-NN approssimativo da usare per l'indicizzazione e la ricerca.	`s3vector`
`method.name`	No	Il metodo del vicino più vicino	""
`store`	N/D	L'abilitazione o la disabilitazione di questo parametro di mappatura è impossibile poiché i dati knn_vector non sono memorizzati in. OpenSearch	Non supportato
`doc_values`	N/D	L'attivazione o la disabilitazione di questo parametro di mappatura non è un'operazione poiché i dati knn_vector non sono archiviati in. OpenSearch	Non supportato

> **Important**
>
> I tipi di campi annidati non sono supportati utilizzando il motore `knn_vector` `s3vector`

Misurazione e fatturazione

Per informazioni sulla misurazione e la fatturazione di questa funzionalità, consulta i prezzi di Amazon OpenSearch Service.

Disattivazione del motore s3vector

Prima di disabilitare il `s3vector` motore, elimina tutti gli indici che lo stanno attualmente utilizzando. In caso contrario, qualsiasi tentativo di disattivare il motore avrà esito negativo.

Tieni inoltre presente che l'attivazione o la disabilitazione del `s3vector` motore attiva una distribuzione blu/verde sul tuo dominio.

Per disabilitare il `s3vector` motore, modifica la configurazione del dominio e impostalo. `S3VectorsEngine.Enabled: false`

Ricerca K-Nearest Neighbor (k-NN) su Amazon Service OpenSearch

Acronimo dell'algoritmo k-Nearest Neighbors associato, k-NN for OpenSearch Amazon Service consente di cercare punti in uno spazio vettoriale e trovare i «vicini più vicini» per tali punti in base alla distanza euclidea o alla somiglianza del coseno. Nei casi d'uso sono inclusi suggerimenti (ad esempio, una funzionalità "altri brani che potrebbero piacerti" in un'applicazione musicale), il riconoscimento delle immagini e il rilevamento delle frodi.

> **Note**
>
> Questa documentazione fornisce una breve panoramica del plug-in k-NN, oltre alle limitazioni relative all'utilizzo del plug-in con un servizio gestito. OpenSearch Per una documentazione completa del plugin k-NN, inclusi esempi semplici e complessi, riferimenti ai parametri e il riferimento completo all'API, consulta la documentazione open source. OpenSearch La documentazione open source copre anche l'ottimizzazione delle prestazioni e k-NN-specific le impostazioni del cluster.

Nozioni di base su k-NN

Per utilizzare k-NN, è necessario creare un indice con l'impostazione `index.knn` e aggiungere uno o più campi del tipo di dati `knn_vector`.

```
PUT my-index

{
  "settings": {
    "index.knn": true
  },
  "mappings": {
    "properties": {
      "my_vector1": {
        "type": "knn_vector",
        "dimension": 2
      },
      "my_vector2": {
        "type": "knn_vector",
        "dimension": 4
      }
    }
  }
}
```

Il tipo di dati `knn_vector` supporta un singolo elenco di fino a 10.000 numeri in virgola mobile, con il numero di numeri in virgola mobile definito dal parametro `dimension` richiesto. Dopo aver creato l'indice, aggiungi alcuni dati.

```
POST _bulk

{ "index": { "_index": "my-index", "_id": "1" } }
{ "my_vector1": [1.5, 2.5], "price": 12.2 }
{ "index": { "_index": "my-index", "_id": "2" } }
{ "my_vector1": [2.5, 3.5], "price": 7.1 }
{ "index": { "_index": "my-index", "_id": "3" } }
{ "my_vector1": [3.5, 4.5], "price": 12.9 }
{ "index": { "_index": "my-index", "_id": "4" } }
{ "my_vector1": [5.5, 6.5], "price": 1.2 }
{ "index": { "_index": "my-index", "_id": "5" } }
{ "my_vector1": [4.5, 5.5], "price": 3.7 }
{ "index": { "_index": "my-index", "_id": "6" } }
```

```
{ "my_vector2": [1.5, 5.5, 4.5, 6.4], "price": 10.3 }
{ "index": { "_index": "my-index", "_id": "7" } }
{ "my_vector2": [2.5, 3.5, 5.6, 6.7], "price": 5.5 }
{ "index": { "_index": "my-index", "_id": "8" } }
{ "my_vector2": [4.5, 5.5, 6.7, 3.7], "price": 4.4 }
{ "index": { "_index": "my-index", "_id": "9" } }
{ "my_vector2": [1.5, 5.5, 4.5, 6.4], "price": 8.9 }
```

Quindi puoi cercare i dati utilizzando il tipo di query `knn`.

```
GET my-index/_search
{
  "size": 2,
  "query": {
    "knn": {
      "my_vector2": {
        "vector": [2, 3, 5, 6],
        "k": 2
      }
    }
  }
}
```

In questo caso, `k` è il numero di neighbors che la query deve restituire, ma è necessario includere anche l'opzione `size`. In caso contrario, vengono ottenuti i risultati `k` per ogni partizione (e ogni segmento) anziché i risultati `k` per l'intera query. k-NN supporta un valore massimo di `k` pari a 10.000.

Se si combina la query `knn` con altre clausole, è possibile che vengano restituiti meno risultati `k`. In questo esempio, la clausola `post_filter` riduce il numero di risultati da 2 a 1.

```
GET my-index/_search

{
  "size": 2,
  "query": {
    "knn": {
      "my_vector2": {
        "vector": [2, 3, 5, 6],
        "k": 2
      }
    }
  },
```

```
  "post_filter": {
    "range": {
      "price": {
        "gte": 6,
        "lte": 10
      }
    }
  }
}
```

Se devi gestire un grande volume di query mantenendo prestazioni ottimali, puoi utilizzare l'_msearchAPI per creare una ricerca di massa con JSON e inviare una singola richiesta per eseguire più ricerche:

```
GET _msearch

{ "index": "my-index"}
{ "query": { "knn": {"my_vector2":{"vector": [2, 3, 5, 6],"k":2 }} } }
{ "index": "my-index", "search_type": "dfs_query_then_fetch"}
{ "query": { "knn": {"my_vector1":{"vector": [2, 3],"k":2 }} } }
```

Il video seguente mostra come configurare ricerche vettoriali di massa per le query K-NN.

Differenze, regolazione e limitazioni di k-NN

OpenSearch consente di modificare tutte le impostazioni k-NN utilizzando l'API. `_cluster/settings` In OpenSearch Service, puoi modificare tutte le impostazioni tranne `knn.memory.circuit_breaker.enabled` e`knn.circuit_breaker.triggered`. Le statistiche k-NN sono incluse come metriche di Amazon CloudWatch .

In particolare, confronta la `KNNGraphMemoryUsage` metrica su ciascun nodo di dati con la `knn.memory.circuit_breaker.limit` statistica e la RAM disponibile per il tipo di istanza. OpenSearch Il servizio utilizza metà della RAM di un'istanza per l'heap Java (fino a una dimensione dell'heap di 32 GiB). Per impostazione predefinita, k-NN utilizza fino al 50% della metà rimanente, quindi un tipo di istanza con 32 GiB di RAM può ospitare 8 GiB di grafici (32 * 0,5 * 0,5). Le prestazioni possono risentirne se l'utilizzo della memoria grafica supera questo valore.

È possibile migrare un indice k-NN creato nella versione 2.x o successiva o archiviare a freddo su un dominio con la versione 2.17 UltraWarmo successiva.

Le API Clear cache e le API di riscaldamento per gli indici k-NN sono bloccate per gli indici caldi. Quando viene avviata la prima query per l'indice, scarica i file grafici da Amazon S3 e lo carica in memoria. Allo stesso modo, quando il TTL è scaduto per i grafici, i file vengono automaticamente rimossi dalla memoria.

Ingestione vettoriale

L'ingestione vettoriale consente di importare e indicizzare rapidamente domini e raccolte Serverless. OpenSearch OpenSearch Il servizio esamina il dominio o la raccolta e crea una pipeline di inserimento per conto dell'utente in cui caricare i dati. OpenSearch L'ingestione e l'indicizzazione del tuo dominio o della tua raccolta sono gestite per te da Vector ingestion.

Puoi accelerare e ottimizzare il processo di indicizzazione abilitando nuove funzionalità. Accelerazione GPU per l'indicizzazione vettoriale Ottimizzazione automatica Con Vector ingestion, non è necessario gestire l'infrastruttura sottostante, applicare patch al software o scalare i cluster per supportare l'indicizzazione e l'ingestione del database vettoriale. Ciò consente di creare rapidamente il database vettoriale per soddisfare le proprie esigenze.

Come funziona

L'ingestione vettoriale esamina il dominio o la raccolta e il relativo indice. È possibile configurare manualmente i campi dell'indice vettoriale o consentire OpenSearch l'utilizzo della configurazione automatica.

L'ingestione vettoriale utilizza OpenSearch Ingestion (OSI) come pipeline di dati tra Amazon S3 e. OpenSearch Il servizio elabora i vettori in parallelo per ottimizzare la velocità di inserimento rispettando i limiti di scalabilità di OSI e. OpenSearch

OpenSearch Prezzi di importazione vettoriale

In un momento specifico, paghi solo per il numero di ingestioni vettoriali assegnate a una pipeline, indipendentemente dal fatto OCUs che i dati scorrano attraverso la pipeline. OpenSearch vector ingestion soddisfa immediatamente i carichi di lavoro aumentando o diminuendo la capacità della pipeline in base all'utilizzo.

Per i dettagli completi sui prezzi, consulta la pagina dei prezzi OpenSearch di Amazon Service.

Prerequisiti

Prima di utilizzare l'ingestione vettoriale, assicurati di disporre delle seguenti risorse:

- Bucket Amazon S3 contenente i tuoi documenti OpenSearch JSON in formato Parquet o JSONL
- OpenSearch risorsa: un dominio o una raccolta
- OpenSearch versione `2.19` o successiva (richiesta per l'integrazione con ottimizzazione automatica)

Crea un database vettoriale

Utilizza il flusso di lavoro di creazione di lavori di inserimento vettoriale per configurare l'ottimizzazione automatica degli indici vettoriali e accelerare la creazione di indici su larga scala.

> **Note**
>
> Il contenuto procedurale di questa sezione è soggetto a modifiche man mano che l'interfaccia utente viene finalizzata. Il flusso di lavoro potrebbe essere aggiornato nelle versioni future per riflettere l'esperienza più recente sulla console.

Per creare un processo di inserimento vettoriale

1. Nella sezione Dettagli del lavoro di ingestione vettoriale, in Nome, inserisci un nome per il tuo lavoro di importazione.
2. Nella sezione Origine dati, configura quanto segue:
   a. Per l'URI di Amazon S3, inserisci la posizione del bucket Amazon S3 contenente OpenSearch i tuoi documenti Service JSON.
   b. Scegli Browse Amazon S3 per selezionare tra i bucket disponibili oppure scegli Visualizza per visualizzare in anteprima il contenuto del bucket.
   c. Per Tipo di contenuto, seleziona una delle seguenti opzioni:
      - Vettori: i documenti contengono già dei vettori e non richiedono un'ulteriore generazione di vettori incorporati.
      - Testo, immagine o audio: i documenti contengono contenuti come testo, immagini o byte audio che devono essere codificati in incorporamenti vettoriali.
3. Nella sezione Autorizzazioni per l'origine dei dati, configura le autorizzazioni di accesso:
   a. Per il ruolo IAM, scegli una delle seguenti opzioni:

- Crea un nuovo ruolo
- Usa un ruolo esistente

b. Per il nome del ruolo IAM, inserisci un nome per il ruolo.

4. Nella sezione Destinazione, configura l'endpoint del OpenSearch servizio:

   a. Per Endpoint, scegli Scegli un'opzione per selezionare tra i domini o le raccolte compatibili nella regione corrente.

   b. Scegli Avanti per procedere con l'endpoint selezionato.

5. Scegli Avanti per continuare con il passaggio successivo oppure scegli Annulla per uscire senza salvare.

Funzionalità correlate

L'ingestione vettoriale funziona con le seguenti funzionalità di Amazon OpenSearch Service per ottimizzare le prestazioni del database vettoriale:

Accelerazione GPU per l'indicizzazione vettoriale

L'accelerazione tramite GPU riduce il tempo necessario per creare, aggiornare ed eliminare gli indici vettoriali. Se utilizzata con l'ingestione vettoriale, è possibile accelerare in modo significativo il processo di inserimento e indicizzazione per database vettoriali su larga scala.

Ottimizzazione automatica

L'ottimizzazione automatica rileva automaticamente i compromessi ottimali tra latenza di ricerca, qualità e requisiti di memoria. L'ingestione vettoriale può applicare consigli di ottimizzazione automatica durante il processo di ingestione per garantire che gli indici vettoriali siano configurati in modo ottimale.

Per ottenere i migliori risultati, prendi in considerazione la possibilità di abilitare sia l'accelerazione GPU che l'ottimizzazione automatica quando utilizzi l'ingestione vettoriale per creare database vettoriali su larga scala.

Esporta l'indice vettoriale di Amazon S3 nel motore vettoriale Service OpenSearch

Un' point-in-timeesportazione dell'indice vettoriale Amazon S3 selezionato in Service. OpenSearch Il motore vettoriale OpenSearch Service fornisce un archivio vettoriale semplice e scalabile con funzionalità di ricerca avanzate.

Per esportare l'indice vettoriale di Amazon S3 nel motore vettoriale Service OpenSearch

1. Nella sezione Sorgente, verifica i dettagli dell'indice vettoriale di Amazon S3:
   - Indice vettoriale Amazon S3: il nome dell'indice di origine
   - Indice vettoriale Amazon S3 ARN: il nome della risorsa Amazon del tuo indice
2. Nella sezione Accesso al servizio, configura OpenSearch l'autorizzazione al servizio:
   a. Per Scegli un metodo per autorizzare il OpenSearch servizio, seleziona una delle seguenti opzioni:
      - Creare e utilizzare un nuovo ruolo di servizio
      - Usa un ruolo di servizio esistente
   b. Per Nome del ruolo di servizio, inserisci un nome per il ruolo di servizio.

      > **Note**
      >
      > Il nome del ruolo di servizio deve contenere da 1 a 64 caratteri. I caratteri validi sono a-z, A-Z, 0-9 e punti (.).

   c. Scegli Visualizza i dettagli delle autorizzazioni per esaminare le autorizzazioni richieste.
3. Espandi Impostazioni aggiuntive: facoltativo per configurare le opzioni avanzate, se necessario.
4. Nella sezione Dettagli di esportazione, configura le seguenti opzioni:
   - Automatizza la creazione OpenSearch di raccolte vettoriali nei OpenSearch servizi: le raccolte di servizi vengono utilizzate per archiviare dati vettoriali. La capacità di elaborazione serverless viene misurata in OpenSearch Service Compute Units (OCUs), per impostazione predefinita la capacità massima dell'OCU è 50.

- Automatizza la creazione di ruoli IAM per l'accesso al servizio: questo ruolo viene utilizzato da OpenSearch Service per leggere l'indice vettoriale di Amazon S3 e scrivere OpenSearch nella raccolta Service.
- Automatizza la creazione della pipeline di importazione dei OpenSearch servizi: le pipeline di importazione dei OpenSearch servizi vengono utilizzate per importare i dati. Un bucket Amazon S3 viene creato come best practice per acquisire e archiviare eventi non riusciti in un bucket Amazon S3 Dead Letter Queue (DLQ), permettendo un facile accesso per la risoluzione dei problemi e l'analisi.

5. Scegli Esporta per avviare il processo di esportazione o scegli Annulla per uscire senza esportare.

Importa lo spazio dei nomi vettoriali Amazon S3 nel motore vettoriale Service OpenSearch

L'analisi dei dati vettoriali con OpenSearch Service richiede una raccolta di servizi una tantum OpenSearch e la configurazione delle autorizzazioni IAM.

Per importare lo spazio dei nomi vettoriali Amazon S3 nel motore vettoriale Service OpenSearch

1. Nella sezione Sorgente, configura l'indice vettoriale Amazon S3:
   - Per l'indice vettoriale Amazon S3 (ARN), inserisci l'ARN del tuo indice vettoriale Amazon S3.

     > **Note**
     >
     > Deve essere nel formato arn:aws:iam: :account-id: /*:index vector-bucket-name

2. Nella sezione Accesso al servizio, configura l'autorizzazione al servizio: OpenSearch
   a. Per Scegli un metodo per autorizzare il OpenSearch servizio, seleziona una delle seguenti opzioni:
      - Creare e utilizzare un nuovo ruolo di servizio
      - Usa un ruolo di servizio esistente
   b. Per Nome del ruolo di servizio, inserisci un nome per il ruolo di servizio.

> **Note**
>
> Il nome del ruolo di servizio deve contenere da 1 a 64 caratteri. I caratteri validi sono a-z, A-Z, 0-9 e punti (.).

c. Scegli Visualizza i dettagli delle autorizzazioni per esaminare le autorizzazioni richieste.

3. Espandi Impostazioni aggiuntive: facoltativo per configurare le opzioni avanzate, se necessario.
4. Nella sezione Passaggi di importazione, configura le seguenti opzioni di automazione:

   - Automatizza la creazione OpenSearch di raccolte vettoriali di OpenSearch Service: le raccolte di servizi vengono utilizzate per archiviare dati vettoriali. La capacità di elaborazione serverless viene misurata in OpenSearch Service Compute Units (OCUs), per impostazione predefinita la capacità massima dell'OCU è 50.
   - Automatizza la creazione di ruoli IAM per l'accesso al servizio: questo ruolo viene utilizzato da OpenSearch Service per leggere l'indice vettoriale di Amazon S3 e scrivere OpenSearch nella raccolta Service.
   - Automatizza la creazione della pipeline OpenSearch di ingestione dei servizi: le pipeline di importazione dei OpenSearch servizi vengono utilizzate per importare i dati. Un bucket Amazon S3 viene creato come best practice per acquisire e archiviare eventi non riusciti in un bucket Amazon S3 Dead Letter Queue (DLQ), permettendo un facile accesso per la risoluzione dei problemi e l'analisi.

5. Scegli Importa per avviare il processo di importazione o scegli Annulla per uscire senza importare.

Visualizza i processi di inserimento vettoriale e la cronologia delle importazioni

I lavori di inserimento vettoriale creano una pipeline per la vettorializzazione dei set di dati, l'automazione dell'ottimizzazione degli indici vettoriali e l'accelerazione della creazione di indici su larga scala.

Per visualizzare i lavori di ingestione vettoriale

1. Nella sezione Lavori di importazione vettoriale, visualizza le informazioni di riepilogo:

   - Lavori: numero totale di lavori di ingestione

- Scegli Crea database vettoriale per creare un nuovo lavoro di importazione

2. Nella sezione Importazioni vettoriali di Amazon S3, visualizza il riepilogo delle importazioni:
   - Importazioni totali - Numero di importazioni completate
   - Scegli Importa vettori Amazon S3 per iniziare una nuova importazione
3. Nella tabella dei lavori di ingestione di Vector, monitora i lavori attivi con le seguenti informazioni:
   - Nome: il nome del lavoro
   - Stato: stato attuale del lavoro (ad esempio, Attivo)
   - Fonte dei dati: posizione della fonte (ad esempio, s3://location)
   - Destinazione: destinazione
   - Ultimo aggiornamento: data e ora dell'aggiornamento più recente
4. Utilizza la casella di ricerca per trovare un lavoro di ingestione vettoriale per individuare lavori specifici.
5. Per gestire i lavori, scegli una delle seguenti azioni:
   - Scegli Elimina per rimuovere i lavori selezionati
   - Scegli Crea database vettoriale per creare lavori aggiuntivi
6. Nella sezione Cronologia delle importazioni di vettori Amazon S3, tieni traccia degli eventi di importazione:
   a. Utilizza il filtro Intervallo di date per specificare un periodo di tempo per la cronologia delle importazioni.
   b. Utilizza il menu a discesa Stato per filtrare in base allo stato dell'importazione (ad esempio, Qualsiasi stato).
   c. Usa la casella di ricerca per trovare le importazioni tramite l'indice vettoriale di Amazon S3 na... per individuare importazioni specifiche.
   d. Visualizza i dettagli dell'importazione, tra cui:
      - Importazione iniziata il (UTC+ 5:30) - Quando è iniziata l'importazione
      - Stato dell'importazione: stato attuale (in corso, completo, non riuscito, parzialmente completato)
      - ARN - Indice vettoriale Amazon S3 - Identificatore dell'indice di origine
      - OpenSearch Raccolta vettoriale di servizio - Raccolta di destinazioni

7. Scegli Importa vettore Amazon S3 per iniziare un nuovo processo di importazione.

Ottimizzazione automatica

Auto-optimize è un servizio che automatizza le ottimizzazioni degli indici vettoriali, consentendo agli utenti di bilanciare qualità, velocità e costi della ricerca senza richiedere settimane di ottimizzazione manuale da parte di esperti. Valuta le configurazioni degli indici in base ai requisiti di latenza e richiamo definiti dall'utente e genera consigli di ottimizzazione, quindi è richiesta una competenza minima. I consigli vengono in genere forniti entro 30-60 minuti.

La configurazione tradizionale degli indici vettoriali richiede competenze e sperimentazioni significative per ottenere prestazioni ottimali. Parametri come `ef_construction` (che controlla la qualità di costruzione dell'indice), `m` (che determina il numero di connessioni grafiche), `ef_search` (che controlla la ricerca HNSW) e metodi di quantizzazione (quantizzazione binaria (32x, 16x, 8x), quantizzazione scalare (4x)) influiscono in modo significativo sia sulla precisione della ricerca che sull'utilizzo delle risorse. L'ottimizzazione automatica utilizza algoritmi di ottimizzazione degli iperparametri per scoprire configurazioni di indici che sono unicamente ottimali per il set di dati entro i requisiti di latenza e richiamo definiti.

Vantaggi

L'ottimizzazione automatica per offre i seguenti vantaggi: OpenSearch

- Regolazione automatica dei parametri: elimina la sperimentazione manuale con algoritmi (HNSW), quantizzazione, rescoring e parametri del motore, risparmiando tempo e riducendo la curva di apprendimento per l'ottimizzazione della ricerca vettoriale.
- Ottimizza la velocità di ricerca: per impostazione predefinita, è configurato per le prestazioni in memoria. OpenSearch L'ottimizzazione automatica rileva compromessi favorevoli che migliorano la qualità della ricerca e consentono di risparmiare sui costi, mantenendo al contempo una velocità di ricerca accettabile.
- Ottimizzazione dei costi: riduce i costi trovando opzioni per ridurre i requisiti di memoria dell'indice, riducendo al minimo i compromessi in termini di qualità della ricerca e velocità.
- Ottimizza la qualità della ricerca: offri potenzialmente un numero di richiami più elevato rispetto alle impostazioni predefinite o scopri compromessi vantaggiosi che offrono risparmi significativi sui costi con una perdita minima di richiami.

L'ottimizzazione automatica funziona insieme ad altre OpenSearch funzionalità, ad esempio Accelerazione GPU per l'indicizzazione vettoriale per fornire un'ottimizzazione completa delle prestazioni per i carichi di lavoro di ricerca vettoriale.

Come funziona

L'ottimizzazione automatica funziona attraverso un'architettura basata sul lavoro che analizza i dati vettoriali e fornisce consigli di ottimizzazione. Punti chiave:

- Gli utenti condividono i propri set di dati in formato Parquet o JSONL in un bucket Amazon S3.
- Configurano i processi di ottimizzazione automatica senza server configurando le soglie di richiamo e latenza accettabili. Soglie più rilassate consentono al servizio di scoprire ottimizzazioni dei costi più significative.
- Ottimizza automaticamente i lavori eseguiti su un'infrastruttura completamente gestita da Amazon Service. OpenSearch I lavori non consumano risorse sul tuo dominio o sulle tue raccolte. I lavoratori eseguono in parallelo per valutare le configurazioni degli indici e utilizzano il campionamento su set di dati di grandi dimensioni per fornire risultati in genere entro 30-60 minuti.
- Ogni lavoro viene fatturato in base a una tariffa forfettaria prevedibile. Per informazioni sui prezzi, consulta la pagina dei prezzi OpenSearch di Amazon Service.

Prerequisiti

- Formato e autorizzazioni del set di dati: devi avere il set di dati disponibile come uno o più file Parquet o JSONL in una cartella bucket Amazon S3. Esempio:
  - Parquet: e `s3://dataset-bucket-us-east-1/dataset_folder/first_half.parquet s3://dataset-bucket-us-east-1/dataset_folder/second_half.parquet`
  - JSONL: `s3://dataset-bucket-us-east-1/dataset_folder/data.jsonl`

  Fornisci l'URI della cartella che la contiene (ad esempio,). `s3://dataset-bucket-us-east-1/dataset_folder/` La cartella deve contenere file di un unico formato: non mischiate file Parquet e JSONL nella stessa cartella. Questo set di dati verrà utilizzato per generare i consigli. Assicurati che il tuo ruolo federato disponga delle seguenti autorizzazioni Amazon S3 su quella risorsa:. `"s3:Get*", "s3:List*", "s3:Describe*"`
- Specificare i metadati corretti del set di dati: il set di dati fornito deve contenere righe di valori float. Il nome di ogni colonna e la dimensionalità di ogni vettore devono corrispondere alle opzioni fornite nella console. Ad esempio, se il set di dati contiene vettori denominati `train_data` che

rappresentano ciascuna 768 dimensione, questi valori devono corrispondere alla console di ottimizzazione automatica.

- (Se si utilizza la funzionalità di inserimento vettoriale): se si prevede di utilizzare la funzionalità di inserimento (seguendo i consigli di ottimizzazione automatica per creare automaticamente dati di indice e inserimento), è necessario configurare il cluster per concedere l'autorizzazione all'ottimizzazione automatica per importare il set di dati nel OpenSearch cluster. OpenSearch Per i OpenSearch domini con una politica di accesso al dominio, concedi l'accesso al ruolo appena creato tramite tale politica. Per i OpenSearch domini con controllo granulare degli accessi, aggiungi il ruolo pipeline come ruolo di backend. Per le raccolte OpenSearch serverless, aggiungi il ruolo pipeline alla politica di accesso ai dati.
- Autorizzazioni IAM: sono necessarie le seguenti autorizzazioni IAM per utilizzare l'ottimizzazione automatica:
  - `opensearch:SubmitAutoOptimizeJob`
  - `opensearch:GetAutoOptimizeJob`
  - `opensearch:DeleteAutoOptimizeJob`
  - `opensearch:CancelAutoOptimizeJob`
  - `opensearch:ListAutoOptimizeJobs`

  > **Note**
  >
  > Si tratta di politiche basate sull'identità. L'ottimizzazione automatica non supporta politiche basate sulle risorse.

- Scadenza delle credenziali: configura la sessione utente federata in modo che abbia una scadenza minima delle credenziali di almeno 1 ora. Per set di dati molto grandi o di dimensioni elevate, valuta la possibilità di aumentare la durata di scadenza fino a 3 ore.

Casi d'uso per l'ottimizzazione automatica

L'ottimizzazione automatica è particolarmente utile nei seguenti scenari:

Ottimizzazione della configurazione iniziale

Quando si implementano per la prima volta applicazioni di ricerca vettoriale, la determinazione dei parametri HNSW ottimali richiede spesso test approfonditi e competenze di settore. L'ottimizzazione

automatica elimina questo trial-and-error processo analizzando i dati e le caratteristiche del carico di lavoro per consigliare configurazioni pronte per la produzione.

Questo caso d'uso è ideale per i team alle prime armi con la ricerca vettoriale o per coloro che migrano da altre piattaforme di database vettoriali e che devono stabilire rapidamente configurazioni di base.

Ottimizzazione della scalabilità

Man mano che il set di dati vettoriali passa da migliaia a milioni di vettori, i parametri che inizialmente funzionavano bene potrebbero diventare non ottimali. L'ottimizzazione automatica consiglia le regolazioni per mantenere le prestazioni su larga scala.

Riduzione dei costi

Gli indici vettoriali possono consumare notevoli risorse di calcolo e archiviazione, in particolare con incorporamenti ad alta dimensione. L'ottimizzazione automatica identifica le opportunità di riduzione dei costi trovando configurazioni di parametri più efficienti che mantengano i livelli di prestazioni richiesti utilizzando meno risorse.

Ad esempio, l'ottimizzazione automatica potrebbe scoprire che il `m` valore corrente (connessione al grafico) è superiore a quello necessario per i requisiti di precisione, consentendoti di ridurre il tempo di indicizzazione e lo spazio di archiviazione senza influire sulla qualità della ricerca.

Risoluzione dei problemi relativi alle prestazioni

In caso di prestazioni di query lente o latenza elevata nelle operazioni di ricerca vettoriale, l'ottimizzazione automatica può analizzare il set di dati e identificare una configurazione più ottimale. Il servizio fornisce raccomandazioni specifiche per risolvere i problemi di prestazioni, come la regolazione della connettività dei grafici o dei parametri di ricerca.

Limitazioni

- Disponibilità regionale: l'ottimizzazione automatica è disponibile solo nelle seguenti regioni: AWS
  - ap-south-1
  - eu-west-1
  - us-west-2
  - us-east-2

  - us-east-1
  - eu-central-1
  - ap-southeast-2
  - ap-northeast-1
  - ap-southeast-1
- Tipi di raccolta: l'ottimizzazione automatica è supportata solo per le raccolte e i OpenSearch domini di ricerca vettoriale (2.19, 3.1 e 3.3).
- Supporto per i motori

  Supporto del motore per tipo di implementazione

Motore	Serverless	OpenSearch Gestito
Lucene	No	Sì
Faiss	Sì	Sì
Nmslib	No	No

- Supporto per algoritmi: l'ottimizzazione automatica supporta solo indici vettoriali basati su HNSW.
- Lavori simultanei: puoi eseguire fino a 10 lavori di ottimizzazione simultanei per account per regione. Nessun nuovo lavoro può essere accettato se viene raggiunto il limite.
- Durata del lavoro: i processi di ottimizzazione possono richiedere da 15 minuti a diverse ore a seconda della dimensione del set di dati, della dimensione e delle metriche prestazionali richieste.
- Consigli: l'ottimizzazione automatica suggerisce solo fino a 3 consigli.
- Set di dati
  - Formati supportati: Parquet, JSONL
  - Archivio dati: Amazon S3

Fatturazione e costi

L'ottimizzazione automatica utilizza un modello di prezzo per processo in cui si paga per ogni lavoro di ottimizzazione riuscito indipendentemente dalle dimensioni del set di dati e dalle configurazioni di ottimizzazione. Non ti verrà addebitato alcun costo per i lavori non riusciti o annullati. Inoltre, l'ottimizzazione automatica viene eseguita su un'infrastruttura separata rispetto ai OpenSearch cluster gestiti o senza server, quindi non influisce sull'utilizzo delle risorse dei cluster preesistenti.

Modello tariffario

I costi di ottimizzazione automatica vengono fatturati separatamente dai costi standard di elaborazione e storage dei domini OpenSearch serverless o gestiti. OpenSearch

Per informazioni sui prezzi, consulta la pagina dei prezzi OpenSearch di Amazon Service.

Formati di dati supportati

L'ottimizzazione automatica supporta i seguenti formati di dati per i set di dati vettoriali archiviati in Amazon S3:

Formato Parquet

Parquet è un formato di archiviazione colonnare ottimizzato per carichi di lavoro analitici. Ogni file Parquet deve contenere una colonna di array float che rappresentano i dati vettoriali.

Esempio di struttura del file Parquet (vista come tabella):

```
| id  | train_data                      |
|-----|---------------------------------|
| 1   | [0.12, 0.45, 0.78, ..., 0.33] |
| 2   | [0.56, 0.89, 0.12, ..., 0.67] |
| 3   | [0.34, 0.67, 0.90, ..., 0.11] |
```

Formato JSONL

JSONL (JSON Lines) è un formato di testo in cui ogni riga è un oggetto JSON valido. Ogni riga deve contenere un campo con una matrice float che rappresenta i dati vettoriali.

File JSONL di esempio:

```
{"id": 1, "train_data": [0.12, 0.45, 0.78, 0.33]}
{"id": 2, "train_data": [0.56, 0.89, 0.12, 0.67]}
{"id": 3, "train_data": [0.34, 0.67, 0.90, 0.11]}
```

Conversione tra formati

Se i dati sono in un formato diverso, è possibile utilizzare i seguenti script Python per convertirli.

Converti JSON o JSONL in Parquet

```
#!/usr/bin/env python3
import json
import pyarrow as pa
import pyarrow.parquet as pq
from pathlib import Path
from typing import Any, Dict, List


def load_json_any(path: Path) -> List[Dict[str, Any]]:
    """
    Load JSON that can be:
      - a list of objects
      - a single object
      - JSON Lines (one object per line)
    Returns list[dict].
    """
    text = path.read_text().strip()

    # Try full JSON file
    try:
        obj = json.loads(text)
        if isinstance(obj, list):
            return obj
        if isinstance(obj, dict):
            return [obj]
    except json.JSONDecodeError:
        pass

    # Fallback # JSON Lines
    records = []
    for i, line in enumerate(text.splitlines(), start=1):
        line = line.strip()
        if not line:
            continue
        try:
            rec = json.loads(line)
        except json.JSONDecodeError as e:
            raise ValueError(f"Invalid JSON on line {i}: {e}")
        if not isinstance(rec, dict):
            raise ValueError(f"Line {i} must contain a JSON object")
        records.append(rec)
```

```
    return records


def json_to_parquet(json_path: str, parquet_path: str, compression: str = "snappy"):
    """Convert ANY JSON to Parquet (schema inferred)."""
    records = load_json_any(Path(json_path))
    table = pa.Table.from_pylist(records)
    pq.write_table(table, parquet_path, compression=compression)
    print(f"Wrote {len(records)} rows to {parquet_path}")


if __name__ == "__main__":
    INPUT_JSON = "vectors.jsonl"
    OUTPUT_PARQUET = "vectors.parquet"
    json_to_parquet(INPUT_JSON, OUTPUT_PARQUET)
```

Converti Parquet in JSONL

```
#!/usr/bin/env python3
import json
import pyarrow.parquet as pq


def parquet_to_jsonl(parquet_path: str, jsonl_path: str):
    """Convert a Parquet file to JSONL format."""
    table = pq.read_table(parquet_path)
    rows = table.to_pylist()
    with open(jsonl_path, "w") as f:
        for row in rows:
            f.write(json.dumps(row) + "\n")
    print(f"Wrote {len(rows)} rows to {jsonl_path}")


if __name__ == "__main__":
    INPUT_PARQUET = "vectors.parquet"
    OUTPUT_JSONL = "vectors.jsonl"
    parquet_to_jsonl(INPUT_PARQUET, OUTPUT_JSONL)
```

Utilizzo dell'ottimizzazione automatica nella console

Puoi utilizzare la console di Amazon OpenSearch Service per creare lavori di inserimento vettoriale, monitorarne l'avanzamento, visualizzare consigli di ottimizzazione e creare indici basati su tali consigli.

Prerequisiti

Prima di poter utilizzare l'ottimizzazione automatica nella console, è necessario disporre di quanto segue:

- Un AWS account attivo con accesso alla OpenSearch console.
- Una raccolta OpenSearch Serverless esistente di Type Vector Search o un dominio gestito OpenSearch .
- Autorizzazioni IAM per le seguenti azioni:
  - `opensearch:SubmitAutoOptimizeJob`
  - `opensearch:GetAutoOptimizeJob`
  - `opensearch:DeleteAutoOptimizeJob`
  - `opensearch:CancelAutoOptimizeJob`
  - `opensearch:ListAutoOptimizeJobs`

  > **Note**
  >
  > Si tratta di politiche basate sull'identità. AWS non supporta politiche basate sulle risorse per l'ottimizzazione automatica delle risorse.

- Configura la tua sessione utente federata in modo che abbia una scadenza minima delle credenziali di almeno 1 ora. Per set di dati molto grandi o di dimensioni elevate, valuta la possibilità di aumentare la durata di scadenza fino a 3 ore.

Creazione di un processo di inserimento vettoriale

Un processo di inserimento vettoriale analizza i dati vettoriali e fornisce consigli di ottimizzazione per la configurazione degli indici.

Per creare un processo di ingestione vettoriale

1. Accedi alla console di Amazon OpenSearch Service all'indirizzo Console di gestione AWS.
2. Nel riquadro di navigazione, scegli Auto-Optimize.
3. Scegli Crea un lavoro di ingestione vettoriale.
4. In Dettagli del lavoro, inserisci un nome per il tuo lavoro di ingestione vettoriale. Questo nome consente di identificare il lavoro nella console.
5. In Origine dati, configura quanto segue:

   a. Per l'URI di Amazon S3, inserisci l'URI Amazon S3 della cartella contenente i tuoi file di dati (Parquet o JSONL). L'URI deve puntare alla cartella che la contiene, non a singoli file. Ad esempio, se il file si trova in `s3://my-bucket/my-folder/file1.parquet` o `s3://my-bucket/my-folder/data.jsonl`, inserisci `s3://my-bucket/my-folder/`.

   > **Note**
   >
   > La cartella deve contenere file in un unico formato. Non mischiate file Parquet e JSONL nella stessa cartella.

   b. Per Regione, seleziona la AWS regione in cui si trova il tuo bucket Amazon S3. La regione deve corrispondere alla posizione del bucket.
6. In OpenSearch Dominio, seleziona un dominio o una raccolta esistente oppure scegli Crea nuovo per crearne uno.

   > **Note**
   >
   > Puoi specificare un dominio OpenSearch gestito o una raccolta OpenSearch serverless serverless.
7. In Autorizzazioni di origine dati, specifica il ruolo IAM che dispone delle autorizzazioni per accedere al bucket OpenSearch e al dominio o alla raccolta Amazon S3. Il ruolo deve disporre delle autorizzazioni necessarie in base alla configurazione del dominio o della raccolta:

   - Per i OpenSearch domini con una politica di accesso al dominio, concedi l'accesso al ruolo tramite tale politica.
   - Per i OpenSearch domini con controllo granulare degli accessi, aggiungi il ruolo come ruolo di backend.

- Per le raccolte OpenSearch serverless, aggiungi il ruolo alla politica di accesso ai dati.

8. Scegli Next (Successivo).
9. In Configura indice, specifica quanto segue:
   a. Per Nome campo, inserisci il nome del campo del set di dati che contiene i dati vettoriali.
   b. Per Tipo di spazio, seleziona la metrica della distanza utilizzata per calcolare la distanza tra i vettori:
      - l2 - Distanza euclidea
      - cosinesimil - Somiglianza con il coseno
      - innerproduct - Prodotto interno
   c. Per Dimension, inserisci il numero di valori in virgola mobile in ogni vettore.
10. In Requisiti prestazionali, configura quanto segue:
    a. Per Recall, specificate la qualità di ricerca desiderata come valore decimale compreso tra 0 e 1. Valori di richiamo più elevati restituiscono risultati più pertinenti. Esempio:
       - 0,95 indica che in media vengono restituiti 19 dei 20 veri vettori di documento più vicini a un vettore di query
       - 0.9 indica 9 su 10
       - 0,8 indica 8 su 10
    b. Per i requisiti di latenza della ricerca, seleziona la tua tolleranza di latenza. I requisiti modesti consentono maggiori risparmi sui costi grazie a metodi di compressione che riducono i requisiti di memoria.
11. Scegli Next (Successivo).
12. Controlla la configurazione e scegli Crea.

Il lavoro inizia l'elaborazione. È possibile monitorarne l'avanzamento nella tabella Vector Ingestion Jobs.

Monitoraggio dei lavori di ottimizzazione

Puoi monitorare lo stato dei tuoi lavori di inserimento vettoriale dalla pagina di destinazione con ottimizzazione automatica.

Per monitorare i lavori di ottimizzazione

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione, scegli Ottimizzazione automatica.
3. La tabella Vector Ingestion Jobs mostra tutti i lavori con il loro stato corrente. Aggiorna la pagina per visualizzare informazioni sullo stato aggiornate.

> **Note**
>
> Non esiste un meccanismo di aggiornamento o notifica automatico. È necessario aggiornare manualmente la console per vedere quando un processo viene completato.

Comprendere gli stati dello stato del lavoro

I lavori di ottimizzazione automatica possono avere i seguenti valori di stato:

Pending (In attesa)

Il lavoro è in coda e in attesa di avvio.

In esecuzione

Il processo di ottimizzazione automatica consiste nell'analisi attiva dei dati e nella generazione di consigli.

Completato

Il processo di ottimizzazione automatica è stato completato correttamente. Tutte le analisi, le valutazioni e i consigli sono completi e disponibili per la visualizzazione.

Non riuscito

Il processo ha riscontrato un errore. Visualizza i dettagli dell'errore nella pagina dei dettagli del lavoro per determinarne la causa.

Attivo

È stato creato un indice nel cluster allegato e i dati sono stati inseriti.

La durata del job dipende principalmente dalla dimensione del set di dati e dal carico di servizio corrente. I lavori tipici vengono completati entro 15 minuti o diverse ore.

Visualizzazione dettagli processo

È possibile visualizzare informazioni dettagliate su uno specifico lavoro di ottimizzazione, inclusi la configurazione e lo stato.

Per visualizzare i dettagli del lavoro

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel pannello di navigazione, scegli Auto-Optimize.
3. Nella tabella Vector Ingestion Jobs, scegli il nome del lavoro.
4. La pagina dei dettagli del lavoro mostra le seguenti informazioni:

   - Nome e stato del Job
   - Configurazione dell'origine dati (URI e regione Amazon S3)
   - OpenSearch dominio o raccolta
   - Configurazione dell'indice (nome del campo, tipo di spazio, dimensione)
   - Requisiti prestazionali (richiamo e latenza)
   - Messaggi di errore (se il processo non è riuscito)

Visualizzazione e comprensione dei risultati

Una volta completato correttamente un lavoro, è possibile visualizzare i consigli di ottimizzazione.

Per visualizzare i risultati dell'ottimizzazione

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel pannello di navigazione, scegli Auto-Optimize.
3. Nella tabella Vector Ingestion Jobs, scegli un lavoro con lo stato Completato.
4. La pagina dei risultati mostra le seguenti sezioni:

   - Panoramica dei risultati: mostra la qualità di richiamo stimata della ricerca rispetto al requisito richiesto e l'indice di memoria occupato rispetto alla configurazione più consigliata.
   - Consigli: elenca fino a tre consigli di ottimizzazione, ordinati in base alla raccomandazione principale, che corrisponde meglio alla configurazione in uso. Ogni raccomandazione include:

- Parametri di configurazione dell'indice
- Parametri di configurazione della ricerca
- Metriche delle prestazioni previste
- Stime dell'impronta di memoria

> **Note**
>
> Sebbene i consigli siano ordinati in base alla migliore corrispondenza, puoi selezionare qualsiasi consiglio più adatto al tuo caso d'uso specifico. Ottimizza automaticamente i tentativi di trovare le corrispondenze più vicine ai criteri di richiamo scelti.

Creazione di un indice basato sui consigli

Dopo aver esaminato i consigli di ottimizzazione, è possibile creare manualmente un indice utilizzando la configurazione consigliata o creare automaticamente un indice con il consiglio selezionato.

Per creare un indice automaticamente

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel pannello di navigazione, scegli Auto-Optimize.
3. Nella tabella Vector Ingestion Jobs, scegli un lavoro con lo stato Completato.
4. Esamina i consigli e seleziona quello che desideri utilizzare.
5. Scegli Build index.
6. Il sistema crea automaticamente un indice nel cluster utilizzando il consiglio selezionato e acquisisce i dati vettoriali dal set di dati.

Per creare un indice manualmente

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel pannello di navigazione, scegli Auto-Optimize.
3. Nella tabella Vector Ingestion Jobs, scegli un lavoro con lo stato Completato.

4. Esamina i consigli e annota la configurazione dell'indice e i parametri di configurazione della ricerca per il consiglio scelto.
5. Utilizza l' OpenSearch API o la console per creare manualmente un indice con i parametri consigliati.

Funzionalità correlate

Auto-optimize collabora con altre funzionalità OpenSearch di Amazon Service per aiutarti a creare e ottimizzare applicazioni di ricerca vettoriale:

- Accelerazione GPU per l'indicizzazione vettoriale- Accelera la creazione di indici vettoriali utilizzando l'accelerazione GPU per ridurre tempi e costi di indicizzazione.
- Ingestione vettoriale- Inserisci e indicizza rapidamente dati vettoriali da Amazon S3 nel tuo dominio o nella tua raccolta.

Accelerazione GPU per l'indicizzazione vettoriale

L'accelerazione GPU consente di creare database vettoriali su larga scala in modo più rapido ed efficiente. Puoi abilitare questa funzionalità su domini e raccolte Serverless nuovi o esistenti. OpenSearch OpenSearch Questa funzionalità utilizza l'accelerazione GPU per ridurre il tempo necessario per indicizzare i dati in indici vettoriali.

Con l'accelerazione GPU, puoi aumentare la velocità di indicizzazione vettoriale fino a 10 volte a un quarto del costo di indicizzazione.

Prerequisiti

L'accelerazione GPU è supportata nei domini che eseguono una versione o successiva e nelle raccolte Serverless. OpenSearch OpenSearch `3.1` OpenSearch Per ulteriori informazioni, vedere, e. Aggiornamento dei domini Amazon Service OpenSearch UpdateDomainConfigUpdateCollection APIs

Come funziona

Gli indici vettoriali richiedono risorse di calcolo significative per creare strutture di dati come i grafici Hierarchical Navigable Small Worlds (HNSW). Quando abiliti l'accelerazione GPU sul tuo dominio o sulla tua raccolta, rileva OpenSearch automaticamente le opportunità per accelerare la creazione dell'indice e trasferisce le build dell'indice su istanze GPU. OpenSearch Il servizio gestisce le istanze

GPU per tuo conto, assegnandole al tuo dominio o alla tua raccolta quando necessario. Ciò significa che non gestisci l'utilizzo né paghi per i tempi di inattività.

Paghi solo per l'elaborazione utile tramite Compute Units (OCU) - Vector Acceleration. Ogni OCU di accelerazione vettoriale è una combinazione di circa 8 GiB di memoria CPU, 2 v CPUs e 6 GiB di memoria GPU. Per ulteriori informazioni, consulta the section called “Prezzi dell'accelerazione GPU”.

Per abilitare l'accelerazione GPU per il tuo dominio o la tua raccolta, consulta. Abilitazione dell'accelerazione GPU

Prezzi dell'accelerazione GPU

AWS ti addebita quando OpenSearch rileva opportunità per accelerare i carichi di lavoro di creazione dell'indice del tuo dominio o della tua raccolta. Ogni OCU di accelerazione vettoriale è una combinazione di circa 8 GiB di memoria CPU, 2 v CPUs e 6 GiB di memoria GPU.

AWS fattura l'OCU con granularità di secondo livello. Nel tuo estratto conto, vedrai una voce relativa al calcolo in OCU-Hours.

Ad esempio, quando utilizzi l'accelerazione GPU per un'ora per creare un indice, utilizzando 2 vCPU e 1 GiB di memoria GPU, ti viene fatturato 1 OCU. Se utilizzi 9 GiB di memoria CPU mentre utilizzi l'accelerazione GPU, ti verranno fatturati 2 OCU.

OpenSearch Serverless ne aggiunge altre OCUs in incrementi di 1 OCU in base alla potenza di elaborazione e allo storage necessari per supportare le tue raccolte. Puoi configurare un numero massimo di OCUs per il tuo account per controllare i costi.

> **Note**
>
> Il numero di OCUs unità fornite in qualsiasi momento può variare e non è esatto. Nel tempo, l'algoritmo utilizzato da OpenSearch Serverless continuerà a migliorare per ridurre al minimo l'utilizzo del sistema. OpenSearch

Per i dettagli completi sui prezzi, consulta la pagina dei prezzi OpenSearch di Amazon Service.

Accelerazione tramite GPU e operazioni di scrittura

L'accelerazione GPU viene attivata quando la velocità di ingestione vettoriale (OpenSearchMB/sec) rientra in un intervallo. Sui OpenSearch domini, hai la flessibilità di

configurare questo intervallo tramite e. `index.knn.remote_index_build.size.min` `index.knn.remote_index_build.size.max` Ad esempio, con l'intervallo inferiore predefinito di 50 MB, la scrittura di 15.000 vettori a piena precisione con una dimensione di 768 tra gli intervalli di aggiornamento attiverà l'accelerazione GPU per impostazione predefinita.

I dati vengono scritti con le seguenti operazioni API:

- Flush
- Bulk
- Reindicizzazione
- Indice
- Aggiorna
- Elimina
- Unione forzata

L'accelerazione GPU viene attivata con l'unione automatica e manuale dei segmenti.

Configurazioni degli indici supportate

Il motore Faiss supporta l'accelerazione GPU.

Le seguenti configurazioni non supportano l'accelerazione GPU:

- Quantizzazione del prodotto Faiss
- Indice dei file invertito (IVF)
- Libreria spaziale non metrica
- Motore Lucene

Supportato Regioni AWS

L'accelerazione GPU è disponibile nelle seguenti versioni: Regioni AWS

- Stati Uniti orientali (Virginia settentrionale)
- Stati Uniti occidentali (Oregon)
- Asia Pacifico (Sydney)

- Asia Pacifico (Tokyo)
- Europa (Irlanda)

Abilitazione dell'accelerazione GPU

Puoi abilitare l'accelerazione GPU durante la creazione o l'aggiornamento di un OpenSearch dominio o di una raccolta Serverless con, o SDK. OpenSearch Console di gestione AWS AWS CLI AWS

Dopo aver abilitato l'accelerazione GPU sul dominio o sulla raccolta, questa funzionalità è abilitata per impostazione predefinita su tutti gli indici. Se devi disabilitare questa funzionalità a livello di indice, consulta. Creazione di indici vettoriali accelerati da GPU

Console

Le seguenti procedure abilitano l'accelerazione GPU per OpenSearch domini e raccolte Serverless utilizzando la console di OpenSearch gestione Serverless. OpenSearch

Create new domain

Per creare un dominio con l'accelerazione GPU abilitata, consulta OpenSearch . Creazione OpenSearch di domini di servizio

Edit existing domain

1. Aprire la console di gestione del servizio. OpenSearch
2. Nel riquadro di navigazione, scegli Domini.
3. Scegli il tuo nome di dominio per aprire la pagina dei dettagli del dominio.
4. Scegli Azioni, quindi Modifica dominio.
5. Nella sezione Funzioni avanzate, seleziona Abilita l'accelerazione GPU. Una volta abilitata questa funzionalità, le operazioni di indicizzazione vettoriale vengono accelerate.
6. Scegli Save changes (Salva modifiche).

Create new collection

Per creare una raccolta OpenSearch Serverless con accelerazione GPU abilitata, consulta. Tutorial: Guida introduttiva ad Amazon OpenSearch Serverless Durante la creazione della raccolta, assicuratevi di selezionare il tipo di raccolta di ricerca vettoriale e di abilitare l'accelerazione GPU nella configurazione della ricerca vettoriale.

Edit existing collection

1. Apri la console di gestione del servizio. OpenSearch
2. Nel riquadro di navigazione, scegli Raccolte.
3. Scegli il nome della collezione per aprire la pagina dei dettagli della collezione.
4. Nella sezione Opzioni di distribuzione, Modifica l'accelerazione GPU vettoriale.
5. Disabilita o abilita l'accelerazione GPU.
6. Scegli Save changes (Salva modifiche).

AWS CLI

Create new domain

L' AWS CLI esempio seguente crea un OpenSearch dominio con accelerazione GPU abilitata negli Stati Uniti orientali (Virginia settentrionale). Sostituiscilo *text* con quello della tua configurazione.

```
aws opensearch create-domain \
    --domain-name my-domain \
    --engine-version OpenSearch_3.1 \
    --cluster-config InstanceType=r6g.xlarge.search,\
        InstanceCount=1,\
        DedicatedMasterEnabled=true,\
        DedicatedMasterCount=3,\
        DedicatedMasterType=m6g.large.search \
    --ebs-options "EBSEnabled=true,\
        VolumeType=gp3,\
        VolumeSize=2000" \
    --encryption-at-rest-options '{"Enabled":true}' \
    --aiml-options '{"ServerlessVectorAcceleration": {"Enabled": true}}' \
    --node-to-node-encryption-options '{"Enabled":true}' \
    --domain-endpoint-options '{"EnforceHTTPS":true,\
        "TLSSecurityPolicy":"Policy-Min-TLS-1-0-2019-07"}' \
    --access-policies '{"Version": "2012-10-17",
        "Statement": [{
            "Effect": "Allow",
            "Principal": {"AWS": "*"},
            "Action": "es:*",
            "Resource": "arn:aws:es:us-east-1:123456789012:domain/my-domain/*"
        }]}' \
```

```
    --advanced-security-options '{
        "Enabled":true,
        "InternalUserDatabaseEnabled":true,
        "MasterUserOptions": {
            "MasterUserName":"USER_NAME",
            "MasterUserPassword":"PASSWORD"
        }}' \
    --region us-east-1
```

Edit existing domain

L' AWS CLI esempio seguente abilita l'accelerazione GPU per un dominio esistente. OpenSearch Sostituiscilo `text` con quello della tua configurazione.

```
aws opensearch update-domain-config \
    --domain-name my-domain \
    --cluster-config InstanceType=r7g.16xlarge.search,InstanceCount=3 \
    --aiml-options '{"ServerlessVectorAcceleration": {"Enabled": true}}'
```

Create new collection

L' AWS CLI esempio seguente crea una raccolta OpenSearch Serverless con accelerazione GPU abilitata negli Stati Uniti orientali (Virginia settentrionale). Sostituiscila `text` con quella della tua configurazione.

```
aws opensearchserverless create-collection \
    --name "my-collection" \
    --type "VECTORSEARCH" \
    --description "My vector collection with GPU acceleration" \
    --vector-options '{"ServerlessVectorAcceleration": "ENABLED"}' \
    --region us-east-1
```

Edit existing collection

L' AWS CLI esempio seguente abilita l'accelerazione GPU per una raccolta Serverless esistente. OpenSearch Sostituiscila `text` con quella della tua configurazione.

```
aws opensearchserverless update-collection \
    --id 07tjusf2h91cunochc \
    --vector-options '{"ServerlessVectorAcceleration": "ENABLED"}' \
    --region us-east-1
```

Creazione di indici vettoriali accelerati da GPU

Dopo aver abilitato l'accelerazione GPU sul tuo dominio o sulla tua raccolta, crea indici vettoriali in grado di sfruttare l'elaborazione GPU.

> **Note**
>
> Quando crei un dominio con l'accelerazione GPU abilitata, l'impostazione è predefinita. `index.knn.remote_index_build.enabled` `true` Non è necessario impostare esplicitamente questa impostazione durante la creazione degli indici. Per le raccolte, è necessario specificare in modo esplicito un valore per questa impostazione.

Creating index with GPU-acceleration

L'esempio seguente crea un indice vettoriale ottimizzato per l'elaborazione tramite GPU. Questo indice memorizza vettori di 768 dimensioni (comuni per gli incorporamenti di testo).

```
PUT my-vector-index
{
  "settings": {
    "index.knn": true,
    "index.knn.remote_index_build.enabled": true
  },
  "mappings": {
    "properties": {
      "vector_field": {
        "type": "knn_vector",
        "dimension": 768
      },
      "text": {
        "type": "text"
      }
    }
  }
}
```

Elementi chiave di configurazione:

- `"index.knn": true`- Abilita la funzionalità k-nearest neighbor

- `"index.knn.remote_index_build.enabled": true`- Abilita l'elaborazione tramite GPU per questo indice. Quando il dominio ha l'accelerazione GPU abilitata, se non si specifica questa impostazione, l'impostazione predefinita è. `true` Per le raccolte, è necessario specificare esplicitamente un valore per questa impostazione.
- `"dimension": 768`- Speciifica la dimensione del vettore (regola in base al modello di incorporamento)

Creating index without GPU-acceleration

L'esempio seguente crea un indice vettoriale in cui l'elaborazione della GPU è disabilitata. Questo indice memorizza vettori di 768 dimensioni (comuni per gli incorporamenti di testo).

```
PUT my-vector-index
{
  "settings": {
    "index.knn": true,
    "index.knn.remote_index_build.enabled": false
  },
  "mappings": {
    "properties": {
      "vector_field": {
        "type": "knn_vector",
        "dimension": 768
      },
      "text": {
        "type": "text"
      }
    }
  }
}
```

Indicizzazione dei dati vettoriali e unione forzata

Dopo aver creato un indice vettoriale accelerato da GPU sul dominio o sulla raccolta, puoi aggiungere dati vettoriali e ottimizzare l'indice utilizzando operazioni standard. OpenSearch L'accelerazione GPU migliora automaticamente sia le prestazioni di indicizzazione che le operazioni di unione forzata, velocizzando la creazione e la manutenzione di applicazioni di ricerca vettoriale su larga scala senza richiedere modifiche ai flussi di lavoro esistenti.

Indicizzazione dei dati vettoriali

Indicizza i dati vettoriali come faresti normalmente. L'accelerazione GPU si applica automaticamente alle operazioni di indicizzazione e unione forzata. L'esempio seguente mostra come aggiungere documenti vettoriali all'indice utilizzando l'API bulk. Ogni documento contiene un campo vettoriale con valori numerici e contenuto testuale associato:

```
POST _bulk
{"index": {"_index": "my-vector-index"}}
{"vector_field": [0.1, 0.2, 0.3, ...], "text": "Sample document 1"}
{"index": {"_index": "my-vector-index"}}
{"vector_field": [0.4, 0.5, 0.6, ...], "text": "Sample document 2"}
```

Operazioni di unione forzata

L'accelerazione GPU si applica anche alle operazioni di unione forzata, che possono ridurre significativamente il tempo necessario per ottimizzare gli indici vettoriali. Tieni presente che le operazioni di unione forzata non sono supportate nelle raccolte. L'esempio seguente dimostra come ottimizzare l'indice vettoriale consolidando tutti i segmenti in un unico segmento:

```
POST my-vector-index/_forcemerge?max_num_segments=1
```

Best practice

Segui queste best practice per massimizzare i vantaggi dell'accelerazione GPU per i tuoi carichi di lavoro di ricerca vettoriale:

- Aumenta i client di indicizzazione: per sfruttare appieno i vantaggi della creazione dell' GPUs indice, aumenta il numero di client di indicizzazione in cui vengono importati i dati. OpenSearch Ciò consente una migliore parallelizzazione e utilizzo delle risorse GPU.
- Regola la soglia approssimativa: modifica l'`index.knn.advanced.approximate_threshold`impostazione per garantire che non si verifichino aumenti di indici di segmenti più piccoli, il che migliora la velocità complessiva di inserimento. Un valore di 10.000 è un buon punto di partenza. Per le raccolte, è necessario specificare esplicitamente un valore per questa impostazione.
- Ottimizza le dimensioni dei frammenti: prova a creare frammenti contenenti almeno 1 milione di documenti. Gli shard con un numero di documenti inferiore a questo numero potrebbero non trarre benefici complessivi dall'accelerazione tramite GPU.

Utilizzo dei OpenSearch pannelli di controllo con Amazon Service OpenSearch

OpenSearch Dashboards è uno strumento di visualizzazione open source progettato per funzionare con. OpenSearch Amazon OpenSearch Service fornisce un'installazione di dashboard con ogni dominio OpenSearch di servizio. Le dashboard vengono eseguite sugli hot data node del dominio.

OpenSearch Dashboards è uno strumento di visualizzazione per esplorare e analizzare i dati all'interno di un singolo dominio. OpenSearch Al contrario, l'interfaccia OpenSearch utente centralizzata (chiamata anche OpenSearch applicazione) è un'interfaccia utente basata sul cloud che si collega a più OpenSearch domini, OpenSearch raccolte Serverless e fonti di dati. AWS Include aree di lavoro per casi d'uso specifici come l'osservabilità e l'analisi della sicurezza e offre un'esperienza unificata tra i set di dati. Sebbene Dashboards sia legato a singoli domini, l'interfaccia utente centralizzata consente l'integrazione e l'analisi dei dati tra domini. Per ulteriori informazioni, consulta OpenSearch Interfaccia utente.

Puoi trovare un link a OpenSearch Dashboards nella dashboard del tuo dominio nella console di servizio. OpenSearch Per i domini in esecuzione OpenSearch, l'URL è. *domain-endpoint*`/_dashboards/` Per i domini che eseguono la versione precedente di Elasticsearch, l'URL è. *domain-endpoint*`/_plugin/kibana`

Le query che utilizzano questa installazione predefinita di Dashboards hanno un timeout di 300 secondi.

> **Note**
>
> Questa documentazione descrive le OpenSearch dashboard nel contesto di Amazon OpenSearch Service, inclusi i diversi modi per connettersi ad esso. Per una documentazione completa, tra cui una guida introduttiva, istruzioni per creare una dashboard, la gestione delle dashboard e Dashboards Query Language (DQL), consulta OpenSearch Dashboards nella documentazione open source. OpenSearch

Controllo dell'accesso ai pannelli di controllo

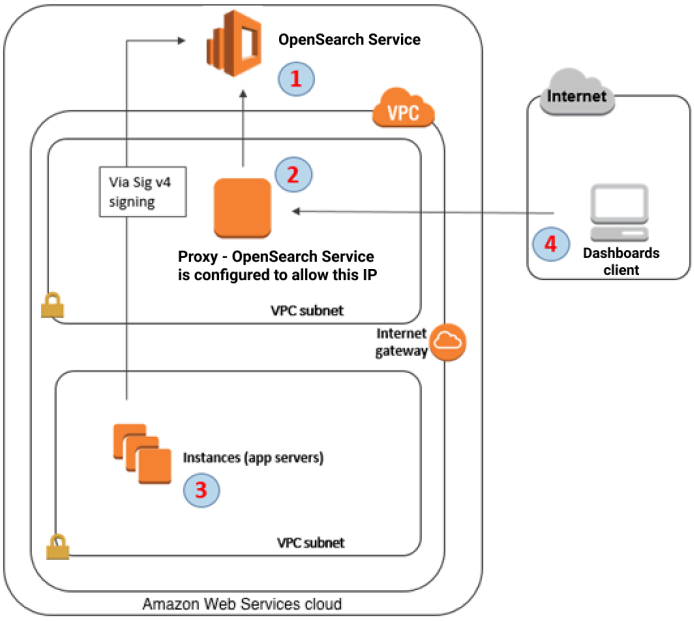
Dashboards non supporta nativamente gli utenti e i ruoli IAM, ma OpenSearch Service offre diverse soluzioni per controllare l'accesso alle dashboard:

- Abilitare l'autenticazione SAML per Dashboards.
- Utilizzare il controllo granulare degli accessi con l'autenticazione di base HTTP.
- Configura l'autenticazione Cognito per Dashboards.
- Per i domini ad accesso pubblico, configurare una policy di accesso basata su IP, che utilizzi o meno un server proxy.
- Per i domini con accesso VPC, utilizzare una policy di accesso aperta, che utilizzi o meno un server proxy, e gruppi di sicurezza per controllare l'accesso. Per ulteriori informazioni, consulta the section called “Informazioni sulle policy d'accesso nei domini VPC”.

Utilizzo di un proxy per accedere al servizio da dashboard OpenSearch

> **Note**
>
> Questo processo è valido solo se il dominio usa l'accesso pubblico e non vuoi usare l'autenticazione Cognito. Per informazioni, consulta the section called “Controllo dell'accesso ai pannelli di controllo”.

Poiché Dashboards è un' JavaScript applicazione, le richieste provengono dall'indirizzo IP dell'utente. Il controllo degli accessi basato su IP può essere poco pratico a causa del numero elevato di indirizzi IP che sarebbe necessario includere nella whitelist perché ogni utente possa accedere a Dashboards. Una soluzione alternativa consiste nel posizionare un server proxy tra Dashboards e Service. OpenSearch È quindi possibile aggiungere una policy d'accesso basata su IP che accetta le richieste provenienti da un solo indirizzo IP, quello del proxy. Il seguente diagramma mostra questa configurazione.



1. Questo è il tuo dominio di OpenSearch servizio. IAM offre accesso autorizzato al dominio. Un'ulteriore policy d'accesso basata su IP permette l'accesso al server proxy.
2. Questo è il server proxy, in esecuzione su un'istanza Amazon EC2.
3. Altre applicazioni possono utilizzare la procedura di firma Signature Version 4 per inviare richieste autenticate al OpenSearch Servizio.
4. I client Dashboards si connettono al dominio OpenSearch di servizio dell'utente tramite il proxy.

Per abilitare questo tipo di configurazione, è necessaria una policy basata su risorse che specifica i ruoli e gli indirizzi IP. Ecco una policy di esempio:

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Resource":"arn:aws:es:us-west-2:111111111111:domain/my-domain/*",
         "Principal":{
            "AWS":"arn:aws:iam::111111111111:role/allowedrole1"
         },
         "Action":[
            "es:ESHttpGet"
         ],
```

```
            "Effect":"Allow"
         },
         {
            "Effect":"Allow",
            "Principal":{
               "AWS":"*"
            },
            "Action":"es:*",
            "Condition":{
               "IpAddress":{
                  "aws:SourceIp":[
                     "203.0.113.0/24",
                     "2001:DB8:1234:5678::/64"
                  ]
               }
            },
            "Resource":"arn:aws:es:us-west-2:111111111111:domain/my-domain/*"
         }
      ]
}
```

Ti consigliamo di configurare l'istanza EC2 che esegue il server proxy con un indirizzo IP elastico (EIP). In questo modo, puoi sostituire l'istanza quando necessario e collegarvi comunque lo stesso indirizzo IP pubblico. Per ulteriori informazioni, consulta Elastic IP Addresses nella Amazon EC2 User Guide.

Se si utilizza un server proxy e l'autenticazione Cognito, potrebbe essere necessario dover aggiungere impostazioni per Dashboards e Amazon Cognito in modo da evitare errori `redirect_mismatch`. Fai riferimento al file `nginx.conf` di esempio seguente:

```
server {
    listen 443;
    server_name $host;
    rewrite ^/$ https://$host/_plugin/_dashboards redirect;

    ssl_certificate          /etc/nginx/cert.crt;
    ssl_certificate_key      /etc/nginx/cert.key;

    ssl on;
    ssl_session_cache  builtin:1000  shared:SSL:10m;
    ssl_protocols  TLSv1 TLSv1.1 TLSv1.2;
    ssl_ciphers HIGH:!aNULL:!eNULL:!EXPORT:!CAMELLIA:!DES:!MD5:!PSK:!RC4;
```

```
    ssl_prefer_server_ciphers on;

    location /_plugin/_dashboards {
        # Forward requests to Dashboards
        proxy_pass https://$dashboards_host/_plugin/_dashboards;

        # Handle redirects to Cognito
        proxy_redirect https://$cognito_host https://$host;

        # Update cookie domain and path
        proxy_cookie_domain $dashboards_host $host;
        proxy_cookie_path / /_plugin/_dashboards/;

        # Response buffer settings
        proxy_buffer_size 128k;
        proxy_buffers 4 256k;
        proxy_busy_buffers_size 256k;
    }

    location ~ \/(log|sign|fav|forgot|change|saml|oauth2) {
        # Forward requests to Cognito
        proxy_pass https://$cognito_host;

        # Handle redirects to Dashboards
        proxy_redirect https://$dashboards_host https://$host;

        # Update cookie domain
        proxy_cookie_domain $cognito_host $host;
    }
}
```

> **Note**
>
> (Facoltativo) Se scegli di fornire un nodo coordinatore dedicato, questo inizierà automaticamente a ospitare OpenSearch le dashboard. Di conseguenza, aumenta la disponibilità delle risorse dei nodi dati come CPU e memoria. Questa maggiore disponibilità di risorse dei nodi di dati può contribuire a migliorare la resilienza complessiva del dominio.

Configurazione delle dashboard per l'utilizzo di un server di mappe WMS

L'installazione predefinita di Dashboards for OpenSearch Service include un servizio di mappe, ad eccezione dei domini nelle regioni di India e Cina. Il servizio di mappe supporta fino a 10 livelli di zoom.

Indipendentemente dalla regione, è possibile configurare Dashboards per l'utilizzo di un server WMS (Web Map Service) diverso per le visualizzazioni delle mappe delle coordinate. Le visualizzazioni delle mappe regionali supportano solo il servizio mappe predefinito.

Per configurare Dashboards per l'uso di un server di mappe WMS:

1. Aprire Dashboards.
2. Scegliere Gestione stack.
3. Scegliere Advanced Settings (Impostazioni avanzate).
4. Individua la visualizzazione: TileMap:. WMSdefaults
5. Modifica `enabled` in `true` e `url` nell'URL di un server di mappe WMS valido:

```
{
  "enabled": true,
  "url": "wms-server-url",
  "options": {
    "format": "image/png",
    "transparent": true
  }
}
```

6. Scegli Save changes (Salva modifiche).

Per applicare il nuovo valore di default alle visualizzazioni, potrebbe essere necessario ricaricare Dashboards. Se sono state salvate le visualizzazioni, scegli Opzioni dopo aver aperto la visualizzazione. Verifica che il server di mappe WMS sia abilitato e l'URL WMS contenga il server di mappe preferito, quindi scegli Applica modifiche.

> **Note**
>
> I servizi mappe spesso prevedono costi di licenza o limitazioni. Sei responsabile di tutti questi aspetti per qualsiasi server di mappe che specifichi. Potresti trovare utile provare i servizi mappe di U.S. Geological Survey.

Connessione di un server Dashboards locale al servizio OpenSearch

Se hai già investito molto tempo nella configurazione della tua istanza Dashboards, puoi utilizzarla al posto (o in aggiunta) all'istanza Dashboards predefinita fornita dal Servizio. OpenSearch La procedura seguente funziona per i domini che utilizzano il controllo granulare degli accessi con una policy di accesso aperto.

Per connettere un server Dashboards locale al servizio OpenSearch

1. Nel tuo dominio OpenSearch di servizio, crea un utente con le autorizzazioni appropriate:
   a. In Dashboards, passare a Sicurezza, Utenti interni, quindi scegliere Crea un utente interno.
   b. Fornire nome utente e password e scegliere Crea.
   c. Passare a Ruoli e selezionare un ruolo.
   d. Selezionare Utenti mappati e scegliere Gestisci mappatura.
   e. In Utenti, aggiungere il nome utente e scegliere Mappa.
2. Scarica e installa la versione appropriata del plug-in di OpenSearch sicurezza sull'installazione OSS di Dashboards autogestita.
3. Sul server Dashboards locale, apri il `config/opensearch_dashboards.yml` file e aggiungi l'endpoint di OpenSearch servizio con il nome utente e la password che hai creato in precedenza:

   ```
   opensearch.hosts: ['https://domain-endpoint']
   opensearch.username: 'username'
   opensearch.password: 'password'
   ```

   È possibile utilizzare il seguente file `opensearch_dashboards.yml` di esempio:

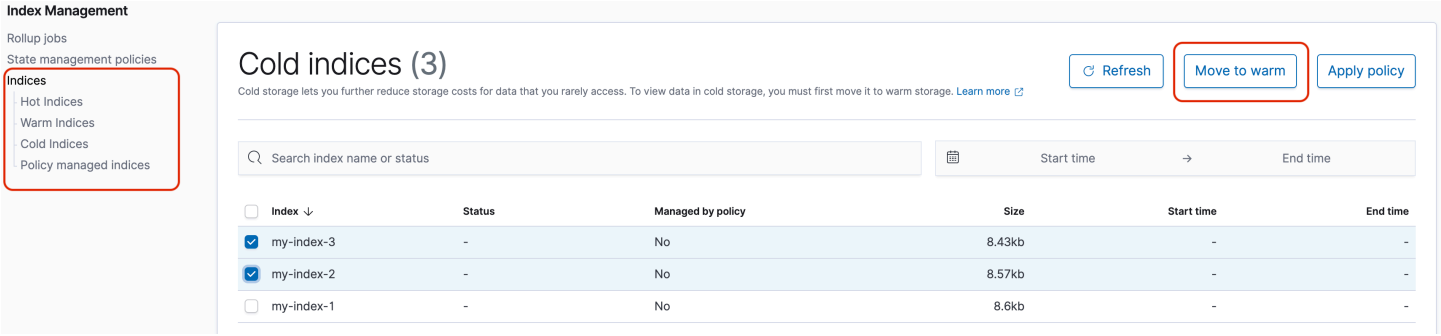
```
server.host: '0.0.0.0'

opensearch.hosts: ['https://domain-endpoint']

opensearchDashboards.index: ".username"

opensearch.ssl.verificationMode: none # if not using HTTPS

opensearch_security.auth.type: basicauth
opensearch_security.auth.anonymous_auth_enabled: false
opensearch_security.cookie.secure: false # set to true when using HTTPS
opensearch_security.cookie.ttl: 3600000
opensearch_security.session.ttl: 3600000
opensearch_security.session.keepalive: false
opensearch_security.multitenancy.enabled: false
opensearch_security.readonly_mode.roles: ['opensearch_dashboards_read_only']
opensearch_security.auth.unauthenticated_routes: []
opensearch_security.basicauth.login.title: 'Please log in using your username and
 password'

opensearch.username: 'username'
opensearch.password: 'password'
opensearch.requestHeadersWhitelist: [authorization, securitytenant,
 security_tenant]
```

Per visualizzare gli indici OpenSearch del servizio, avvia il server Dashboards locale, vai su Dev Tools ed esegui il seguente comando:

```
GET _cat/indices
```

Gestione degli indici nelle dashboard

L'installazione di Dashboards sul dominio di OpenSearch servizio fornisce un'interfaccia utente utile per la gestione degli indici in diversi livelli di storage del dominio. Scegliete Gestione degli indici dal menu principale delle dashboard per visualizzare tutti gli indici nelle aree di archiviazione a caldo e a freddo UltraWarm, nonché gli indici gestiti dalle politiche di Index State Management (ISM). Utilizza la gestione degli indici per spostarli tra l'archiviazione a caldo e a freddo e per monitorare le migrazioni tra i tre livelli.



Tieni presente che non vedrai le opzioni degli indici caldi, caldi e freddi a meno che tu non abbia abilitato e/o abilitato la conservazione a freddo. UltraWarm

Funzionalità aggiuntive

L'installazione predefinita di Dashboards in ogni dominio OpenSearch di servizio presenta alcune funzionalità aggiuntive:

- Interfacce utente per i vari plugin OpenSearch
- Tenant
- Report

  Utilizzare il menu Creazione di report per generare report CSV on demand dalla pagina Individua e report PDF o PNG di pannelli di controllo o visualizzazioni. I report CSV hanno un limite di 10.000 righe.
- Grafici Gantt
- Notebook

Utilizzo OpenSearch dell'interfaccia utente in Amazon OpenSearch Service

OpenSearch L'interfaccia utente (interfaccia utente) è un'esperienza di analisi operativa modernizzata per Amazon OpenSearch Service che offre una vista unificata per interagire con i dati da più fonti. A differenza di OpenSearch Dashboards, che funziona solo con un dominio o una raccolta che lo ospita, l' OpenSearch interfaccia utente è ospitata in. Cloud AWS Ciò consente all' OpenSearch interfaccia utente di raggiungere un'elevata disponibilità e di rimanere funzionale durante gli aggiornamenti del cluster e di connettersi in modo nativo a più fonti di dati. Per informazioni sui OpenSearch pannelli di controllo, consulta. OpenSearch Pannelli di controllo

Le seguenti sono le caratteristiche principali dell' OpenSearch interfaccia utente:

- Supporto per più fonti di dati: l' OpenSearch interfaccia utente può connettersi a più fonti di dati per creare una visualizzazione completa. Ciò include OpenSearch domini e raccolte serverless, nonché fonti di AWS dati integrate come Amazon, Amazon Security Lake e Amazon Simple Storage Service (CloudWatchAmazon S3) Simple Storage Service (Amazon S3).
- Zero tempi di inattività durante gli aggiornamenti: l' OpenSearch interfaccia utente è ospitata in. Cloud AWS Ciò significa che OpenSearch rimane operativo e può recuperare i dati dai cluster durante i processi di aggiornamento.
- Spazi di lavoro: spazi curati per la collaborazione in team per vari flussi di lavoro, come Observability, Security Analytics e Search. Puoi definire le impostazioni sulla privacy e gestire le autorizzazioni per i collaboratori nel tuo spazio di lavoro.
- Single Sign-On: l' OpenSearch interfaccia utente funziona con la federazione SAML tramite AWS Identity and Access Management (IAM) per integrarsi con i tuoi provider di identità AWS IAM Identity Center e creare un'esperienza Single Sign-On per gli utenti finali.
- Analisi basata su GenAI: l' OpenSearch interfaccia utente supporta la generazione di query in linguaggio naturale per aiutarti a generare le query giuste per le tue analisi. OpenSearch L'interfaccia utente collabora anche con Amazon Q Developer per fornire la chat di Amazon Q e aiutare a generare visualizzazioni, riepilogo degli avvisi, approfondimenti e rilevatori di anomalie consigliati.
- Supporto per più linguaggi di interrogazione: l' OpenSearch interfaccia utente supporta Piped Processing Language (PPL), SQL, Lucene e Dashboards Query Language (DQL).

- Supporto per più regioni e più account: l' OpenSearch interfaccia utente può utilizzare la funzionalità di ricerca tra cluster per connettersi a OpenSearch domini di account diversi e regioni diverse per analisi e visualizzazioni aggregate.

Per iniziare a creare la tua prima interfaccia utente, segui le istruzioni riportate in. OpenSearch the section called “Nozioni di base”

Per informazioni sulle funzionalità più recenti rilasciate per l' OpenSearch interfaccia utente, consultathe section called “Cronologia delle versioni”.

Argomenti

- Cronologia delle versioni dell'interfaccia utente di Amazon OpenSearch Service
- Guida introduttiva all'interfaccia OpenSearch utente in Amazon OpenSearch Service
- AI agentica in Amazon Service OpenSearch
- Crittografia dei metadati delle applicazioni dell' OpenSearch interfaccia utente con chiavi gestite dal cliente
- Abilitazione della federazione SAML con AWS Identity and Access Management
- Gestione delle associazioni delle fonti di dati e delle autorizzazioni di accesso al Virtual Private Cloud
- Utilizzo delle aree di lavoro OpenSearch di Amazon Service
- Accesso ai dati tra regioni e account diversi
- Gestione dell'accesso all' OpenSearch interfaccia utente da un endpoint VPC
- Migrazione degli oggetti salvati dalle dashboard all'interfaccia utente OpenSearch OpenSearch
- OpenSearch Endpoint e quote dell'interfaccia utente

Cronologia delle versioni dell'interfaccia utente di Amazon OpenSearch Service

La tabella seguente elenca tutte le versioni del supporto Amazon OpenSearch Service per l' OpenSearch interfaccia utente e le funzionalità e i miglioramenti inclusi in ciascuna versione.

Modifica	Data di rilascio	Description
Nuova funzionalità	10/03/2025	OpenSearch L'interfaccia utente ora supporta l'accesso ai dati tra account diversi ai domini. OpenSearch Puoi configurare le applicazioni in un account per accedere ai domini di account diversi utilizzando IAM o l' AWS IAM Identity Center autenticazione, sia per i domini pubblici che per quelli VPC. Per ulteriori informazioni, consulta the section called “Accesso ai dati di più account ai domini OpenSearch ”.
Nuova funzionalità	2025-12-29	OpenSearch L'interfaccia utente ora supporta le chiavi gestite dai clienti (CMK) per la crittografia dei metadati delle applicazioni. Questa funzionalità consente di soddisfare i requisiti normativi e di conformità offrendoti il pieno controllo sulle chiavi di crittografia. Questo lancio aumenta anche il limite di dimensione dei metadati per gli oggetti salvati nell' OpenSearch interfaccia utente, consentendoti di creare e archiviare query complesse, visualizzazioni estese e aree di lavoro su larga scala. Per ulteriori informazioni, consulta the section called “Crittografia e chiave gestita dal cliente”.
Nuova funzionalità	2025-04-16	OpenSearch L'interfaccia utente ora funziona con la ricerca tra cluster. In questo modo è possibile utilizzare l' OpenSearch interfaccia utente in un'unica interfaccia Regione AWS per accedere ai cluster in una regione diversa. Questo viene fatto configurandolo come un cluster remoto connesso a un cluster nella stessa regione. Per ulteriori informazioni, consulta the section called “Funzionalità di ricerca tra cluster”.
Nuova funzionalità	2025-03-31	Amazon Q Developer è ora disponibile a livello generale in Amazon OpenSearch Service. Per informazioni, consulta Assistente AI.
Nuova funzionalità	05/02/2025	La federazione da Security Assertion Markup Language 2.0) (SAML) a (IAM) ora funziona con l'interfaccia AWS Identity and Access Management utente. OpenSearch In questo modo è possibile creare un'esperienza Single Sign-On (SSO)

Modifica	Data di rilascio	Description
		avviata dal provider di identità per gli utenti finali. Per ulteriori informazioni, consulta the section called “Abilitazione della federazione SAML con IAM”.
Nuova integrazione	2024-12-01	L'integrazione zero-ETL con Amazon CloudWatch semplifica l'analisi e la visualizzazione dei dati di log, riducendo le spese tecniche generali e i costi operativi. Per ulteriori informazioni, consulta New Amazon CloudWatch e Amazon OpenSearch Service lanciano un'esperienza di analisi integrata sul AWS News Blog.
Nuova integrazione	2024-12-01	L'integrazione zero-ETL con Amazon Security Lake consente alle organizzazioni di cercare, analizzare e ottenere informazioni utili dai propri dati di sicurezza in modo efficiente. Per ulteriori informazioni, consulta la sezione Introduzione all'integrazione tra Amazon OpenSearch Service e Amazon Security Lake per semplificare l'analisi della sicurezza nel AWS News Blog.
Rilascio iniziale	2024-11-07	La versione pubblica iniziale dell'interfaccia utente. OpenSearch Per ulteriori informazioni, consulta Amazon OpenSearch Service lancia l' OpenSearch interfaccia utente di nuova generazione sul AWS Big Data Blog.

Guida introduttiva all'interfaccia OpenSearch utente in Amazon OpenSearch Service

In Amazon OpenSearch Service, un'applicazione è un'istanza dell'interfaccia OpenSearch utente (OpenSearch UI). Ogni applicazione può essere associata a più fonti di dati e una singola fonte può essere associata a più applicazioni. È possibile creare più applicazioni per amministratori diversi utilizzando diverse opzioni di autenticazione supportate.

Utilizzate le informazioni contenute in questo argomento per guidarvi nel processo di creazione di un'applicazione OpenSearch UI utilizzando Console di gestione AWS o il AWS CLI.

Argomenti

- Autorizzazioni richieste per la creazione di applicazioni Amazon OpenSearch Service
- Creazione di un'applicazione OpenSearch UI
- Gestione degli amministratori delle applicazioni

Autorizzazioni richieste per la creazione di applicazioni Amazon OpenSearch Service

Prima di creare un'applicazione, verifica che ti siano state concesse le autorizzazioni necessarie per l'attività. Se necessario, contatta un amministratore dell'account per ricevere assistenza.

Autorizzazioni generali

Per utilizzare le applicazioni in OpenSearch Service, sono necessarie le autorizzazioni indicate nella seguente politica. Le autorizzazioni servono ai seguenti scopi:

- Le cinque `es:*Application` autorizzazioni sono necessarie per creare e gestire un'applicazione.
- Le tre `es:*Tags` autorizzazioni sono necessarie per aggiungere, elencare e rimuovere tag dall'applicazione.
- Le `aoss:BatchGetCollection` `es:GetDirectQueryDataSource` autorizzazioni `es:DescribeDomain` e sono necessarie per associare le fonti di dati.
- Sono necessarie le `opensearch:*DirectQuery*` autorizzazioni `aoss:APIAccessAlles:ESHttp*`, e 4 per accedere alle fonti di dati.
- `iam:CreateServiceLinkedRole`Fornisce l'autorizzazione ad Amazon OpenSearch Service per creare un ruolo collegato al servizio (SLR) nel tuo account. Questo ruolo viene utilizzato e consente all'applicazione OpenSearch UI di pubblicare i CloudWatch parametri Amazon nel tuo account. Per ulteriori informazioni, consulta the section called “Permissions” nell'argomento Utilizzo di ruoli collegati ai servizi per creare domini VPC e fonti di dati di interrogazione diretta.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "VisualEditor0",
```

```
            "Effect": "Allow",
            "Action": [
                "es:CreateApplication",
                "es:DeleteApplication",
                "es:GetApplication",
                "es:ListApplications",
                "es:UpdateApplication",
                "es:AddTags",
                "es:ListTags",
                "es:RemoveTags",
                "aoss:APIAccessAll",
                "es:ESHttp*",
                "opensearch:StartDirectQuery",
                "opensearch:GetDirectQuery",
                "opensearch:CancelDirectQuery",
                "opensearch:GetDirectQueryResult",
                "aoss:BatchGetCollection",
                "aoss:ListCollections",
                "es:DescribeDomain",
                "es:DescribeDomains",
                "es:ListDomainNames",
                "es:GetDirectQueryDataSource",
                "es:ListDirectQueryDataSources"
            ],
            "Resource": "*"
        },
        {
            "Sid": "VisualEditor1",
            "Effect": "Allow",
            "Action": "iam:CreateServiceLinkedRole",
            "Resource": "arn:aws:iam::*:role/aws-service-role/
opensearchservice.amazonaws.com/AWSServiceRoleForAmazonOpenSearchService"
        }
    ]
}
```

Autorizzazioni per creare un'applicazione che utilizza l'autenticazione IAM Identity Center (opzionale)

Per impostazione predefinita, le applicazioni dashboard vengono autenticate utilizzando AWS Identity and Access Management (IAM) per gestire le autorizzazioni per AWS gli utenti delle risorse. Tuttavia, puoi scegliere di fornire un'esperienza Single Sign-On utilizzando IAM Identity Center, che

consente di utilizzare i provider di identità esistenti per accedere alle applicazioni dell'interfaccia utente. OpenSearch In questo caso, selezionerai l'opzione Autenticazione con IAM Identity Center nella procedura riportata più avanti in questo argomento, quindi concederai agli utenti di IAM Identity Center le autorizzazioni necessarie per accedere all' OpenSearch applicazione UI.)

Per creare un'applicazione che utilizza l'autenticazione IAM Identity Center, avrai bisogno delle seguenti autorizzazioni. Sostituisci *`placeholder values`* con le informazioni appropriate. Contatta un amministratore dell'account per ricevere assistenza, se necessario.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "IDCPermissions",
            "Effect": "Allow",
            "Action": [
                "es:CreateApplication",
                "es:DeleteApplication",
                "es:GetApplication",
                "es:ListApplications",
                "es:UpdateApplication",
                "es:AddTags",
                "es:ListTags",
                "es:RemoveTags",
                "aoss:BatchGetCollection",
                "aoss:ListCollections",
                "es:DescribeDomain",
                "es:DescribeDomains",
                "es:ListDomainNames",
                "es:GetDirectQueryDataSource",
                "es:ListDirectQueryDataSources",
                "sso:CreateApplication",
                "sso:DeleteApplication",
                "sso:PutApplicationGrant",
                "sso:PutApplicationAccessScope",
                "sso:PutApplicationAuthenticationMethod",
                "sso:ListInstances",
                "sso:DescribeApplicationAssignment",
                "sso:DescribeApplication",
                "sso:CreateApplicationAssignment",
```

```
                "sso:ListApplicationAssignments",
                "sso:DeleteApplicationAssignment",
                "sso-directory:SearchGroups",
                "sso-directory:SearchUsers",
                "sso:ListDirectoryAssociations",
                "identitystore:DescribeUser",
                "identitystore:DescribeGroup",
                "iam:ListRoles"
            ],
            "Resource": "*"
        },
        {
            "Sid": "SLRPermission",
            "Effect": "Allow",
            "Action": "iam:CreateServiceLinkedRole",
            "Resource": "arn:aws:iam::*:role/aws-service-role/
opensearchservice.amazonaws.com/AWSServiceRoleForAmazonOpenSearchService"
        },
        {
            "Sid": "PassRolePermission",
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": "arn:aws:iam::111122223333:role/iam-role-for-identity-
center"
        }
    ]
}
```

Creazione di un'applicazione OpenSearch UI

Create un'applicazione che specifichi il nome dell'applicazione, il metodo di autenticazione e gli amministratori utilizzando una delle seguenti procedure.

Argomenti

- Creazione di un'applicazione OpenSearch UI che utilizza l'autenticazione IAM nella console
- Creazione di un'applicazione OpenSearch UI che utilizza AWS IAM Identity Center l'autenticazione nella console
- Creazione di un'applicazione OpenSearch UI che utilizza AWS IAM Identity Center l'autenticazione tramite AWS CLI

Creazione di un'applicazione OpenSearch UI che utilizza l'autenticazione IAM nella console

Per creare un'applicazione OpenSearch UI che utilizzi l'autenticazione IAM nella console

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli OpenSearch UI (dashboard).
3. Scegli Crea applicazione.
4. Per Nome applicazione, inserisci un nome per l'applicazione.
5. Non selezionare la casella di controllo Autenticazione con IAM Identity Center. Per informazioni sulla creazione di un'applicazione con autenticazione tramite AWS IAM Identity Center, consulta the section called “Creazione di un'applicazione OpenSearch UI che utilizza AWS IAM Identity Center l'autenticazione nella console” più avanti in questo argomento.
6. (Facoltativo) L'utente viene aggiunto automaticamente come amministratore dell'applicazione che si sta creando. Nell'area di gestione degli amministratori OpenSearch dell'applicazione, puoi concedere le autorizzazioni di amministratore ad altri utenti.

   > **Note**
   >
   > Il ruolo di amministratore dell'applicazione OpenSearch UI concede le autorizzazioni per modificare ed eliminare un'applicazione UI. OpenSearch Gli amministratori dell'applicazione possono anche creare, modificare ed eliminare aree di lavoro in un'applicazione UI. OpenSearch

   Per concedere le autorizzazioni di amministratore ad altri utenti, scegliete una delle seguenti opzioni:

   - Concedi l'autorizzazione di amministratore a utenti specifici: nel campo amministratori dell'OpenSearchapplicazione, nell'elenco pop-up Proprietà, seleziona utenti IAM o

     AWS IAM Identity Center utenti, quindi scegli i singoli utenti a cui concedere le autorizzazioni di amministratore.
   - Concedi l'autorizzazione di amministratore a tutti gli utenti: a tutti gli utenti dell'organizzazione o dell'account vengono concesse le autorizzazioni di amministratore.

7. (Facoltativo) Configura le impostazioni di crittografia. Per impostazione predefinita, OpenSearch i metadati dell'interfaccia utente sono crittografati con chiavi AWS proprietarie. Per utilizzare la propria chiave gestita dal cliente (CMK) per la crittografia, consulta. the section called “Crittografia e chiave gestita dal cliente”
8. (Facoltativo) Nell'area Tag, applicate una o più name/value coppie di chiavi di tag all'applicazione.

   I tag sono metadati facoltativi assegnati a una risorsa. Consentono di categorizzare una risorsa in diversi modi, ad esempio in base allo scopo, al proprietario o all'ambiente.
9. Scegli Create (Crea).

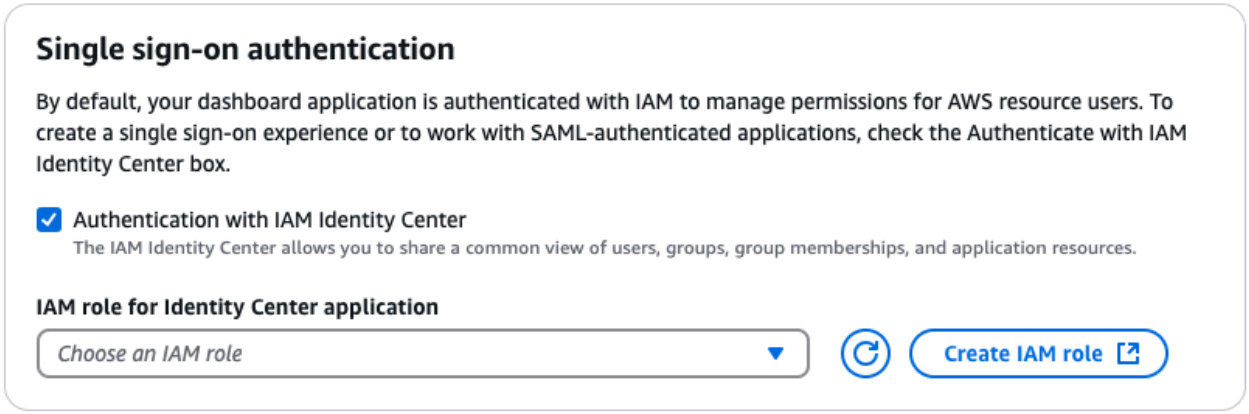
Creazione di un'applicazione OpenSearch UI che utilizza AWS IAM Identity Center l'autenticazione nella console

Per creare un'applicazione OpenSearch UI che utilizzi l' AWS IAM Identity Center autenticazione, è necessario disporre delle autorizzazioni IAM descritte in precedenza in questo argomento inthe section called “Autorizzazioni per creare un'applicazione che utilizza l'autenticazione IAM Identity Center (opzionale)”.

Per creare un'applicazione OpenSearch UI che utilizzi AWS IAM Identity Center l'autenticazione nella console

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli OpenSearch UI (dashboard).
3. Scegli Crea applicazione.
4. Per Nome applicazione, inserisci un nome per l'applicazione.
5. (Facoltativo) Per abilitare il Single Sign-On per la tua organizzazione o il tuo account, procedi come segue:

   a. Seleziona la casella di controllo Autenticazione con IAM Identity Center, come mostrato nell'immagine seguente:



b. Esegui una delle seguenti operazioni:

- Nell'elenco delle applicazioni IAM role for Identity Center, scegli un ruolo IAM esistente che fornisca le autorizzazioni richieste a IAM Identity Center per accedere all' OpenSearch interfaccia utente e alle fonti di dati associate. Vedi le politiche nel prossimo bullet per le autorizzazioni che il ruolo deve avere.
- Crea un nuovo ruolo con le autorizzazioni richieste. Utilizza le seguenti procedure nella Guida per l'utente IAM con le opzioni specificate per creare un nuovo ruolo e con la politica di autorizzazione e la politica di fiducia necessarie.
  - Procedura: creazione di politiche IAM (console)

    Mentre segui i passaggi di questa procedura, incolla la seguente policy nel campo JSON dell'editor di policy:

    JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "IdentityStoreOpenSearchDomainConnectivity",
            "Effect": "Allow",
            "Action": [
                "identitystore:DescribeUser",
                "identitystore:ListGroupMembershipsForMember",
                "identitystore:DescribeGroup"
            ],
            "Resource": "*",
            "Condition": {
                "ForAnyValue:StringEquals": {
```

```
                    "aws:CalledViaLast": "es.amazonaws.com"
                }
            }
        },
        {
            "Sid": "OpenSearchDomain",
            "Effect": "Allow",
            "Action": [
                "es:ESHttp*"
            ],
            "Resource": "*"
        },
        {
            "Sid": "OpenSearchServerless",
            "Effect": "Allow",
            "Action": [
                "aoss:APIAccessAll"
            ],
            "Resource": "*"
        }
    ]
}
```

- Procedura: creare un ruolo utilizzando politiche di attendibilità personalizzate

  Man mano che segui i passaggi di questa procedura, sostituisci il segnaposto JSON nella casella Custom trust policy con quanto segue:

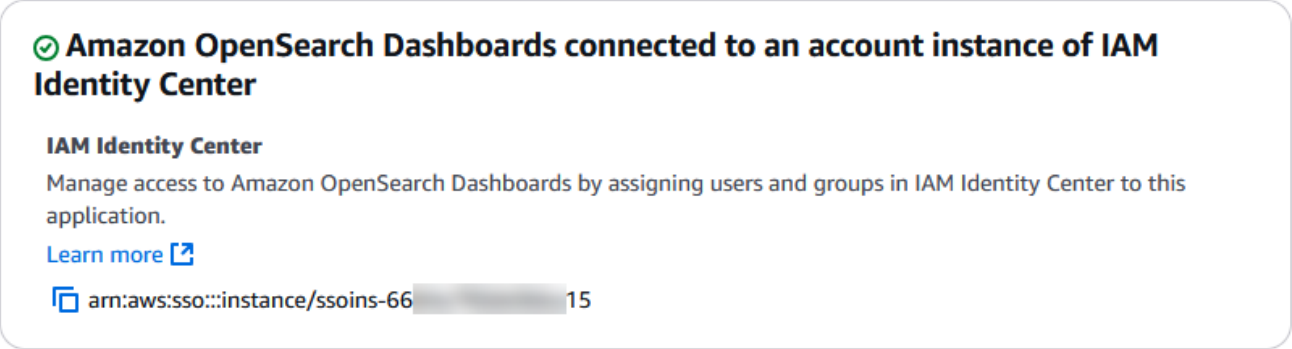
  > **Tip**
  >
  > Se stai aggiungendo la politica di fiducia a un ruolo esistente, aggiungi la politica nella scheda Relazione di fiducia del ruolo.

  JSON

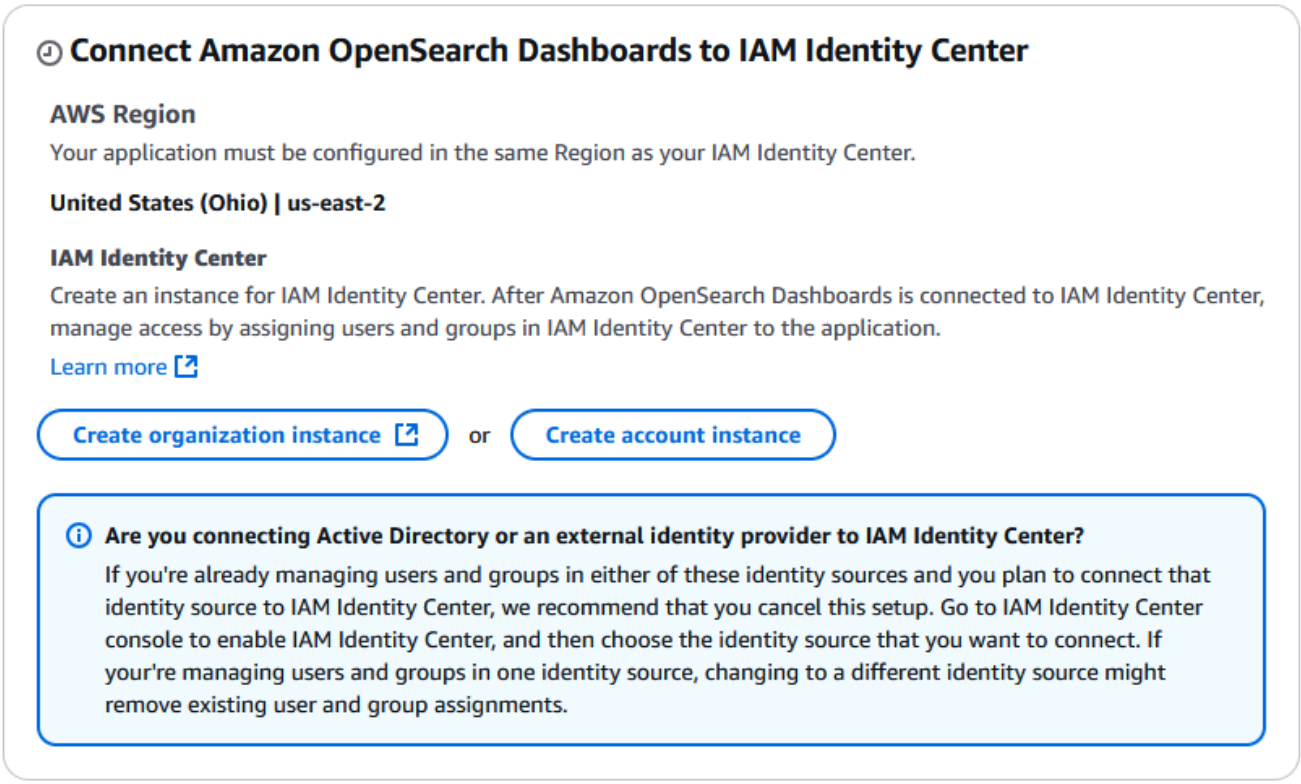
```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
```

```
                "Service":
 "application.opensearchservice.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        },
        {
            "Effect": "Allow",
            "Principal": {
                "Service":
 "application.opensearchservice.amazonaws.com"
            },
            "Action": "sts:SetContext",
            "Condition": {
                "ForAllValues:ArnEquals": {
                "sts:RequestContextProviders":
 "arn:aws:iam::123456789012:oidc-provider/portal.sso.us-
east-1.amazonaws.com/apl/application-id"
                }
            }
        }
    ]
}
```

c. Se un'istanza IAM Identity Center è già stata creata nella tua organizzazione o nel tuo account, la console segnala che Amazon OpenSearch Dashboards è già connesso a un'istanza organizzativa di IAM Identity Center, come mostrato nell'immagine seguente.



Se IAM Identity Center non è ancora disponibile nella tua organizzazione o nel tuo account, tu o un amministratore con le autorizzazioni necessarie potete creare un'istanza dell'organizzazione o un'istanza di account. L'area Connect Amazon OpenSearch Dashboards to IAM Identity Center offre opzioni per entrambi, come mostrato nell'immagine seguente:

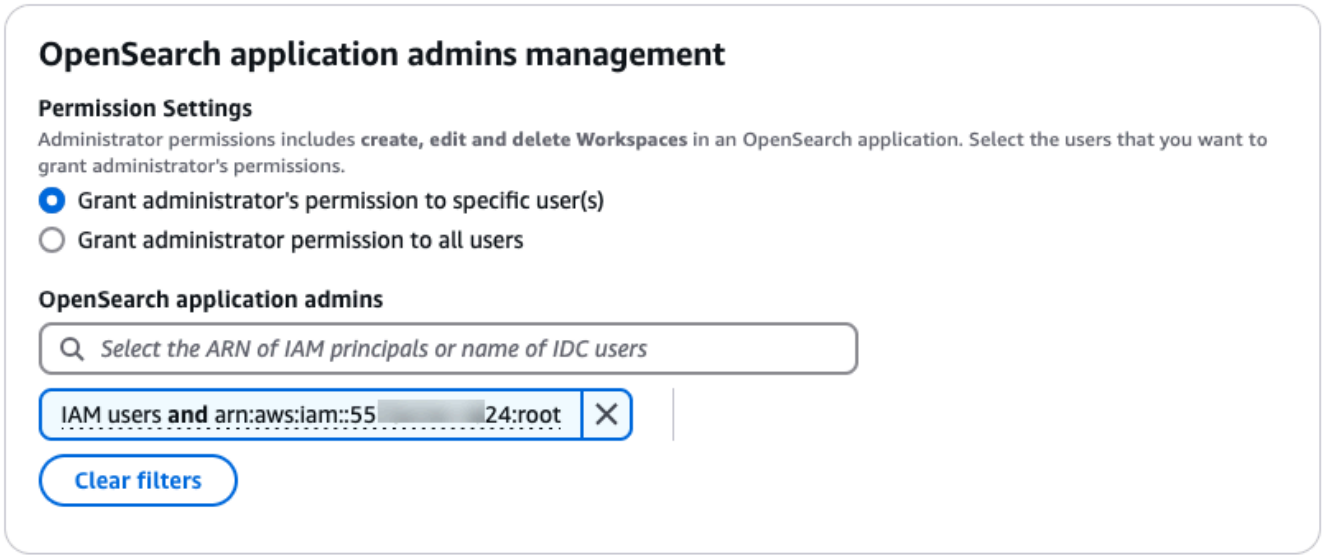


In questo caso, puoi creare un'istanza di account in IAM Identity Center per il test o richiedere a un amministratore di creare un'istanza organizzativa in IAM Identity Center. Per ulteriori informazioni, consulta gli argomenti seguenti nella Guida per l'utente AWS IAM Identity Center :

> **Note**
>
> Se desideri creare applicazioni OpenSearch UI in un'istanza organizzativa diversa Regione AWS dalla tua istanza organizzativa IAM Identity Center, consulta Utilizzo di IAM Identity Center su più istanze Regioni AWS.

- Istanze organizzative di IAM Identity Center
- Istanze account del Centro identità IAM
- Attiva AWS IAM Identity Center

6. (Facoltativo) Verrai aggiunto automaticamente come amministratore dell'applicazione che stai creando. Nell'area di gestione degli amministratori OpenSearch dell'applicazione, puoi concedere le autorizzazioni di amministratore ad altri utenti, come mostrato nell'immagine seguente:



> **Note**
>
> Il ruolo di amministratore dell'applicazione OpenSearch UI concede le autorizzazioni per modificare ed eliminare un'applicazione UI. OpenSearch Gli amministratori dell'applicazione possono anche creare, modificare ed eliminare aree di lavoro in un'applicazione UI. OpenSearch

Per concedere le autorizzazioni di amministratore ad altri utenti, scegliete una delle seguenti opzioni:

- Concedi l'autorizzazione di amministratore a utenti specifici: nel campo amministratori dell'OpenSearchapplicazione, nell'elenco pop-up Proprietà, seleziona utenti IAM o

  AWS IAM Identity Center utenti, quindi scegli i singoli utenti a cui concedere le autorizzazioni di amministratore.
- Concedi l'autorizzazione di amministratore a tutti gli utenti: a tutti gli utenti dell'organizzazione o dell'account vengono concesse le autorizzazioni di amministratore.

7. (Facoltativo) Nell'area Tag, applicate una o più name/value coppie di chiavi di tag all'applicazione.

   I tag sono metadati facoltativi assegnati a una risorsa. Consentono di categorizzare una risorsa in diversi modi, ad esempio in base allo scopo, al proprietario o all'ambiente.
8. Scegli Create (Crea).

Creazione di un'applicazione OpenSearch UI che utilizza AWS IAM Identity Center l'autenticazione tramite AWS CLI

Per creare un'applicazione OpenSearch UI che utilizzi AWS IAM Identity Center l'autenticazione tramite AWS CLI, utilizzate il comando create-application con le seguenti opzioni:

- `--name`— Il nome dell'applicazione.
- `--iam-identity-center-options`— (Facoltativo) L'istanza IAM Identity Center e il ruolo IAM da utilizzare per l'autenticazione e il controllo degli accessi. OpenSearch

Sostituisci *`placeholder values`* con le informazioni appropriate.

```
aws opensearch create-application \
    --name application-name \
    --iam-identity-center-options "
          {
          \"enabled\":true,
          \"iamIdentityCenterInstanceArn\":\"arn:aws:sso:::instance/sso-instance\",
          \"iamRoleForIdentityCenterApplicationArn\":\"arn:aws:iam::account-
id:role/role-name\"
          }
    "
```

Gestione degli amministratori delle applicazioni

L'amministratore di un'applicazione OpenSearch UI è un ruolo definito con l'autorizzazione a modificare ed eliminare un'applicazione OpenSearch UI.

Per impostazione predefinita, in qualità di creatore di un'applicazione OpenSearch UI, sei il primo amministratore dell'applicazione OpenSearch UI.

Gestione degli amministratori OpenSearch dell'interfaccia utente tramite la console

È possibile aggiungere altri amministratori a un'applicazione OpenSearch UI in Console di gestione AWS, durante il flusso di lavoro di creazione dell'applicazione o nella pagina Modifica dopo la creazione dell'applicazione.

Il ruolo di amministratore dell'applicazione OpenSearch UI concede le autorizzazioni per modificare ed eliminare un' OpenSearch applicazione UI. Gli amministratori dell'applicazione possono anche creare, modificare ed eliminare aree di lavoro in un'applicazione UI. OpenSearch

Nella pagina dei dettagli dell'applicazione, puoi cercare l'Amazon Resource Name (ARN) di un principale IAM o cercare il nome di un utente IAM Identity Center.

Per gestire gli amministratori OpenSearch dell'interfaccia utente utilizzando la console

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli OpenSearch UI (dashboard).
3. Nell'area delle OpenSearch applicazioni, scegli il nome di un'applicazione esistente.
4. Scegli Edit (Modifica).
5. Per concedere le autorizzazioni di amministratore ad altri utenti, scegliete una delle seguenti opzioni:
   - Concedi l'autorizzazione di amministratore a utenti specifici: nel campo amministratori dell'OpenSearchapplicazione, nell'elenco pop-up Proprietà, seleziona utenti IAM o

     AWS IAM Identity Center utenti, quindi scegli i singoli utenti a cui concedere le autorizzazioni di amministratore.
   - Concedi l'autorizzazione di amministratore a tutti gli utenti: a tutti gli utenti dell'organizzazione o dell'account vengono concesse le autorizzazioni di amministratore.
6. Scegliere Aggiorna.

È possibile rimuovere amministratori aggiuntivi, ma ogni applicazione dell' OpenSearch interfaccia utente deve mantenere almeno un amministratore.

Gestione degli amministratori OpenSearch dell'interfaccia utente tramite AWS CLI

È possibile creare e aggiornare gli amministratori delle applicazioni dell' OpenSearch interfaccia utente utilizzando. AWS CLI

Creazione di amministratori OpenSearch dell'interfaccia utente utilizzando il AWS CLI

Di seguito sono riportati alcuni esempi di aggiunta di dirigenti IAM e utenti IAM Identity Center come amministratori durante la creazione di un' OpenSearch applicazione UI.

Esempio 1: creazione di un'applicazione OpenSearch UI che aggiunge un utente IAM come amministratore

Esegui il comando seguente per creare un'applicazione OpenSearch UI che aggiunga un utente IAM come amministratore. Sostituisci *placeholder values* con le informazioni appropriate.

```
aws opensearch create-application \
    --name application-name \
    --app-configs "
        {
        \"key\":\"opensearchDashboards.dashboardAdmin.users\",
        \"value\":\"arn:aws:iam::account-id:user/user-id\"
        }
    "
```

Esempio 2: creare un'applicazione OpenSearch UI che abiliti IAM Identity Center e aggiunga un ID utente IAM Identity Center come amministratore dell'applicazione OpenSearch UI

Esegui il comando seguente per creare un'applicazione OpenSearch UI che abiliti IAM Identity Center e aggiunga un ID utente IAM Identity Center come amministratore dell'applicazione OpenSearch UI. Sostituisci *placeholder values* con le informazioni appropriate.

`key`specifica l'elemento di configurazione da impostare, ad esempio il ruolo di amministratore per l'applicazione OpenSearch UI. I valori validi includono `opensearchDashboards.dashboardAdmin.users` e `opensearchDashboards.dashboardAdmin.groups`.

*xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx*rappresenta il valore assegnato alla chiave, ad esempio l'Amazon Resource Name (ARN) di un utente IAM.

```
aws opensearch create-application \
    --name myapplication \
    --iam-identity-center-options "
        {
        \"enabled\":true,
        \"iamIdentityCenterInstanceArn\":\"arn:aws:sso:::instance/ssoins-instance-id\",
        \"iamRoleForIdentityCenterApplicationArn\":\"arn:aws:iam::account-id:role/role-
name\"
        }
    " \
    --app-configs "
        {
```

```
        \"key\":\"opensearchDashboards.dashboardAdmin.users\",
        \"value\":\"xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx\"
        }
    "
```

Aggiornamento degli amministratori OpenSearch dell'interfaccia utente tramite AWS CLI

Di seguito sono riportati alcuni esempi di aggiornamento dei principali IAM e degli utenti di IAM Identity Center assegnati come amministratori per un'applicazione esistente. OpenSearch

Esempio 1: aggiungere un utente IAM come amministratore per un'applicazione esistente OpenSearch

Esegui il comando seguente per aggiornare un'applicazione OpenSearch UI e aggiungere un utente IAM come amministratore. Sostituisci *placeholder values* con le informazioni appropriate.

```
aws opensearch update-application \
    --id myapplication \
    --app-configs "
        {
        \"key\":\"opensearchDashboards.dashboardAdmin.users\",
        \"value\":\"arn:aws:iam::account-id:user/user-id\"
        }
    "
```

Esempio 2: aggiorna un'applicazione OpenSearch UI per aggiungere un ID utente IAM Identity Center come amministratore dell'applicazione OpenSearch UI

Esegui il comando seguente per aggiornare un'applicazione OpenSearch UI e aggiungere un ID utente IAM Identity Center come amministratore dell'applicazione OpenSearch UI. Sostituisci *placeholder values* con le informazioni appropriate.

`key`specifica l'elemento di configurazione da impostare, ad esempio il ruolo di amministratore per l'applicazione OpenSearch UI. I valori validi includono `opensearchDashboards.dashboardAdmin.users` e `opensearchDashboards.dashboardAdmin.groups`.

*xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx*rappresenta il valore assegnato alla chiave, ad esempio l'Amazon Resource Name (ARN) di un utente IAM.

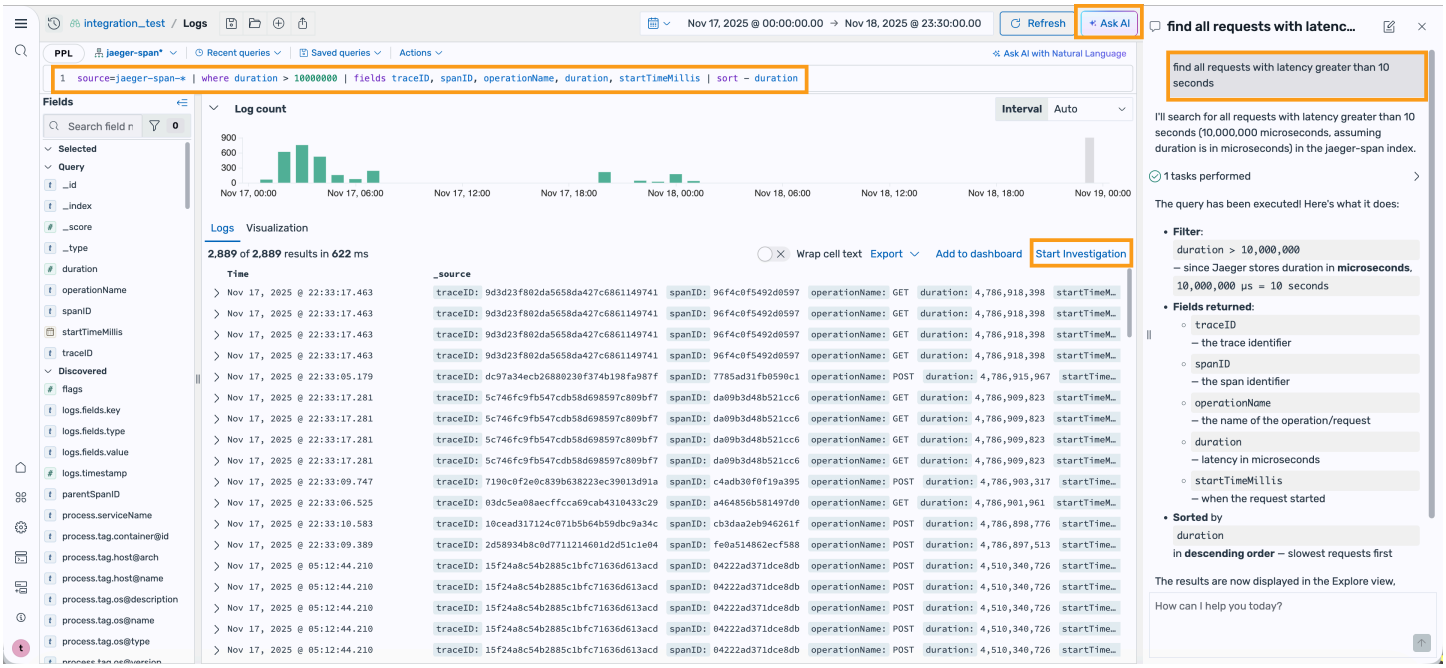
```
aws opensearch update-application \
```

```
    --id myapplication \
    --app-configs "
        {
        \"key\":\"opensearchDashboards.dashboardAdmin.users\",
        \"value\":\"xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx\"
        }
    "
```

AI agentica in Amazon Service OpenSearch

OpenSearch L'interfaccia utente include una suite di funzionalità di intelligenza artificiale agentica progettate per semplificare l'analisi operativa e accelerare le indagini sugli incidenti per i team di progettazione e supporto.

In ogni pagina dell'applicazione OpenSearch UI, scegli il pulsante Ask AI per aprire Agentic Chat. Agentic Chat comprende il contesto della pagina che stai visualizzando, utilizza strumenti agentici per analizzare i dati sottostanti, aggiorna le query per visualizzare i risultati in Discover e risponde alle tue domande con chiarezza e precisione. Quando un'analisi complessa della causa principale richiede un'indagine più approfondita, puoi far sì che l'agente investigativo pianifichi, esegua e rifletta autonomamente attraverso un flusso di lavoro in più fasi, fornendo ipotesi strutturate supportate da prove di dati. Alla base di entrambe le funzionalità agentiche c'è Agentic Memory, un livello di memoria che conserva il contesto all'interno della conversazione o dell'indagine, in modo da poter continuare a porre domande di follow-up con lo stesso contesto coerente in diverse sessioni web.



Funzionalità chiave

- Agentic Chat: integrata nella funzione Ask AI in ogni pagina dell' OpenSearch interfaccia utente, Agentic Chat comprende il contesto attuale e utilizza strumenti per analizzare i dati per rispondere al meglio alle tue domande. Su Discover, inserisci il linguaggio naturale nell'interfaccia di chat per generare e iterare le query PPL (Piped Processing Language), quindi non devi essere un esperto di PPL per ottenere informazioni utili. Puoi anche avviare conversazioni direttamente dalle visualizzazioni, dove Agentic Chat può identificare anomalie nei grafici, correlarle con i dati sottostanti e generare analisi. Quando è necessaria un'analisi complessa della causa principale, utilizzate il comando `/investigate` slash per avviare l'Investigation Agent in base alla conversazione in corso o con nuove istruzioni.
- Agente investigativo: un agente di ricerca approfondita orientato agli obiettivi che puoi attivare digitando `/investigate` in chat o scegliendo il pulsante Avvia indagine nelle pagine delle funzionalità. Inserisci un obiettivo di indagine, ad esempio «trova la causa principale di questo picco di latenza». L'agente investigativo pianifica in modo autonomo utilizzando i tuoi dati e l'obiettivo dichiarato, esegue interrogazioni e analisi e riflette attraverso un flusso di lavoro in più fasi. Una volta completata l'indagine, in genere entro pochi minuti, genera ipotesi strutturate classificate per verosimiglianza, ciascuna supportata da prove di dati. Fornisce la massima trasparenza in ogni fase del ragionamento, in modo da poter verificare e fidarsi dei risultati.
- Agentic Memory: un livello di memoria gestito dai servizi che conserva il contesto all'interno della conversazione o dell'indagine, in modo che possiate continuare a porre domande o indagare all'interno di un contesto coerente. Agentic Memory consente sia ad Agentic Chat che all'Investigation Agent di mantenere la continuità durante una sessione.

Limiti di prezzo e utilizzo

Le funzionalità di Agentic AI sono gratuite. I limiti di utilizzo basati sui token si applicano per account per prevenire gli abusi.

Le funzionalità di intelligenza artificiale agentica utilizzano un modello di base in Amazon Bedrock. Tutte le configurazioni di sicurezza e governance dei dati di Amazon Bedrock si applicano alle funzionalità di intelligenza artificiale agentica. Per ulteriori informazioni, consulta la sezione Protezione dei dati in Amazon Bedrock.

Sicurezza

Le funzionalità di intelligenza artificiale agentica operano all'interno del framework di sicurezza esistente e non aggirano né sostituiscono i controlli di accesso configurati. Quando interagisci con Agentic Chat o avvii un agente investigativo, le funzionalità di intelligenza artificiale si autenticano utilizzando le tue attuali credenziali IAM o IAM Identity Center ed ereditano tutte le autorizzazioni associate. L'assistente AI può solo interrogare fonti di dati, indici e visualizzazioni a cui hai già accesso esplicito tramite le politiche di controllo degli accessi basate sui ruoli (RBAC) esistenti.

Se il tuo OpenSearch dominio implementa un controllo granulare degli accessi con sicurezza a livello di campo o di documento, l'IA rispetta queste restrizioni e non può far emergere dati da campi o documenti con restrizioni che non rientrano nell'ambito delle autorizzazioni. Allo stesso modo, tutte le query PPL generate da Agentic Chat o i recuperi di dati eseguiti durante le indagini vengono eseguiti nel contesto dell'utente, garantendo che tutte le politiche di sicurezza configurate rimangano pienamente applicate.

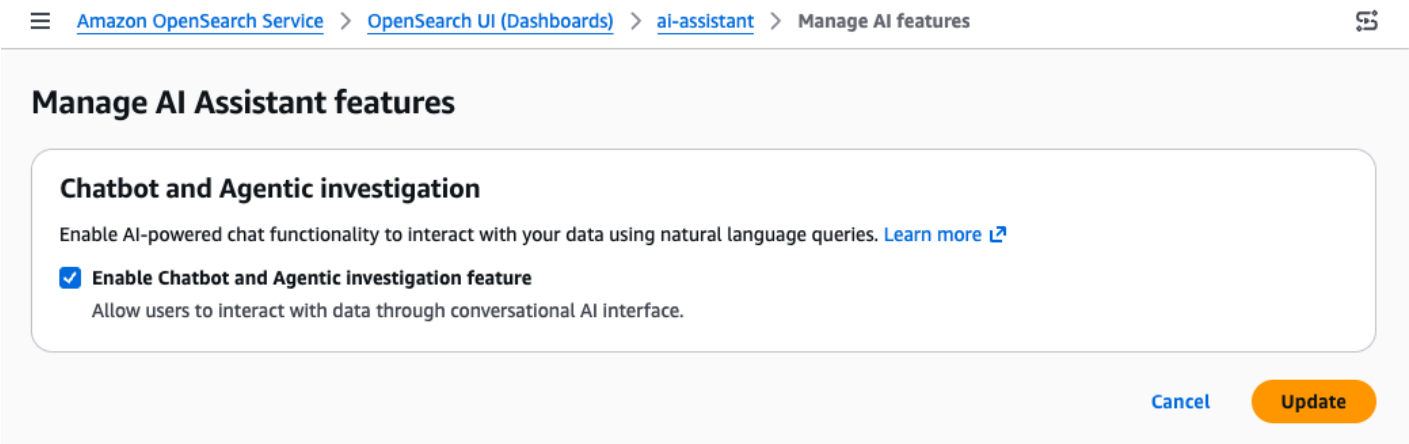
Questo modello di ereditarietà delle autorizzazioni garantisce che l'attivazione delle funzionalità di intelligenza artificiale non crei nuove vulnerabilità di sicurezza né richieda agli amministratori di gestire strutture di autorizzazione separate per le interazioni con l'IA.

Abilitazione e disabilitazione delle funzionalità di intelligenza artificiale agentica

Le funzionalità di intelligenza artificiale agentica vengono abilitate automaticamente per le nuove applicazioni OpenSearch dell'interfaccia utente e per le applicazioni esistenti, a meno che tu non abbia disabilitato esplicitamente le funzionalità di intelligenza artificiale in uno dei tuoi domini. OpenSearch

Per abilitare o disabilitare le funzionalità di intelligenza artificiale, puoi utilizzare l'o l' Console di gestione AWS API:

- Console: accedi alla pagina dei dettagli dell'applicazione OpenSearch UI e aggiorna le funzionalità AI da lì.



- API: utilizza l'RegisterCapabilityAPI per abilitare le funzionalità di intelligenza artificiale o utilizza l'DeregisterCapabilityAPI per disabilitarle.

Disponibilità nelle regioni

Le funzionalità di intelligenza artificiale agentica sono disponibili nelle seguenti versioni: Regioni AWS

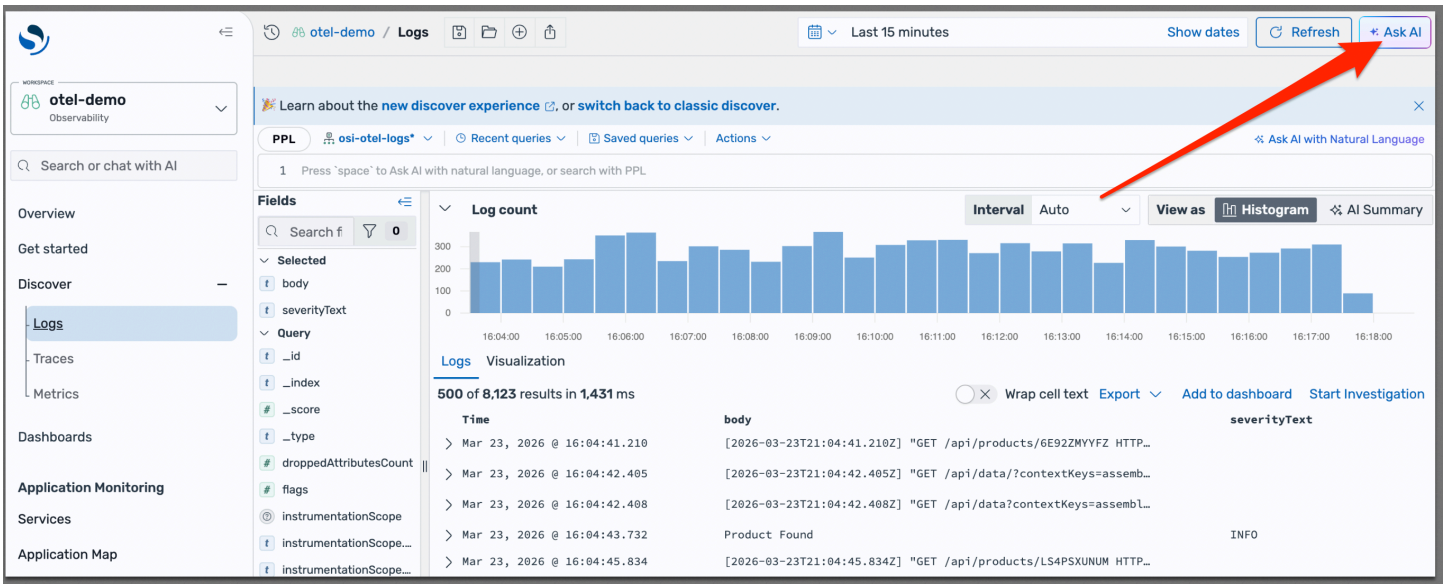
- Asia Pacifico (Tokyo) - ap-northeast-1
- Asia Pacifico (Sydney) - ap-southeast-2
- UE (Francoforte) - eu-central-1
- Europa (Stoccolma) - eu-nord-1
- Europa (Spagna): eu-south-2
- Europa (Irlanda) - eu-west-1
- Stati Uniti orientali (Virginia settentrionale) - us-east-1
- Stati Uniti orientali (Ohio) - us-east-2
- Stati Uniti occidentali (Oregon) - us-west-2

Argomenti

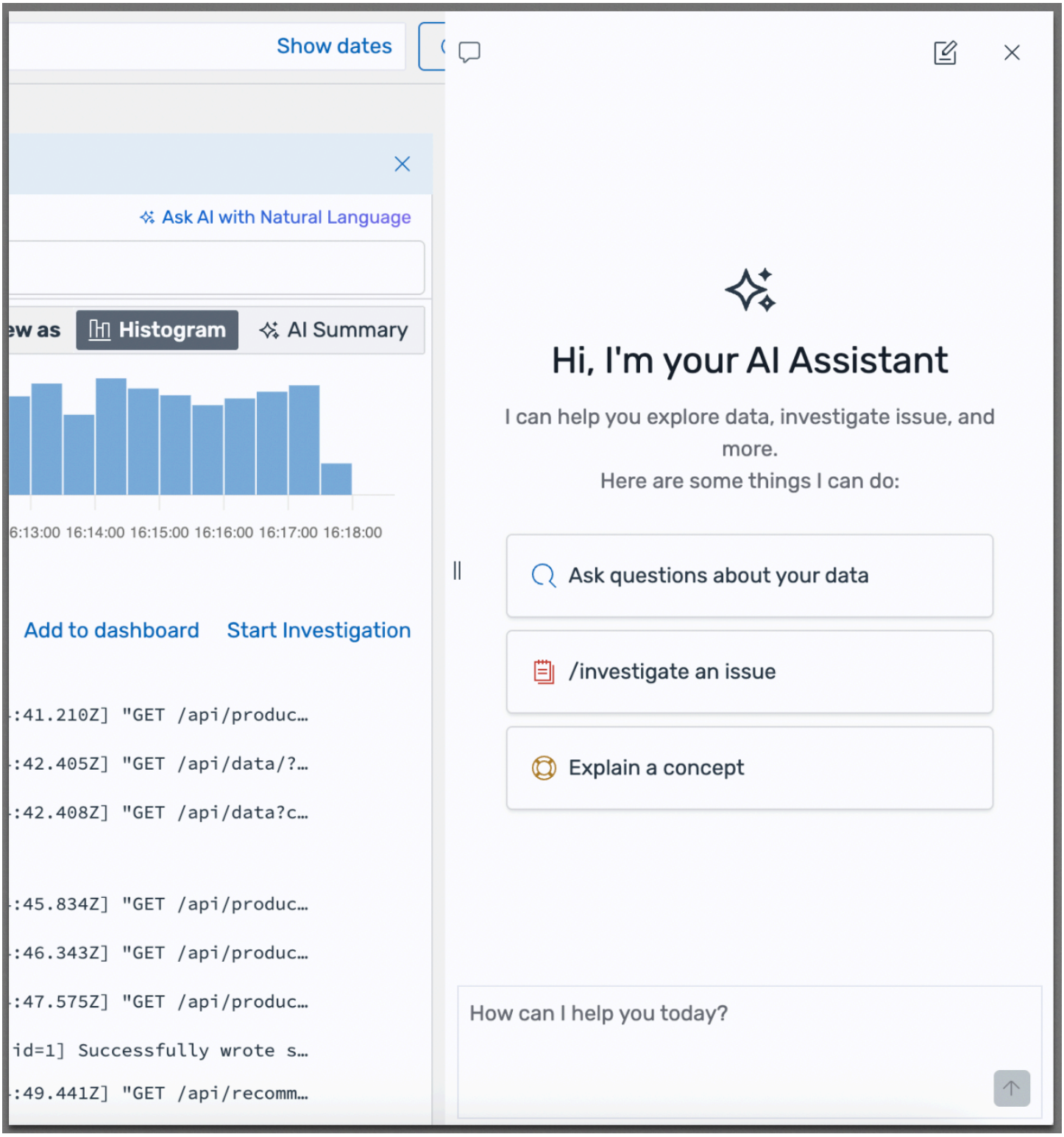
- Chat agentica nel servizio Amazon OpenSearch
- Agente investigativo in Amazon OpenSearch Service
- Memoria agentica in Amazon Service OpenSearch

Chat agentica nel servizio Amazon OpenSearch

Agentic Chat è un assistente AI incorporato in ogni pagina dell'interfaccia utente. OpenSearch Scegli il pulsante Ask AI per aprire il pannello della chat, dove puoi porre domande sui tuoi dati, generare domande e avviare indagini. Agentic Chat comprende il contesto della pagina che stai visualizzando in Discover and Investigation e utilizza strumenti agentici per analizzare i dati sottostanti.



Quando apri il pannello della chat, Agentic Chat presenta delle opzioni per aiutarti a iniziare: porre domande sui tuoi dati, esaminare un problema o spiegare un concetto. Se hai iniziato una conversazione in precedenza, questa rimane visibile nel pannello della chat mentre navighi tra le pagine, in modo da poter continuare da dove avevi interrotto. In alternativa, scegli il pulsante Nuova chat nell'angolo in alto a destra per iniziare una nuova conversazione.



Utilizzo di Agentic Chat con Discover

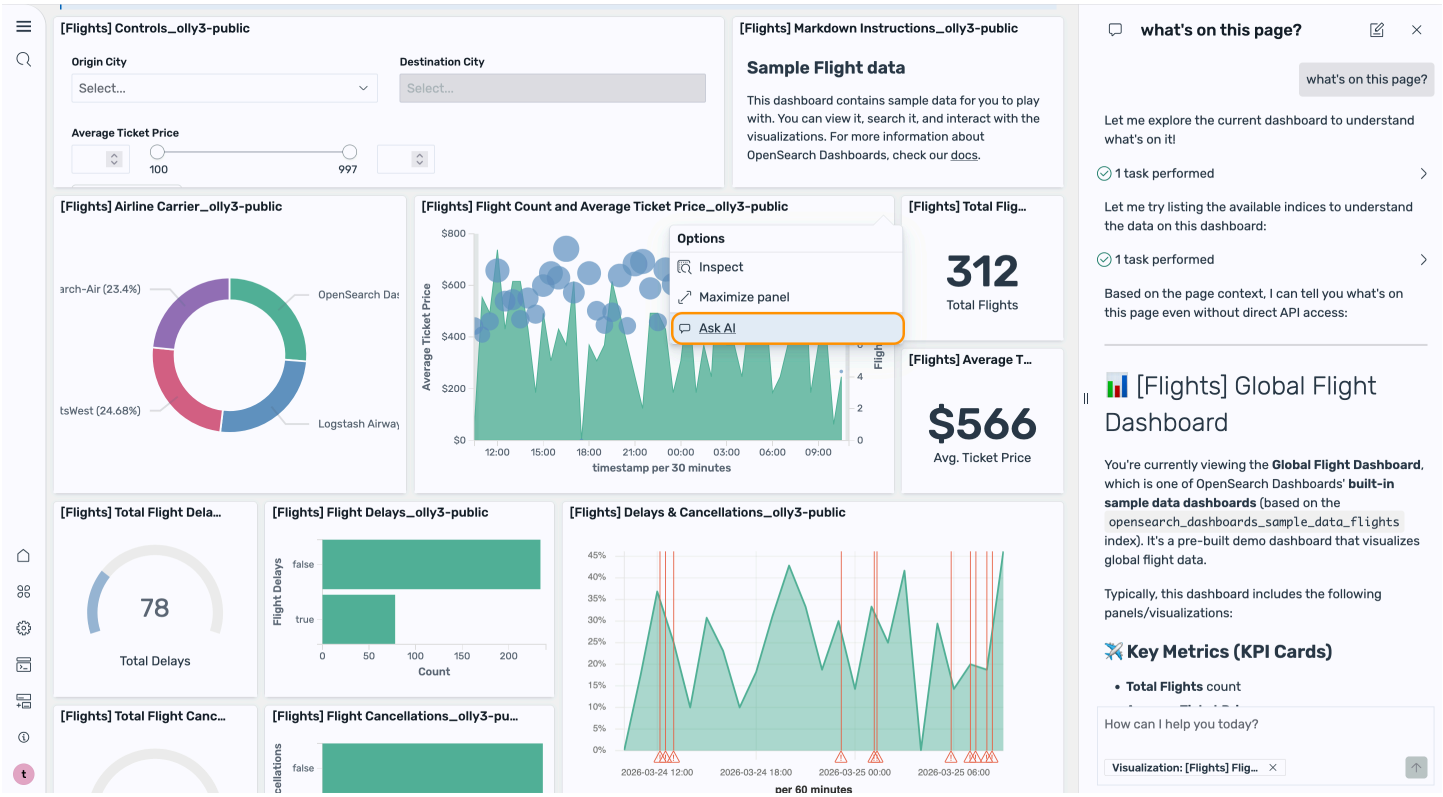
Nella pagina Discover degli spazi di lavoro di Observability, puoi inserire il linguaggio naturale nell'interfaccia di chat per generare query PPL. Agentic Chat traduce le tue domande in PPL, esegue

la query e visualizza i risultati direttamente nella vista Discover. Non devi essere un esperto di PPL per ottenere informazioni utili dai tuoi dati.

Per perfezionare una query generata, fai domande di follow-up in linguaggio naturale, ad esempio «aggiungi un filtro per il codice di stato 500". Agentic Chat comprende il contesto della query corrente e lo modifica di conseguenza. Puoi anche chiedere di modificare le aggregazioni, modificare gli intervalli di tempo o aggiungere campi aggiuntivi ai risultati. Ogni iterazione aggiorna la vista Discover con i nuovi risultati della query.

Utilizzo di Agentic Chat con visualizzazioni

Puoi iniziare una conversazione con Agentic Chat direttamente da una visualizzazione. Apri il menu contestuale su un pannello di visualizzazione e scegli Ask AI. Agentic Chat analizza la visualizzazione, identifica le anomalie nei grafici, si correla con i dati sottostanti e genera analisi.



Avvio di un'indagine dalla chat

Quando è necessaria un'analisi complessa della causa principale, puoi avviare Investigation Agent direttamente da Agentic Chat. Usa il comando `/investigate` slash nell'input della chat o scegli il pulsante Avvia indagine nelle pagine delle funzionalità.

Per ulteriori informazioni sull'agente investigativo, consultathe section called “Agente investigativo”.

Strumenti supportati

Agentic Chat utilizza i seguenti strumenti per analizzare i dati e rispondere alle domande. Per visualizzare la maggior parte up-to-date degli strumenti disponibili, digita «quali strumenti puoi usare» nell'interfaccia della chat.

Strumenti di frontend

Questi strumenti aggiornano l' OpenSearch interfaccia utente:

- `create_investigation`— Crea un nuovo taccuino per le indagini sugli agenti con dettagli come obiettivi, sintomi, indice e intervallo di tempo.
- `execute_ppl_query`— Esegue interrogazioni PPL sul set di dati corrente e visualizza i risultati nella pagina Discover.
- `update_time_range`— Aggiorna il filtro dell'intervallo temporale globale nella pagina Discover corrente (ad esempio, «ultime 24 ore» o «ultima settimana»).

Strumenti di backend

Questi strumenti interagiscono direttamente con OpenSearch i dati e APIs:

- `SearchIndexTool`— Cerca in un indice utilizzando query DSL.
- `MsearchTool`— Esegue più operazioni di ricerca in un'unica richiesta.
- `CountTool`— Restituisce il numero di documenti corrispondenti a una query.
- `ExplainTool`— Spiega perché un documento corrisponde o non corrisponde a una query.
- `ListIndexTool`— Elenca gli indici nel cluster con dettagli opzionali.
- `IndexMappingTool`— Recupera le mappature e le impostazioni degli indici.
- `GetShardsTool`— Ottiene informazioni frammentate per un indice.
- `ClusterHealthTool`— Restituisce informazioni sullo stato del cluster.
- `LogPatternAnalysisTool`— Analizza i modelli di log, confronta gli intervalli di tempo o esegue l'analisi delle sequenze di traccia.
- `MetricChangeAnalysisTool`— Confronta le distribuzioni percentili dei campi numerici tra due intervalli di tempo.
- `DataDistributionTool`— Analizza le distribuzioni dei valori dei campi in un intervallo di tempo target, facoltativamente rispetto a una linea di base.

- `GenericOpenSearchApiTool`— Uno strumento flessibile per chiamare direttamente qualsiasi OpenSearch endpoint API.

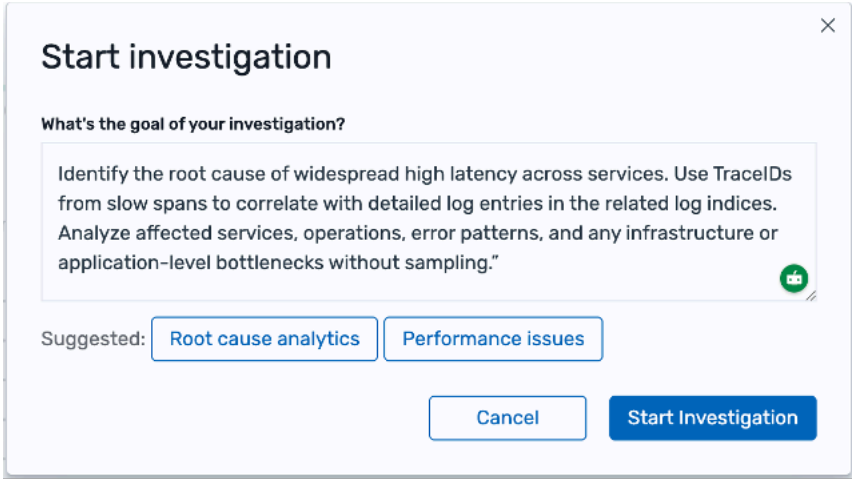
Agente investigativo in Amazon OpenSearch Service

L'Investigation Agent è un agente di ricerca approfondita nell' OpenSearch interfaccia utente orientato agli obiettivi che indaga autonomamente su questioni complesse per tuo conto. Pianifica l'utilizzo dei dati e dell'obiettivo dichiarato, esegue interrogazioni e analisi e riflette attraverso un flusso di lavoro in più fasi. Una volta completata l'indagine, in genere entro pochi minuti, genera ipotesi strutturate classificate per verosimiglianza, ciascuna supportata da prove di dati. Fornisce la massima trasparenza in ogni fase del ragionamento, in modo da poter verificare e fidarsi dei risultati.

Avvio di un'indagine

Puoi avviare un'indagine in due modi:

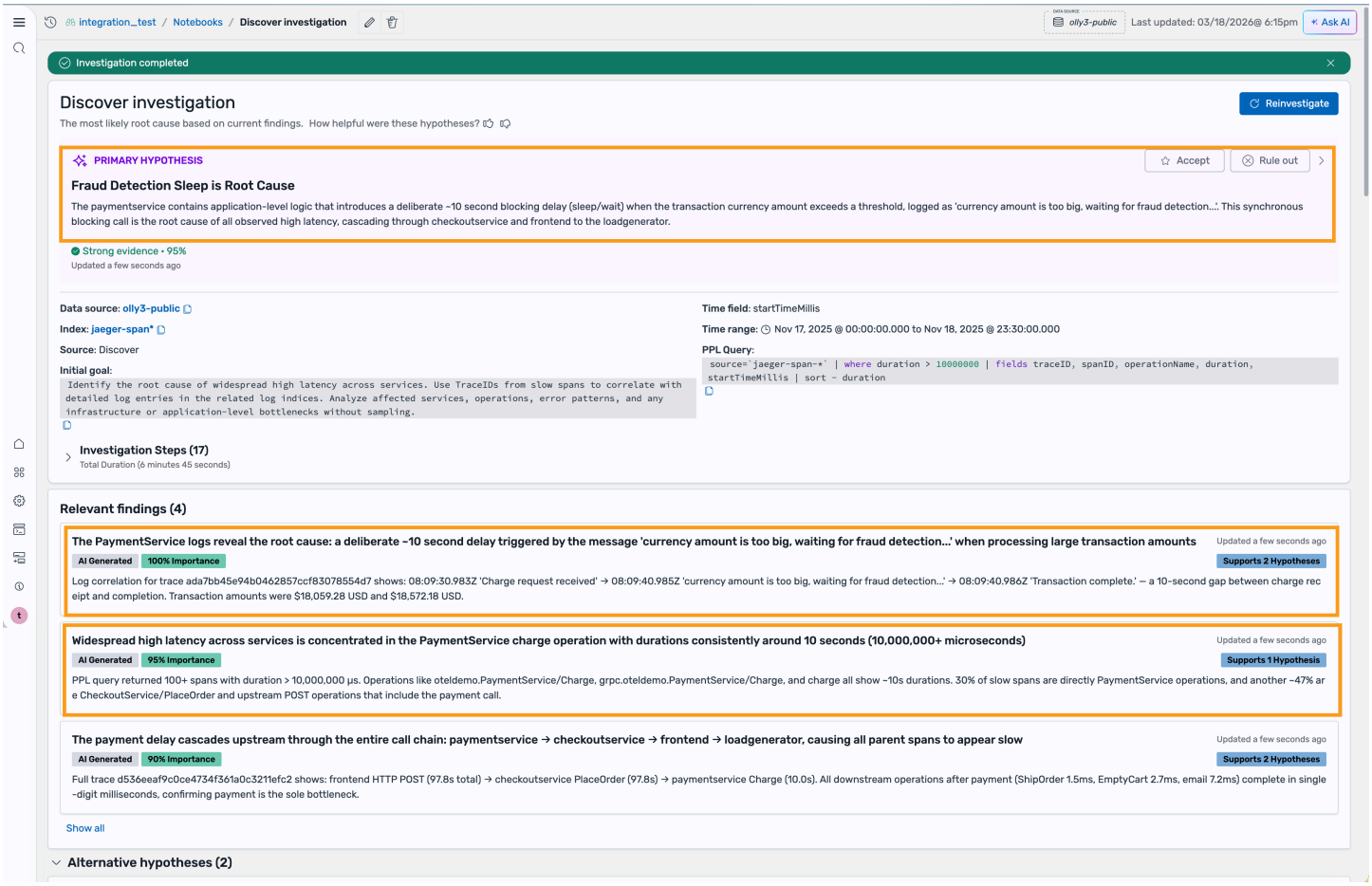
- Da Discover, Visualization o altre pagine di funzionalità supportate, scegli il pulsante Avvia indagine. Viene visualizzata una finestra di dialogo in cui puoi inserire l'obiettivo dell'indagine e scegliere tra modelli suggeriti come «Root cause analytics» o «Problemi di prestazioni».
- Da Agentic Chat, digita il comando `/investigate` slash nell'input della chat con l'obiettivo dell'indagine.



Per ulteriori informazioni su Agentic Chat, consulta. the section called “Chat agentica”

Revisione dei risultati delle indagini

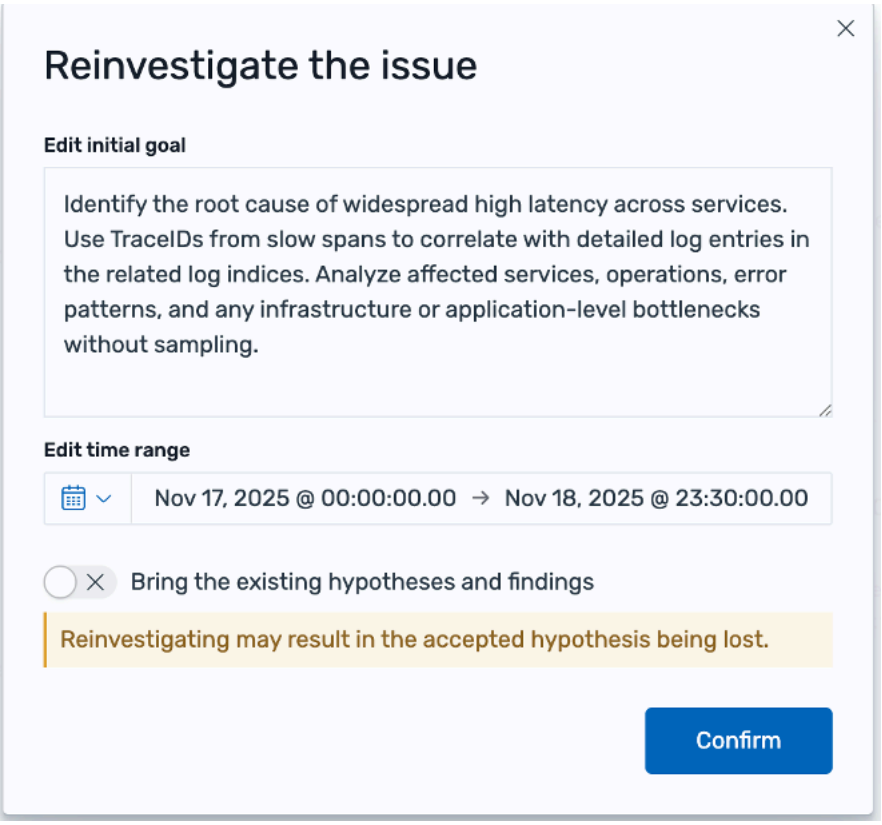
Una volta completata l'indagine, l'agente investigativo presenta un'ipotesi primaria con un livello di confidenza e prove a sostegno. La pagina dei risultati mostra le fasi dell'indagine intraprese, i risultati pertinenti classificati per importanza e le ipotesi alternative.



Puoi esaminare i risultati alla base di ogni ipotesi, quindi scegliere Accetta per confermare l'ipotesi o Escludi per rifiutarla. È possibile esaminare anche ipotesi alternative con minore probabilità. Puoi selezionare qualsiasi ipotesi alternativa come conclusione finale se corrisponde meglio alla tua valutazione.

Ri-investigare

Se i risultati dell'indagine richiedono ulteriori chiarimenti o se l'agente investigativo stabilisce che alla domanda dell'indagine non è possibile rispondere con i set di dati disponibili, è possibile utilizzare l'opzione Reinvestiga per modificare e rieseguire l'indagine. Scegliete Riinvestiga per modificare l'obiettivo iniziale, modificare l'intervallo di tempo e, facoltativamente, inserire le ipotesi e i risultati esistenti nella nuova indagine.



Memoria agentica in Amazon Service OpenSearch

Agentic Memory è un livello di memoria gestito dai servizi che alimenta Agentic Chat e Investigation Agent. Mantiene il contesto all'interno della conversazione o dell'indagine, in modo da offrire un'esperienza coerente tra diverse pagine di funzionalità, schede del browser e aggiornamenti delle pagine. Agentic Memory funziona automaticamente e non richiede alcuna configurazione da parte dell'utente.

Agentic Memory è basato sul framework di memoria degli OpenSearch agenti. L'archiviazione della memoria è isolata dall'ID utente per motivi di privacy.

Protezione dei dati

L'uso di Agentic Memory è gratuito. I dati dei clienti archiviati in Agentic Memory sono crittografati con una chiave gestita dal servizio. Se hai abilitato la crittografia a chiave gestita dal cliente (CMK) per la tua applicazione OpenSearch UI, i dati di memoria verranno invece crittografati con la CMK. La memoria viene archiviata in una raccolta Amazon OpenSearch Serverless gestita dal servizio.

Per ulteriori informazioni sulla crittografia CMK, consulta. the section called “Crittografia e chiave gestita dal cliente”

Limitazioni

Agentic Memory non è in grado di conservare il contesto tra diversi thread di conversazione.

Crittografia dei metadati delle applicazioni dell' OpenSearch interfaccia utente con chiavi gestite dal cliente

Le risorse visive e le configurazioni vengono archiviate come metadati per le applicazioni dell'interfaccia utente. OpenSearch Ciò include interrogazioni, visualizzazioni e dashboard salvate. I dati provenienti dalle fonti di dati associate non vengono archiviati nei metadati. Per informazioni sulla crittografia dei dati nelle tue fonti di dati, consulta Protezione dei dati in Amazon OpenSearch Service per i OpenSearch domini e Crittografia in Amazon OpenSearch Serverless per le raccolte serverless.

I metadati OpenSearch dell'interfaccia utente sono protetti con crittografia a riposo. Ciò impedisce l'accesso non autorizzato. La crittografia utilizza AWS Key Management Service (AWS KMS) per archiviare e gestire le chiavi di crittografia. Per impostazione predefinita, OpenSearch i metadati dell'interfaccia utente sono crittografati con chiavi AWS proprietarie.

Puoi anche utilizzare la funzionalità Customer Managed Key (CMK) per gestire le tue chiavi di crittografia. Ciò consente di soddisfare i requisiti normativi e di conformità. Per utilizzare CMK, è necessario creare una nuova applicazione OpenSearch UI e abilitare CMK nel processo di creazione. Al momento non è supportato l'aggiornamento di un'applicazione OpenSearch UI esistente dalla chiave AWS proprietaria a CMK.

Quando utilizzare le chiavi gestite dal cliente:

- La tua organizzazione ha requisiti di conformità normativa per la gestione delle chiavi
- Sono necessari audit trail per l'utilizzo delle chiavi di crittografia
- Vuoi controllare i programmi di rotazione dei tasti
- È necessario integrarsi con i flussi di lavoro di gestione delle chiavi esistenti

Quando utilizzi una chiave gestita dal cliente, hai il pieno controllo sulla chiave. Ciò include la capacità di:

- Stabilire e mantenere le policy della chiave
- Stabilire e mantenere concessioni e policy IAM
- Attivare e disattivare la chiave

- Ruota il materiale crittografico della chiave
- Aggiungi tag alla chiave
- Creare alias delle chiavi
- Pianifica l'eliminazione della chiave

> **Note**
>
> La chiave gestita dal cliente deve trovarsi nella Regione AWS stessa applicazione dell' OpenSearch interfaccia utente. Non è possibile utilizzare una chiave proveniente da un'altra regione.

Argomenti

- Prerequisiti per l'utilizzo delle chiavi gestite dal cliente
- Creazione di un'applicazione con crittografia a chiave gestita dal cliente utilizzando la console
- Creazione di un'applicazione con crittografia a chiave gestita dal cliente utilizzando AWS CLI
- Monitoraggio dell'utilizzo delle chiavi gestite dal cliente
- Aggiornamento delle impostazioni di crittografia

Prerequisiti per l'utilizzo delle chiavi gestite dal cliente

Prima di poter utilizzare una chiave gestita dal cliente per crittografare i metadati dell'applicazione dell' OpenSearch interfaccia utente, è necessario creare una chiave di crittografia simmetrica. AWS KMSPer istruzioni sulla creazione di chiavi, consulta Creating keys nella Developer Guide.AWS KMS

La politica chiave per la chiave gestita dal cliente deve concedere all' OpenSearch interfaccia utente l'autorizzazione all'uso della chiave. Utilizza la seguente politica chiave, sostituendola *`placeholder values`* con le tue informazioni:

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowOpenSearchUIToUseKey",
            "Effect": "Allow",
            "Principal": {
```

```
                "Service": [
                    "application.opensearchservice.amazonaws.com"
                ]
            },
            "Action": [
                "kms:Encrypt",
                "kms:Decrypt",
                "kms:GenerateDataKey",
                "kms:DescribeKey"
            ],
            "Resource": "*"
        },
        {
            "Sid": "AllowKeyAdministration",
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::account-id:root"
            },
            "Action": "kms:*",
            "Resource": "*"
        }
    ]
}
```

La policy include due dichiarazioni:

- La prima istruzione consente all' OpenSearch interfaccia utente di utilizzare la chiave per le operazioni di crittografia.
- La seconda istruzione consente agli utenti dell'utente Account AWS di amministrare la chiave. Ciò include le autorizzazioni per aggiornare la politica delle chiavi, abilitare o disabilitare la chiave e pianificare l'eliminazione della chiave. Puoi limitare ulteriormente queste autorizzazioni sostituendo il root principal con utenti o ruoli IAM specifici.

Per ulteriori informazioni sulle politiche chiave, consulta Using key policy AWS KMS nella AWS KMS Developer Guide.

Creazione di un'applicazione con crittografia a chiave gestita dal cliente utilizzando la console

Quando OpenSearch crei un'applicazione UI nella console, puoi specificare una chiave gestita dal cliente per crittografare i metadati dell'applicazione.

Per creare un'applicazione OpenSearch UI con crittografia a chiave gestita dal cliente utilizzando la console

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli OpenSearch UI (dashboard).
3. Scegli Crea applicazione.
4. Per Nome applicazione, inserisci un nome per l'applicazione.
5. Configura le impostazioni di autenticazione e amministratore in base alle esigenze. Per ulteriori informazioni, consulta the section called “Nozioni di base”.
6. Nella sezione Crittografia, per Encryption at rest, scegli Usa chiave gestita dal cliente.
7. Seleziona una chiave gestita dal cliente esistente dall'elenco oppure scegli Crea una chiave per crearne una nuova AWS KMS.

   > **Note**
   >
   > La chiave deve essere nella Regione AWS stessa dell'applicazione che stai creando.

8. (Facoltativo) Aggiungi tag all'applicazione.
9. Scegli Create (Crea).

Creazione di un'applicazione con crittografia a chiave gestita dal cliente utilizzando AWS CLI

Per creare un'applicazione OpenSearch UI con crittografia a chiave gestita dal cliente utilizzando il AWS CLI, utilizzate il comando create-application con il `--kms-key-arn` parametro.

Sostituisci *`placeholder values`* con le informazioni appropriate.

```
aws opensearch create-application \
```

```
    --name my-application \
    --kms-key-arn arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012
```

Se non specificate il `--kms-key-arn` parametro, OpenSearch utilizza una chiave AWS-managed per crittografare i metadati dell'applicazione.

Monitoraggio dell'utilizzo delle chiavi gestite dal cliente

Quando si utilizza una chiave gestita dal cliente con un'applicazione di OpenSearch interfaccia utente, AWS KMS registra ogni utilizzo della chiave nei AWS CloudTrail registri. È possibile utilizzare questi registri per monitorare come e quando viene utilizzata la chiave. I registri mostrano quale utente o servizio ha avuto accesso alla chiave.

AWS AWS KMS ruota automaticamente le chiavi gestite dal cliente ogni anno. È inoltre possibile ruotare manualmente i tasti in base alle esigenze. Per ulteriori informazioni sulla rotazione dei tasti, consulta Rotating KMS keys nella Developer Guide.AWS KMS

Per ulteriori informazioni sul monitoraggio dell'utilizzo delle chiavi, consulta Logging AWS KMS API call with AWS CloudTrail nella Developer Guide.AWS KMS

> **Note**
>
> L'utilizzo di chiavi gestite dal cliente comporta dei costi. AWS KMS I costi si basano sul numero di richieste API e di chiavi archiviate. Per i dettagli sui prezzi, consulta la sezione Prezzi del servizio di gestione delle AWS chiavi.

Aggiornamento delle impostazioni di crittografia

Dopo aver creato un'applicazione OpenSearch UI, non è possibile modificarne le impostazioni di crittografia. Se è necessario utilizzare una chiave gestita dal cliente diversa, è necessario creare una nuova applicazione. Se è necessario passare da una chiave gestita AWS a una chiave gestita dal cliente e viceversa, è inoltre necessario creare una nuova applicazione con le impostazioni di crittografia desiderate.

> **Important**
>
> Prima di disabilitare o eliminare una chiave gestita dal cliente, considera quanto segue:

- Se disabiliti la chiave, l'applicazione perderà l'accesso ai metadati crittografati. È necessario riattivare la stessa chiave per ripristinare l'accesso.
- Se si elimina la chiave, gli oggetti salvati nell'applicazione diventano definitivamente inaccessibili. Ciò include interrogazioni, visualizzazioni e dashboard. Le chiavi eliminate non possono essere recuperate.
- Ti consigliamo di documentare l'ARN della tua chiave prima di apportare modifiche allo stato della chiave.

Fasi successive

Dopo aver configurato la crittografia CMK per la tua applicazione, puoi:

- Associare fonti di dati all'applicazione. Per ulteriori informazioni, consulta the section called “Gestione delle associazioni delle fonti di dati e delle autorizzazioni di accesso VPC”.
- Crea spazi di lavoro per il tuo team. Per ulteriori informazioni, consulta the section called “Utilizzo delle aree di lavoro OpenSearch di Amazon Service”.
- Imposta il AWS CloudTrail monitoraggio dell'utilizzo delle chiavi. Per ulteriori informazioni, consulta the section called “Monitoraggio dell'utilizzo delle chiavi gestite dal cliente”.

Abilitazione della federazione SAML con AWS Identity and Access Management

OpenSearch L'interfaccia utente supporta Security Assertion Markup Language 2.0 (SAML), uno standard aperto utilizzato da molti provider di identità. Ciò consente la federazione delle identità con AWS Identity and Access Management (IAM). Con questo supporto, gli utenti del tuo account o della tua organizzazione possono accedere direttamente all' OpenSearch interfaccia utente assumendo ruoli IAM. Puoi creare un'esperienza Single Sign-On avviata dal provider di identità (IdP) per gli utenti finali, in cui possono autenticarsi nel provider di identità esterno ed essere indirizzati direttamente alla pagina definita nell'interfaccia utente. OpenSearch Puoi anche implementare un controllo granulare degli accessi configurando gli utenti finali o i gruppi in modo che assumano ruoli IAM diversi con autorizzazioni diverse per l'accesso all'interfaccia utente e alle fonti di dati associate. OpenSearch

Questo argomento presenta le step-by-step istruzioni per configurare l'uso di SAML con l'interfaccia utente. OpenSearch In queste procedure, utilizziamo come esempio i passaggi per configurare

l'applicazione Okta per la gestione delle identità e degli accessi. I passaggi di configurazione per altri provider di identità, come Azure Active Directory e Ping, sono simili.

Argomenti

- Passaggio 1: configurare l'applicazione del provider di identità (Okta)
- Passaggio 2: configura la configurazione per Okta AWS
- Fase 3: Creare la policy di accesso ad Amazon OpenSearch Service in IAM
- Passaggio 4: verifica l'esperienza Single Sign-On avviata dal provider di identità con SAML
- Passaggio 5: configura il controllo granulare degli accessi basato sugli attributi SAML

Passaggio 1: configurare l'applicazione del provider di identità (Okta)

Per utilizzare SAML con l' OpenSearch interfaccia utente, il primo passaggio consiste nel configurare il provider di identità.

Attività 1: creare utenti Okta

1. Accedi alla tua organizzazione Okta all'indirizzo https://login.okta.com/come utente con privilegi amministrativi.
2. Nella console di amministrazione, in Directory nel pannello di navigazione, scegli Persone.
3. Scegli Add person (Aggiungi persona).
4. In First Name, inserisci il nome dell'utente.
5. Per Cognome, inserisci il cognome dell'utente.
6. Per Nome utente, inserisci il nome utente dell'utente in formato e-mail.
7. Scegli Imposterò la password e inserisci una password
8. (Facoltativo) Deseleziona la casella L'utente deve cambiare la password al primo accesso se non desideri che l'utente modifichi la password al primo accesso.
9. Scegli Save (Salva).

Attività 2: Creare e assegnare gruppi

1. Accedi alla tua organizzazione Okta all'indirizzo https://login.okta.com/come utente con privilegi amministrativi.
2. Nella console di amministrazione, in Directory nel pannello di navigazione, scegli Gruppi.

3. Scegliere Add Group (Aggiungi gruppo).
4. Inserisci il nome di un gruppo e scegli Salva.
5. Scegli il gruppo appena creato, quindi scegli Assegna persone.
6. Scegli il segno più (+), quindi scegli Fine.
7. (Facoltativo) Ripeti i passaggi da 1 a 6 per aggiungere altri gruppi.

Attività 3: Creare applicazioni Okta

1. Accedi alla tua organizzazione Okta all'indirizzo https://login.okta.com/come utente con privilegi amministrativi.
2. Nella console di amministrazione, in Applicazioni nel pannello di navigazione, scegli Applicazioni.
3. Scegli Create App Integration (Crea integrazione app).
4. Scegli SAML 2.0 come metodo di accesso, quindi scegli Avanti.
5. Inserisci un nome per l'integrazione dell'app (ad esempio,**`OpenSearch_UI`**), quindi scegli Avanti.
6. Inserisci i seguenti valori nell'app; non è necessario modificare altri valori:
    a. 1. Per l'URL Single Sign On, inserisci **`https://signin.aws.amazon.com/saml`** AWS le aree commerciali o l'URL specifico della tua regione.
    b. 2. Per Audience URI (SP Entity ID), inserisci**`urn:amazon:webservices`**.
    c. 3. Per il formato Name ID, immettete**`EmailAddress`**.
7. Scegli Next (Successivo).
8. Scegli Sono un cliente Okta e aggiungo un'app interna, quindi scegli Questa è un'app interna che abbiamo creato.
9. Scegli Fine.
10. Scegli Assegnazioni, quindi scegli Assegna.
11. Scegli Assegna ai gruppi, quindi seleziona Assegna accanto ai gruppi che desideri aggiungere.
12. Seleziona Fatto.

Attività 4: configurare la configurazione avanzata di Okta

Dopo aver creato l'applicazione SAML personalizzata, completa i seguenti passaggi:

1. Accedi alla tua organizzazione Okta all'indirizzo https://login.okta.com/come utente con privilegi amministrativi.

Nella console dell'amministratore, nell'area Generale, scegli Modifica nelle impostazioni SAML.

2. Scegli Next (Successivo).
3. Imposta lo stato di inoltro predefinito sull'endpoint dell' OpenSearch interfaccia utente, utilizzando il formato:

   `https://region.console.aws.amazon.com/aos/home?region=region#opensearch/applications/application-id/redirectToDashboardURL`.

   Di seguito è riportato un esempio:

   `https://us-east-2.console.aws.amazon.com/aos/home?region=us-east-2#opensearch/applications/abc123def4567EXAMPLE/redirectToDashboardURL`
4. In Attribute Statements (opzionale), aggiungi le seguenti proprietà:

   a. Fornisci il ruolo IAM e il provider di identità in formato separato da virgole utilizzando l'attributo Role. Utilizzerai lo stesso ruolo IAM e provider di identità in una fase successiva durante la configurazione. AWS

   b. Imposta user.login per. RoleSessionName Viene utilizzato come identificatore per le credenziali temporanee emesse quando si assume il ruolo.

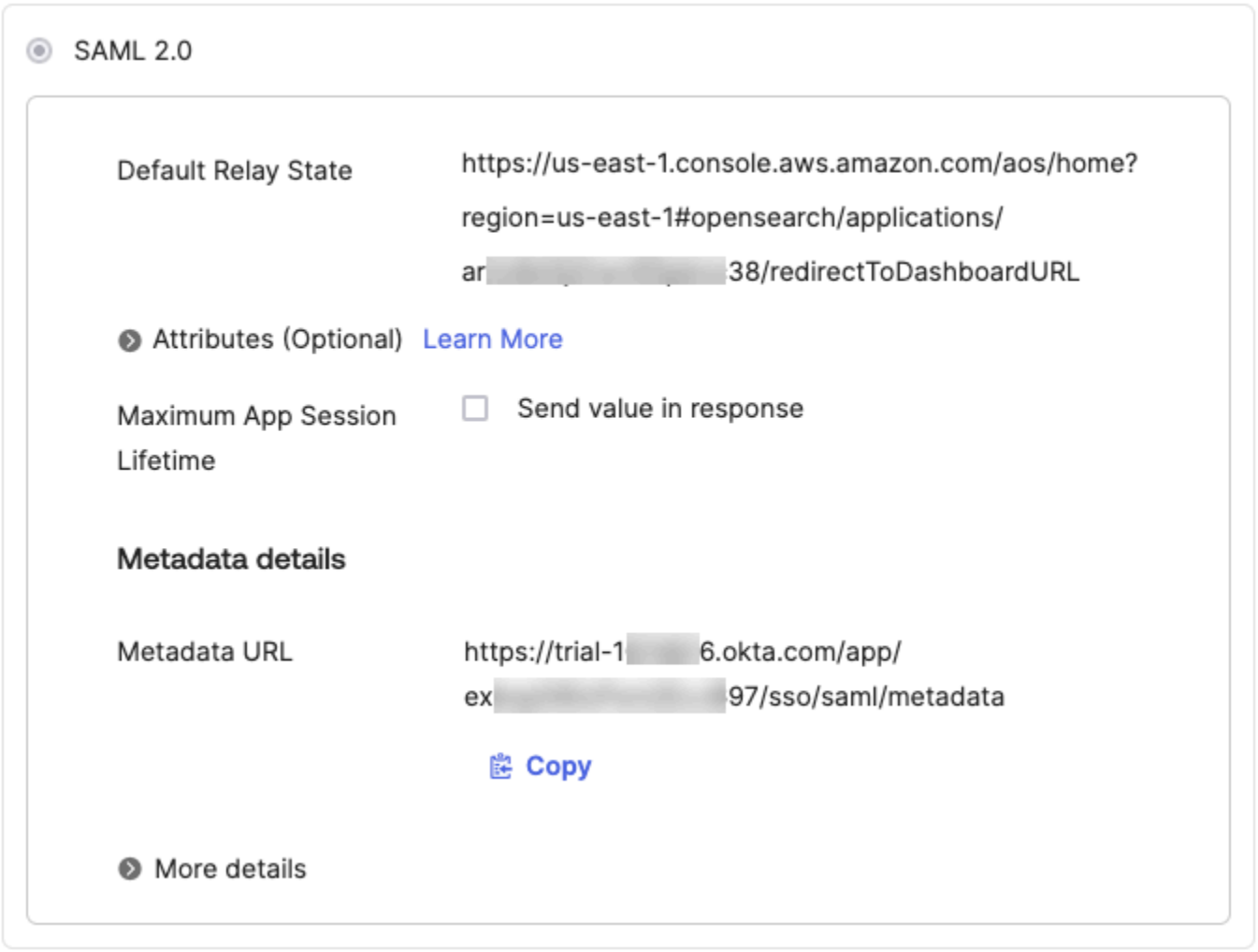
   Per riferimento:

Nome	Formato del nome	Formato	Esempio
`https://aws.amazon.com/SAML/Attributes/Role`	Imprecisata	`arn:aws:iam:: aws-account-id :role/role-name,arn:aws:iam:: aws-account-id :saml-provider/ provider-name`	`arn:aws:iam::111222333444:role/oktarole,arn:aws:iam::111223334 44:saml-provider/oktaidp`

Nome	Formato del nome	Formato	Esempio
`https://aws.amazon.com/SAML/Attributes/RoleSessionName`	Imprecisata	`user.login`	`user.login`

5. Dopo aver aggiunto le proprietà dell'attributo, scegliete Avanti, quindi Fine.

Gli attributi devono avere un formato simile a quelli mostrati nell'immagine seguente. Il valore Default Relay State è l'URL per definire la pagina di destinazione per gli utenti finali dell'account o dell'organizzazione dopo aver completato la convalida del Single Sign-On di Okta. Puoi impostarlo su qualsiasi pagina dell' OpenSearch interfaccia utente e quindi fornire quell'URL agli utenti finali previsti.



Passaggio 2: configura la configurazione per Okta AWS

Completa le seguenti attività per configurare la AWS configurazione per Okta.

Attività 1: raccogliere informazioni su Okta

Per questo passaggio, dovrai raccogliere le tue informazioni su Okta in modo da poterle configurare in seguito. AWS

1. Accedi alla tua organizzazione Okta all'indirizzo https://login.okta.com/come utente con privilegi amministrativi.
2. Nella scheda Accedi, nell'angolo in basso a destra della pagina, scegli Visualizza le istruzioni di configurazione SAML.

3. Prendi nota del valore dell'URL Single Sign-on di Identity Provider. È possibile utilizzare questo URL per la connessione a qualsiasi client SQL di terze parti come SQL Workbench/J.
4. Utilizza i metadati del provider di identità nel blocco 4, quindi salva il file di metadati in formato.xml (ad esempio,). `metadata.xml`

Attività 2: creare il provider IAM

Per creare il tuo provider IAM, completa i seguenti passaggi:

1. Accedi Console di gestione AWS e apri la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro di navigazione, in Gestione degli accessi, scegli Provider di identità.
3. Scegli Aggiungi provider.
4. Per il tipo di provider, seleziona SAML.
5. Per Nome del provider, inserisci un nome.
6. Per il documento di metadati, scegli Scegli file e carica il file di metadati (.xml) che hai scaricato in precedenza.
7. Scegli Aggiungi provider.

Attività 3: creare un ruolo IAM

Per creare il tuo AWS Identity and Access Management ruolo, completa i seguenti passaggi:

1. Accedi Console di gestione AWS e apri la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel pannello di navigazione, in Gestione degli accessi, scegli Ruoli.
3. Scegli Crea ruolo.
4. Per il tipo di entità affidabile, seleziona la federazione SAML 2.0.
5. Per un provider basato su SAML 2.0, scegli il provider di identità che hai creato in precedenza.
6. Seleziona Consenti accesso programmatico e. Console di gestione AWS
7. Scegli Next (Successivo).
8. Nell'elenco delle politiche di autorizzazione, seleziona le caselle di controllo relative alle politiche che concedono le autorizzazioni OpenSearch del servizio, AWS ad esempio le politiche gestite. AmazonOpenSearchServiceFullAccess

9. Scegli Next (Successivo).
10. Nell'area Revisione, per Nome ruolo, inserisci il nome del tuo ruolo, ad esempio. **`oktarole`**
11. (Facoltativo) In Descrizione, inserisci una breve descrizione dello scopo del ruolo.
12. Scegli Crea ruolo.
13. Passa al ruolo che hai appena creato, scegli la scheda Relazioni di fiducia, quindi scegli Modifica politica di fiducia.
14. Nel riquadro Modifica dichiarazione, in Aggiungi azioni per STS, seleziona la casella per TagSession.
15. Scegli Aggiorna policy.

Fase 3: Creare la policy di accesso ad Amazon OpenSearch Service in IAM

Scopri come configurare i tuoi ruoli IAM per il controllo degli OpenSearch accessi. Con i ruoli IAM, puoi implementare un controllo granulare degli accessi per consentire ai gruppi di utenti Okta di accedere alle risorse. OpenSearch Questo argomento illustra la configurazione basata sui ruoli IAM utilizzando due gruppi di esempio.

Sample group: Alice

Richiesta:

```
GET _plugins/_security/api/roles/alice-group
```

Risultato:

```
{
  "alice-group": {
    "reserved": false,
    "hidden": false,
    "cluster_permissions": [
      "unlimited"
    ],
    "index_permissions": [
      {
        "index_patterns": [
          "alice*"
        ],
        "dls": "",
```

```
        "fls": [],
        "masked_fields": [],
        "allowed_actions": [
          "indices_all"
        ]
      }
    ],
    "tenant_permissions": [
      {
        "tenant_patterns": [
          "global_tenant"
        ],
        "allowed_actions": [
          "kibana_all_write"
        ]
      }
    ],
    "static": false
  }
}
```

Sample group: Bob

Richiesta:

```
GET _plugins/_security/api/roles/bob-group
```

Risultato:

```
{
  "bob-group": {
    "reserved": false,
    "hidden": false,
    "cluster_permissions": [
      "unlimited"
    ],
    "index_permissions": [
      {
        "index_patterns": [
          "bob*"
        ],
        "dls": "",
        "fls": [],
```

```
        "masked_fields": [],
        "allowed_actions": [
          "indices_all"
        ]
      }
    ],
    "tenant_permissions": [
      {
        "tenant_patterns": [
          "global_tenant"
        ],
        "allowed_actions": [
          "kibana_all_write"
        ]
      }
    ],
    "static": false
  }
}
```

Puoi mappare i ruoli del dominio Amazon OpenSearch Service ai ruoli IAM utilizzando la mappatura dei ruoli di backend, come dimostrato nell'esempio seguente:

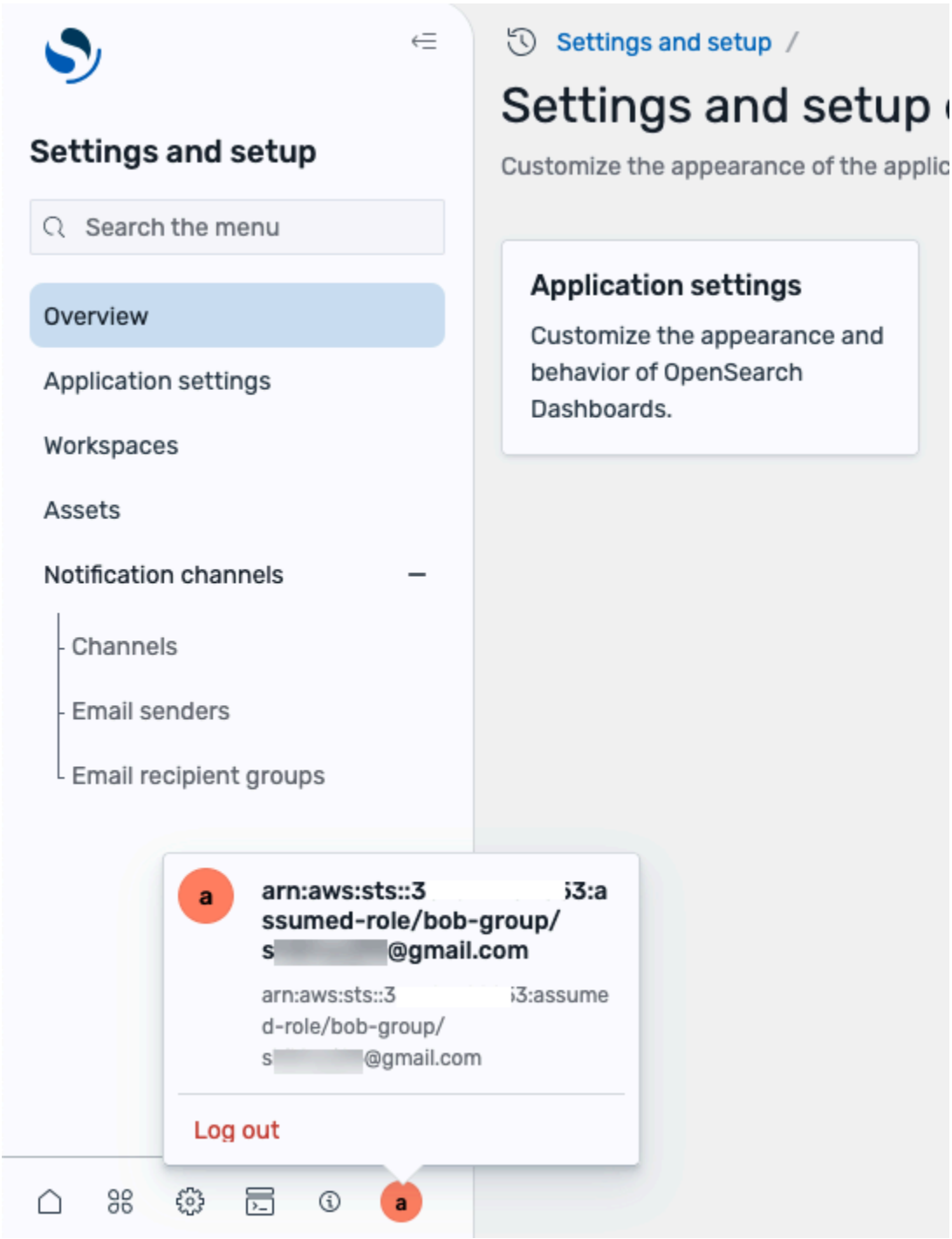
```
{
  "bob-group": {
    "hosts": [],
    "users": [],
    "reserved": false,
    "hidden": false,
    "backend_roles": [
      "arn:aws:iam::111222333444:role/bob-group"
    ],
    "and_backend_roles": []
  },
  "alice-group": {
    "hosts": [],
    "users": [],
    "reserved": false,
    "hidden": false,
    "backend_roles": [
      "arn:aws:iam::111222333444:role/alice-group"
    ],
```

```
    "and_backend_roles": []
  }
}
```

Passaggio 4: verifica l'esperienza Single Sign-On avviata dal provider di identità con SAML

Apri l'URL per Default Relay State per aprire la pagina di autenticazione Okta. Inserisci le credenziali di un utente finale. Verrai reindirizzato automaticamente all'interfaccia utente. OpenSearch

Puoi verificare le tue credenziali correnti scegliendo l'icona utente nella parte inferiore del pannello di navigazione, come illustrato nell'immagine seguente:



Puoi anche verificare le autorizzazioni di controllo degli accessi dettagliate per l'utente accedendo agli Strumenti per sviluppatori nella parte inferiore del pannello di navigazione ed eseguendo query nella console. Di seguito sono riportati alcuni esempi di interrogazioni.

Example 1: Displays information about the current user

Richiesta:

```
GET _plugins/_security/api/account
```

Risultato:

```
{
  "user_name": "arn:aws:iam::XXXXXXXXXXXX:role/bob-group",
  "is_reserved": false,
  "is_hidden": false,
  "is_internal_user": false,
  "user_requested_tenant": null,
  "backend_roles": [
    "arn:aws:iam::XXXXXXXXXXXX:role/bob-group"
  ],
  "custom_attribute_names": [],
  "tenants": {
    "global_tenant": true,
    "arn:aws:iam::XXXXXXXXXXXX:role/bob-group": true
  },
  "roles": [
    "bob-group"
  ]
}
```

Example 2: Displays actions permitted for a user

Richiesta:

```
GET bob-test/_search
```

Risultato:

```
{
  "took": 390,
  "timed_out": false,
  "_shards": {
    "total": 5,
    "successful": 5,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": {
```

```
      "value": 1,
      "relation": "eq"
    },
    "max_score": 1,
    "hits": [
      {
        "_index": "bob-test",
        "_id": "ui01N5UBCIHpjO8Jlvfy",
        "_score": 1,
        "_source": {
          "title": "Your Name",
          "year": "2016"
        }
      }
    ]
  }
}
```

Example 3: Displays actions not permitted for a user

Richiesta:

```
GET alice-test
```

Risultato:

```
{
  "error": {
    "root_cause": [
      {
        "type": "security_exception",
        "reason": "no permissions for [indices:admin/get]
 and User [name=arn:aws:iam::111222333444:role/bob-group,
 backend_roles=[arn:aws:iam::111222333444:role/bob-group], requestedTenant=null]"
      }
    ],
    "type": "security_exception",
    "reason": "no permissions for [indices:admin/get]
 and User [name=arn:aws:iam::111222333444:role/bob-group,
 backend_roles=[arn:aws:iam::111222333444:role/bob-group], requestedTenant=null]"
  },
  "status": 403
}
```

Passaggio 5: configura il controllo granulare degli accessi basato sugli attributi SAML

Con Amazon OpenSearch Service, puoi utilizzare il controllo granulare degli accessi con SAML per mappare utenti e gruppi dal tuo provider di identità a utenti e ruoli di controllo degli accessi OpenSearch granulari in modo dinamico. Puoi assegnare questi ruoli a OpenSearch domini specifici e raccolte serverless e definire autorizzazioni a livello di indice e sicurezza a livello di documento.

> **Note**
>
> Per ulteriori informazioni sul controllo granulare degli accessi, consulta. the section called “Controllo granulare degli accessi”

Argomenti

- Attributi SAML per un controllo granulare degli accessi
- Attività 1: configura Okta per un controllo granulare degli accessi
- Attività 2: configurare SAML nel dominio OpenSearch
- Attività 3: configurare SAML nelle raccolte OpenSearch Serverless

Attributi SAML per un controllo granulare degli accessi

Soggetto chiave

Si associa a un attributo utente univoco, ad esempio e-mail o nome utente, che identifica l'utente per l'autenticazione.

RuolesKey

Si associa agli attributi di gruppo o ruolo nel tuo IdP che determinano i ruoli o le autorizzazioni per l'autorizzazione.

Attività 1: configura Okta per un controllo granulare degli accessi

Per configurare Okta per un controllo granulare degli accessi

1. Aggiungi un nuovo attributo per l' OpenSearch utente principale nella sezione Attribute Statements:

- Valore: `UserName`
- Valore: `${user-email}`

Questo attributo viene utilizzato come chiave Subject nella configurazione OpenSearch granulare del controllo degli accessi per l'autenticazione.

2. Aggiungi un attributo di gruppo per i ruoli nella sezione Group Attribute Statement:

   - Valore: `groups`
   - Filtro: `OpenSearch_xxx`

   Questo attributo viene utilizzato come chiave Role per mappare i gruppi a ruoli di controllo degli accessi OpenSearch dettagliati per l'autorizzazione.

Attività 2: configurare SAML nel dominio OpenSearch

Per configurare SAML nel dominio OpenSearch

1. Nella console di AWS gestione, identifica il dominio di OpenSearch servizio per il quale desideri abilitare il controllo di accesso granulare per gli utenti dell'interfaccia utente. OpenSearch
2. Vai alla pagina dei dettagli del dominio specifico.
3. Seleziona la scheda Configurazione di sicurezza e fai clic su Modifica.
4. Espandi SAML tramite IAM Federate.
5. Inserisci `subjectKey` e `roleKey` che hai definito in Okta.
6. Seleziona Salva modifiche.

Puoi anche configurare un controllo granulare degli accessi utilizzando. AWS CLI

```
aws opensearch create-domain \
--domain-name testDomain \
--engine-version OpenSearch_1.3 \
--cluster-config
 InstanceType=r5.xlarge.search,InstanceCount=1,DedicatedMasterEnabled=false,ZoneAwarenessEnable
 \
```

```
--access-policies '{"Version": "2012-10-17",       "Statement":
[{"Effect":"Allow","Principal":{"AWS":"*"},"Action":"es:*","Resource":"arn:aws:es:us-
east-1:12345678901:domain/neosaml10/*"}]}' \
--domain-endpoint-options '{"EnforceHTTPS":true,"TLSSecurityPolicy":"Policy-Min-
TLS-1-2-2019-07"}' \
--node-to-node-encryption-options '{"Enabled":true}' \
--encryption-at-rest-options '{"Enabled":true}' \
--advanced-security-options
 '{"Enabled":true,"InternalUserDatabaseEnabled":true,"MasterUserOptions":
{"MasterUserName":"********","MasterUserPassword":"********"}, "IAMFederationOptions":
{"Enabled": true,"SubjectKey":"TestSubjectKey","RolesKey":"TestRolesKey"}}' \
--ebs-options "EBSEnabled=true,VolumeType=gp2,VolumeSize=300" \
--no-verify-ssl \
--endpoint-url https://es.us-east-1.amazonaws.com \
--region us-east-1
```

Per aggiornare un dominio esistente:

```
aws opensearch update-domain-config \
--domain-name testDomain \
--advanced-security-options
 '{"Enabled":true,"InternalUserDatabaseEnabled":true,"MasterUserOptions":
{"MasterUserName":"********","MasterUserPassword":"********"}, "IAMFederationOptions":
{"Enabled": true,"SubjectKey":"TestSubjectKey","RolesKey":"TestRolesKey"}}' \
--ebs-options "EBSEnabled=true,VolumeType=gp2,VolumeSize=300" \
--no-verify-ssl \
--endpoint-url https://es.us-east-1.amazonaws.com \
--region us-east-1
```

Attività 3: configurare SAML nelle raccolte OpenSearch Serverless

Per configurare il controllo granulare degli accessi basato su SAML in Serverless OpenSearch

1. Apri Console di gestione AWS e accedi al OpenSearch servizio Amazon.
2. Nel pannello di navigazione, in Serverless, scegli Sicurezza, quindi scegli Autenticazione.
3. Nella sezione IAM Federation, seleziona Modifica.

   Puoi controllare il controllo granulare degli accessi basato sugli attributi SAML utilizzando questa configurazione. IAM Federation è disabilitata per impostazione predefinita.
4. Seleziona Abilita IAM Federation.
5. Inserisci i `roleKey` valori `subjectKey` and che hai definito in Okta.

Per ulteriori informazioni, consulta the section called “Attributi SAML per un controllo granulare degli accessi”.

6. Seleziona Salva.
7. Nel riquadro di navigazione sotto Serverless, scegli Politica di accesso ai dati.
8. Aggiorna una politica esistente o creane una nuova.
9. Espandi una regola, scegli Aggiungi principi, quindi seleziona Utenti e gruppi della Federazione IAM.
10. Aggiungi i principali richiesti e scegli Salva.
11. Scegliere Concedi.
12. In base a questa regola, procedi come segue:
    - Seleziona le autorizzazioni che desideri definire per i principali selezionati.
    - Specificate le raccolte a cui desiderate applicare le autorizzazioni.
    - Facoltativamente, definisci le autorizzazioni a livello di indice.

    > **Note**
    >
    > È possibile creare più regole per assegnare autorizzazioni diverse a diversi gruppi di responsabili.

13. Al termine, scegli Save (Salva).
14. Scegli Create (Crea).

In alternativa, puoi utilizzare la CLI per creare le configurazioni di sicurezza per le raccolte, come indicato di seguito:

```
aws opensearchserverless create-security-config --region "region"  --type iamfederation
 --name "configuration_name" --description "description" --iam-federation-options
 '{"groupAttribute":"GroupKey","userAttribute":"UserKey"}'
```

Gestione delle associazioni delle fonti di dati e delle autorizzazioni di accesso al Virtual Private Cloud

Utilizza le procedure in questa sezione per gestire le associazioni delle fonti di dati e configurare le autorizzazioni di accesso necessarie per un cloud privato virtuale (VPC).

Argomenti

- Associazione di una fonte di dati a un'applicazione UI OpenSearch
- Gestione dell'accesso ai domini in un VPC
- Configurazione dell'accesso alle raccolte OpenSearch Serverless in un VPC

Associazione di una fonte di dati a un'applicazione UI OpenSearch

Dopo aver creato un'applicazione OpenSearch UI, puoi utilizzare la console o associarla AWS CLI a una o più fonti di dati. Dopodiché, gli utenti finali possono recuperare i dati da queste fonti di dati per effettuare ricerche, lavorare con i dashboard e così via.

Associa un'origine dati a un'applicazione OpenSearch UI (console)

Per associare un'origine dati a un'applicazione OpenSearch UI utilizzando la console

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Scegli OpenSearch UI (dashboard), quindi scegli il nome di un'applicazione OpenSearch UI.
3. Nell'area Origini dati associate, scegli Gestisci fonti di dati.
4. Scegli tra i OpenSearch domini e le raccolte che desideri associare all'applicazione.

> **Tip**
>
> Se non riesci a trovare le fonti di dati che stai cercando, contatta i tuoi amministratori per concederti le autorizzazioni necessarie. Per ulteriori informazioni, consulta the section called “Autorizzazioni per creare un'applicazione che utilizza l'autenticazione IAM Identity Center (opzionale)”.

5. Scegli Avanti, quindi scegli Salva.

Dopo aver associato un'origine dati all'applicazione, il pulsante Avvia applicazione viene abilitato nella pagina dei dettagli dell'applicazione. Puoi scegliere Launch Application per aprire la OpenSearch pagina Benvenuto, in cui puoi creare e gestire aree di lavoro.

Per informazioni sull'utilizzo delle aree di lavoro, consulta. the section called “Utilizzo delle aree di lavoro OpenSearch di Amazon Service”

Gestione dell'accesso ai domini in un VPC

Se all'applicazione era associato un OpenSearch dominio in un VPC, un amministratore VPC deve autorizzare l'accesso tra l'interfaccia utente OpenSearch e il VPC utilizzando la console o. AWS CLI

Gestione dell'accesso ai domini in un VPC (console)

Per configurare l'accesso a un dominio VPC utilizzando: Console di gestione AWS

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli Domini, quindi scegli il nome del dominio VPC.

   oppure

   Scegli Crea dominio, quindi configura i dettagli per il dominio.
3. Scegli la scheda Endpoint VPC, quindi scegli Autorizza principale.
4. Nella finestra di dialogo Autorizza i principali, seleziona Autorizza i principali da altri AWS servizi, quindi scegli OpenSearch le applicazioni (Dashboard) dall'elenco.
5. Seleziona Authorize (Autorizza).

Gestione dell'accesso ai domini in un AWS CLI VPC ()

Per autorizzare un dominio VPC utilizzando AWS CLI

Per autorizzare il dominio VPC utilizzando, eseguire AWS CLI il comando seguente. Sostituisci *placeholder values* con le informazioni appropriate.

```
aws opensearch authorize-vpc-endpoint-access \
 --domain-name domain-name \
 --service application.opensearchservice.amazonaws.com \
 --region region-id
```

Per revocare un'associazione di dominio VPC utilizzando la console

Quando un'associazione non è più necessaria, il proprietario del dominio VPC può revocare l'accesso utilizzando la procedura seguente.

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nel riquadro di navigazione a sinistra, scegli Domini, quindi scegli il nome del dominio VPC.
3. Scegli la scheda Endpoint VPC, quindi seleziona il pulsante per la riga OpenSearch Applicazioni (Dashboard).
4. Scegli Revoca accesso.

Per revocare un'associazione di dominio VPC utilizzando AWS CLI

Per revocare un'associazione di dominio VPC con l'applicazione OpenSearch UI, esegui il comando seguente. Sostituisci *placeholder values* con le informazioni appropriate.

```
aws opensearch revoke-vpc-endpoint-access \
    --domain-name domain-name \
    --service application.opensearchservice.amazonaws.com \
    --region region-id
```

Configurazione dell'accesso alle raccolte OpenSearch Serverless in un VPC

Se all'applicazione è stata associata una raccolta Amazon OpenSearch Serverless in un VPC, un amministratore VPC può autorizzare l'accesso creando una nuova policy di rete e allegandola alla raccolta.

Configurazione dell'accesso alle raccolte OpenSearch Serverless in un VPC (console)

Per configurare l'accesso alle raccolte OpenSearch Serverless in un VPC utilizzando la console

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nella barra di navigazione a sinistra, scegli Politiche di rete, scegli il nome della politica di rete, quindi scegli Modifica.

   oppure

Scegli Crea politica di rete, quindi configura i dettagli della politica.

3. Nell'area Tipo di accesso, scegli Privato (consigliato), quindi seleziona Accesso privato al AWS servizio.
4. Nel campo di ricerca, scegli Servizio, quindi scegli`application.opensearchservice.amazonaws.com`.
5. Nell'area Tipo di risorsa, seleziona la casella Abilita l'accesso all' OpenSearch endpoint.
6. Per Cerca raccolte, o inserisci termini di prefisso specifici, nel campo di ricerca, seleziona Nome raccolta, quindi inserisci o seleziona il nome delle raccolte da associare alla politica di rete.
7. Scegli Crea per una nuova politica di rete o Aggiorna per una politica di rete esistente.

Configurazione dell'accesso alle raccolte OpenSearch Serverless in un VPC ()AWS CLI

Per configurare l'accesso alle raccolte OpenSearch Serverless in un VPC utilizzando AWS CLI

1. Crea un file.json simile al seguente. Sostituisci *`placeholder values`* con le informazioni appropriate.

```
{
    "Description" : "policy-description",
    "Rules": [{
       "ResourceType" : "collection",
        "Resource" : ["collection/collection_name"]
     }],
    "SourceServices" : [
          "application.opensearchservice.amazonaws.com"
      ],
      "AllowFromPublic" : false
}
```

2. Crea o aggiorna una policy di rete per una raccolta in un VPC per lavorare con le applicazioni OpenSearch dell'interfaccia utente.

   Create a network policy

   Eseguire il seguente comando seguente. Sostituisci *`placeholder values`* con le informazioni appropriate.

```
aws opensearchserverless create-security-policy \
    --type network  \
    --region region \
    --endpoint-url endpoint-url \
    --name network-policy-name \
    --policy file:/path_to_network_policy_json_file
```

Questo comando restituisce informazioni simili alle seguenti:

```
{
    "securityPolicyDetail": {
        "createdDate": ******,
        "lastModifiedDate": ******,
        "name": "network-policy-name",
        "policy": [
            {
                "SourceVPCEs": [],
                "AllowFromPublic": false,
                "Description": "",
                "Rules": [
                    {
                        "Resource": [
                            "collection/network-policy-name"
                        ],
                        "ResourceType": "collection"
                    }
                ],
                "SourceServices": [
                    "application.opensearchservice.amazonaws.com"
                ]
            }
        ],
        "policyVersion": "******",
        "type": "network"
    }
}
```

Update a network policy

Eseguire il seguente comando seguente. Sostituisci *placeholder values* con le informazioni appropriate.

```
aws opensearchserverless update-security-policy \
    --type network  \
    --region region \
    --endpoint-url endpoint-url \
    --name network-policy-name \
    --policy-version "policy_version_from_output_of_network_policy_creation" \
    --policy file:/path_to_network_policy_json_file
```

Questo comando restituisce informazioni simili alle seguenti:

```
{
    "securityPolicyDetail": {
        "createdDate": ******,
        "lastModifiedDate": ******,
        "name": "network-policy-name",
        "policy": [
            {
                "SourceVPCEs": [],
                "AllowFromPublic": false,
                "Description": "",
                "Rules": [
                    {
                        "Resource": [
                            "collection/network-policy-name"
                        ],
                        "ResourceType": "collection"
                    }
                ],
                "SourceServices": [
                    "application.opensearchservice.amazonaws.com"
                ]
            }
        ],
        "policyVersion": "******",
        "type": "network"
    }
}
```

Utilizzo delle aree di lavoro OpenSearch di Amazon Service

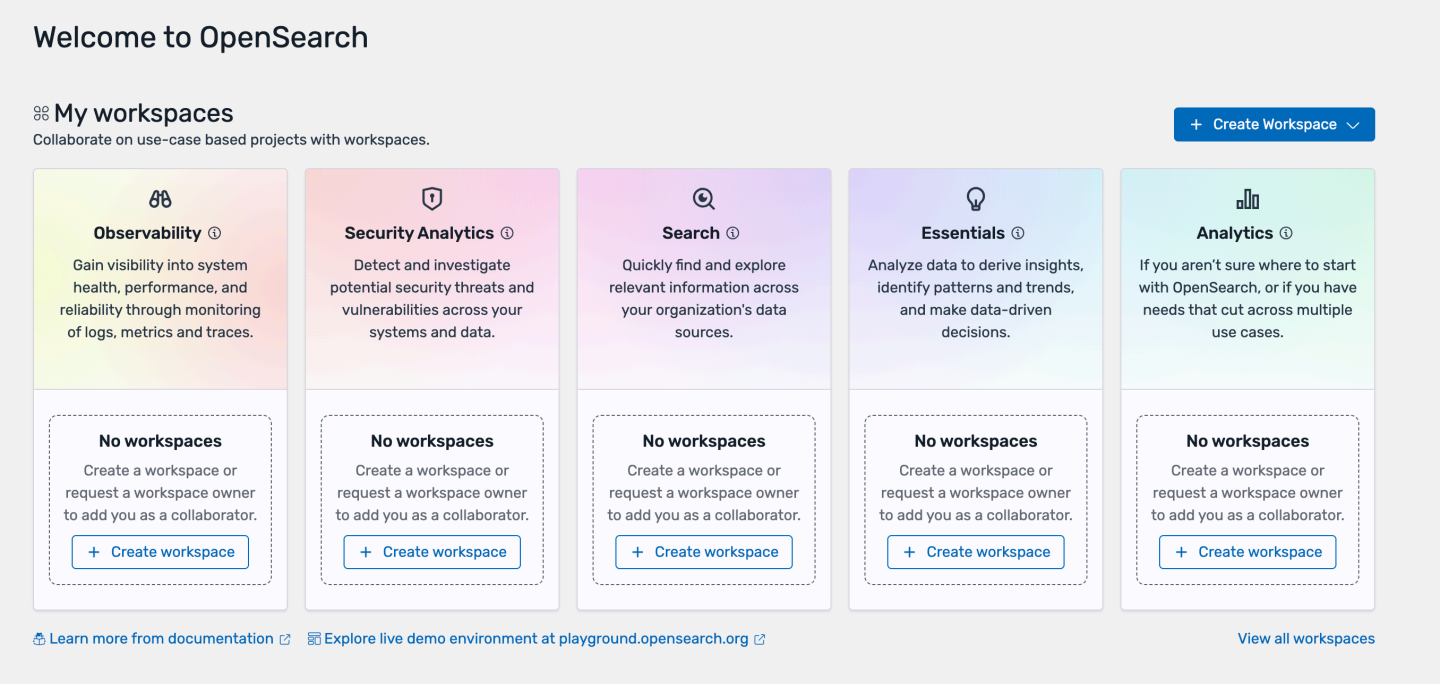
Amazon OpenSearch Service supporta la creazione di più aree di lavoro specifiche per ogni caso d'uso. Ogni area di lavoro offre un'esperienza curata per casi d'uso comuni come Observability, Security Analytics e Search. Workspace supporta anche la gestione dei collaboratori, in modo da poter condividere l'area di lavoro solo con i collaboratori previsti e gestire le autorizzazioni per ciascuno di essi.

OpenSearch Creazione di aree di lavoro per applicazioni UI

Dopo aver creato e associato un'applicazione OpenSearch UI alle fonti di dati e aver configurato le autorizzazioni utente per l'applicazione, puoi avviare l'applicazione OpenSearch UI per creare aree di lavoro.

Per iniziare a creare un'area di lavoro, potete selezionare il pulsante Avvia applicazione nella pagina dei dettagli dell'applicazione o utilizzare l'URL dell'applicazione OpenSearch UI per aprire la home page dell'applicazione OpenSearch UI in una nuova finestra del browser.

L'applicazione OpenSearch UI offre opzioni per la creazione di aree di lavoro ed elenca tutte le aree di lavoro esistenti nella home page, classificate per caso d'uso.



Per ulteriori informazioni sui tipi di aree di lavoro supportati, consulta. the section called “Tipi di aree di lavoro”

Privacy dello spazio di lavoro e collaboratori

Puoi definire un'impostazione di privacy per un'area di lavoro come livello di autorizzazione predefinito per tutti gli utenti. Puoi farlo durante la creazione di un'area di lavoro o modificare un'area di lavoro esistente (nella scheda Collaboratori dell'area di lavoro). Esistono tre opzioni di privacy tra cui scegliere:

- Privato per i collaboratori: solo i collaboratori che aggiungi esplicitamente all'area di lavoro possono accedere all'area di lavoro. Puoi definire i livelli di autorizzazione per ogni collaboratore.
- Tutti possono visualizzare: chiunque abbia accesso all'applicazione OpenSearch UI può accedere all'area di lavoro e visualizzarne le risorse, ma non può apportare modifiche all'area di lavoro.
- Chiunque può modificare: chiunque abbia accesso all'applicazione OpenSearch UI può accedere all'area di lavoro, visualizzare le risorse al suo interno e apportare modifiche alle risorse nell'area di lavoro.

Nella scheda Collaboratori dello spazio di lavoro, puoi aggiungere utenti o ruoli IAM e AWS IAM Identity Center utenti come collaboratori in uno spazio di lavoro. I collaboratori possono scegliere tra tre livelli di autorizzazioni:

- Sola lettura: il collaboratore può solo visualizzare le risorse nell'area di lavoro. Questa impostazione viene sostituita se l'area di lavoro è configurata per utilizzare l'impostazione sulla privacy Chiunque può modificare.
- Lettura e scrittura: il collaboratore può visualizzare e modificare le risorse nell'area di lavoro. Se l'area di lavoro è configurata per utilizzare l'impostazione sulla privacy Chiunque può visualizzare, il collaboratore può comunque apportare modifiche.
- Amministratore: il collaboratore può aggiornare le impostazioni ed eliminare l'area di lavoro. Il collaboratore può anche modificare le impostazioni sulla privacy dell'area di lavoro e gestire i collaboratori. L'utente che crea l'area di lavoro viene automaticamente assegnato come amministratore dell'area di lavoro.

Tipi di aree di lavoro

Amazon OpenSearch Service offre cinque tipi di spazio di lavoro, ognuno con caratteristiche diverse per i diversi casi d'uso:

- L'area di lavoro Observability è progettata per ottenere visibilità sullo stato, le prestazioni e l'affidabilità del sistema attraverso il monitoraggio di log, metriche e tracce.
- L'area di lavoro Security Analytics è progettata per rilevare e indagare su potenziali minacce e vulnerabilità alla sicurezza nei sistemi e nei dati.
- L'area di lavoro Search è progettata per trovare ed esplorare rapidamente le informazioni pertinenti tra le fonti di dati dell'organizzazione.
- L'area di lavoro Essentials è progettata per OpenSearch Serverless come fonte di dati e consente di analizzare i dati per ricavare informazioni, identificare modelli e tendenze e prendere rapidamente decisioni basate sui dati. Puoi trovare ed esplorare le informazioni pertinenti nelle fonti di dati della tua organizzazione in un'area di lavoro Essentials.
- L'area di lavoro Analytics (tutte le funzionalità) è progettata per casi d'uso multiuso e supporta tutte le funzionalità disponibili nell'interfaccia utente di OpenSearch servizio (dashboard).

Accesso ai dati tra regioni e account diversi

OpenSearch L'interfaccia utente supporta l'accesso ai dati da OpenSearch domini su diversi server Account AWS. Regione AWS Puoi scegliere tra due approcci a seconda delle tue esigenze. La tabella seguente mette a confronto i due approcci.

> **Note**
>
> Sia l'accesso ai dati tra account che la ricerca tra cluster funzionano solo con i domini. OpenSearch Nessuno dei due approcci supporta OpenSearch le raccolte Serverless.

Aspetto	Accesso ai dati tra account	Funzionalità di ricerca tra cluster
Funzionalità	Associa i domini di altri account come fonti di dati dirette nell'interfaccia utente OpenSearch	Interroga i dati su domini connessi utilizzando connessioni di ricerca tra cluster
Meccanismo	Accesso diretto: l' OpenSearch interfaccia utente si connette direttamente al dominio di destinazione in un altro account	Accesso indiretto: richiede un dominio locale nello stesso account dell' OpenSearch interfaccia utente

Aspetto	Accesso ai dati tra account	Funzionalità di ricerca tra cluster
		per inoltrare le richieste ai domini remoti
Supporto multi-account	Sì	Sì
Supporto tra regioni	No, i domini di origine e di destinazione devono appartenere allo stesso Regione AWS	Sì, i domini di origine e di destinazione possono trovarsi in diversi Regione AWS
Unisci i dati tra i domini	No, ogni dominio viene interrogato in modo indipendente come fonte di dati separata	Sì, una singola query può aggregare i risultati di più domini connessi
Metodi di autenticazione	IAM e AWS IAM Identity Center	IAM (con controllo granulare degli accessi)
Complessità della configurazione	Inferiore: richiede un ruolo IAM su più account per la convalida	Superiore: richiede connessioni tra cluster, politiche di accesso su entrambi i domini e un controllo granulare degli accessi
Visibilità delle fonti di dati nell'interfaccia utente OpenSearch	Ogni dominio tra account viene visualizzato come una fonte di dati separata	Si accede ai domini remoti tramite gli alias di connessione del dominio di origine locale
Accesso in scrittura al dominio remoto	Sì, controllato dalla politica di accesso del dominio di destinazione	No: la ricerca tra cluster fornisce l'accesso in sola lettura ai domini remoti

Argomenti

- Accesso ai dati tra account diversi ai domini OpenSearch
- Funzionalità di ricerca tra cluster

Accesso ai dati tra account diversi ai domini OpenSearch

Puoi configurare le applicazioni OpenSearch dell'interfaccia utente in un account per accedere ai OpenSearch domini di account diversi. Quando OpenSearch crei un'applicazione UI con fonti di dati tra account diversi, fornisci un elemento `iamRoleForDataSourceArn` che rimanda a un ruolo IAM nell'account di destinazione. OpenSearch L'interfaccia utente convalida la richiesta assumendo questo ruolo e chiamando `es:DescribeDomain` per verificare l'accessibilità del dominio. Il ruolo tra account diversi viene utilizzato solo per la convalida del piano di controllo. L'accesso al piano dati è controllato separatamente dalla politica di accesso del dominio di destinazione.

> ⚠ Codice di esempio
>
> Gli esempi di codice in questo argomento sono solo a scopo illustrativo. Dimostrano funzionalità di base e potrebbero non includere la gestione degli errori, le migliori pratiche di sicurezza o le funzionalità pronte per la produzione. Prima di utilizzare il codice di esempio in produzione, esaminatelo e modificatelo per soddisfare i vostri requisiti specifici e testatelo accuratamente nel vostro ambiente.

Concetti chiave

Account di origine

: Il Account AWS che ospita la tua applicazione OpenSearch UI.

Account Target

: Il Account AWS luogo in cui risiede il OpenSearch dominio.

Ruolo tra account

: Un ruolo IAM nell'account di destinazione utilizzato solo per la convalida del piano di controllo. Questo ruolo richiede solo l'`es:DescribeDomain`autorizzazione.

Ruolo applicativo IAM Identity Center

: Un ruolo IAM nell'account di origine utilizzato per l'accesso al piano dati degli utenti di IAM Identity Center.

Prerequisiti

Prima di configurare l'accesso ai dati tra account diversi, assicurati di disporre di quanto segue:

- AWS CLI installato e configurato
- Accesso sia ai file di origine che a quelli Account AWS di destinazione
- Per i flussi di IAM Identity Center: un'istanza AWS IAM Identity Center dell'organizzazione

Scenari

Scegli lo scenario che corrisponde al tuo metodo di autenticazione e alla configurazione del dominio:

- the section called “Scenario 1: utente IAM che accede a un dominio pubblico”
- the section called “Scenario 2: utente di IAM Identity Center che accede a un dominio pubblico”
- the section called “Scenario 3: utente IAM che accede a un dominio VPC”
- the section called “Scenario 4: utente IAM Identity Center che accede a un dominio VPC”

Scenario 1: utente IAM che accede a un dominio pubblico

Fase 1: Creare il ruolo IAM per più account (account di destinazione)

Crea un ruolo IAM nell'account di destinazione che consenta all'account di origine di assumerlo per la convalida del dominio.

Per creare il ruolo tra account

1. Crea una politica di fiducia che consenta all'account di origine di assumere il ruolo:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account-id:root"
    },
    "Action": "sts:AssumeRole"
  }]
}
```

2. Crea il ruolo:

```
aws iam create-role \
```

```
  --role-name OpenSearchUIAccessRole \
  --assume-role-policy-document file://trust-policy.json
```

3. Crea una politica di autorizzazioni con la sola `es:DescribeDomain` azione:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": "es:DescribeDomain",
    "Resource": "arn:aws:es:region:target-account-id:domain/*"
  }]
}
```

4. Allega la politica delle autorizzazioni al ruolo:

```
aws iam put-role-policy \
  --role-name OpenSearchUIAccessRole \
  --policy-name ValidationOnly \
  --policy-document file://permissions-policy.json
```

Fase 2: Creare il OpenSearch dominio (account di destinazione)

Crea un OpenSearch dominio nell'account di destinazione con il controllo granulare degli accessi e la crittografia abilitati:

```
aws opensearch create-domain \
  --domain-name domain-name \
  --engine-version OpenSearch_2.19 \
  --cluster-config InstanceType=m5.large.search,InstanceCount=1 \
  --ebs-options "EBSEnabled=true,VolumeType=gp3,VolumeSize=100" \
  --advanced-security-options
 '{"Enabled":true,"InternalUserDatabaseEnabled":true,"MasterUserOptions":
{"MasterUserName":"admin","MasterUserPassword":"master-password"}}' \
  --node-to-node-encryption-options '{"Enabled":true}' \
  --encryption-at-rest-options '{"Enabled":true}' \
  --domain-endpoint-options '{"EnforceHTTPS":true,"TLSSecurityPolicy":"Policy-Min-
TLS-1-2-2019-07"}' \
  --access-policies '{"Version":"2012-10-17",       "Statement":
[{"Effect":"Allow","Principal":
{"AWS":"*"},"Action":"es:ESHttp*","Resource":"arn:aws:es:region:target-account-
id:domain/domain-name/*"}]}' \
```

```
  --region region
```

Attendi che lo stato del dominio diventi `Active` tale prima di procedere.

Fase 3: Creare l'applicazione OpenSearch UI (account di origine)

Crea l'applicazione nell'account di origine con l'origine dati tra account:

```
aws opensearch create-application \
  --region region \
  --name "cross-account-iam-app" \
  --data-sources '[{
    "dataSourceArn":"arn:aws:es:region:target-account-id:domain/domain-name",
    "dataSourceDescription":"Cross-account domain",
    "iamRoleForDataSourceArn":"arn:aws:iam::target-account-
id:role/OpenSearchUIAccessRole"
  }]' \
  --app-configs '[{"key":"opensearchDashboards.dashboardAdmin.users","value":"[\"*
\"]"}]'
```

Fase 4: Verifica e accedi

Recupera i dettagli dell'applicazione per ottenere l'URL dell'endpoint:

```
aws opensearch get-application \
  --region region \
  --id application-id
```

- Passa all'URL dell'endpoint dell'applicazione dalla risposta.
- Accedi con le credenziali IAM.
- L'utente IAM firma le richieste del piano dati con le proprie credenziali.
- La policy di accesso al dominio di destinazione controlla a quali dati l'utente può accedere.

Scenario 2: utente di IAM Identity Center che accede a un dominio pubblico

Fase 1: Creare il ruolo IAM per più account (account di destinazione)

Crea un ruolo IAM nell'account di destinazione che consenta all'account di origine di assumerlo per la convalida del dominio.

Per creare il ruolo tra account

1. Crea una politica di fiducia che consenta all'account di origine di assumere il ruolo:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account-id:root"
    },
    "Action": "sts:AssumeRole"
  }]
}
```

2. Crea il ruolo:

```
aws iam create-role \
  --role-name OpenSearchUIAccessRole \
  --assume-role-policy-document file://trust-policy.json
```

3. Crea una politica di autorizzazioni con la sola `es:DescribeDomain` azione:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": "es:DescribeDomain",
    "Resource": "arn:aws:es:region:target-account-id:domain/*"
  }]
}
```

4. Allega la politica delle autorizzazioni al ruolo:

```
aws iam put-role-policy \
  --role-name OpenSearchUIAccessRole \
  --policy-name ValidationOnly \
  --policy-document file://permissions-policy.json
```

Fase 2: Creare il OpenSearch dominio (account di destinazione)

Crea un OpenSearch dominio nell'account di destinazione. Utilizza lo stesso comandothe section called “Fase 2: Creare il OpenSearch dominio (account di destinazione)”, ma aggiorna la policy di accesso per consentire il ruolo dell'applicazione IAM Identity Center dall'account di origine:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account-id:role/NeoIdCAppRole"
    },
    "Action": "es:ESHttp*",
    "Resource": "arn:aws:es:region:target-account-id:domain/domain-name/*"
  }]
}
```

Attendi che lo stato del dominio diventi tale `Active` prima di procedere.

Fase 3: Creare il ruolo IAM per l'applicazione IAM Identity Center (account di origine)

Crea un ruolo IAM nell'account di origine che l' OpenSearch interfaccia utente utilizza per l'accesso al piano dati degli utenti di IAM Identity Center.

Per creare il ruolo applicativo IAM Identity Center

1. Crea una politica di fiducia:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "application.opensearchservice.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    },
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "application.opensearchservice.amazonaws.com"
```

```
      },
      "Action": "sts:SetContext",
      "Condition": {
        "ForAllValues:ArnEquals": {
          "sts:RequestContextProviders": "arn:aws:iam::source-account-id:oidc-
provider/portal.sso.region.amazonaws.com/apl/application-id"
        }
      }
    }
  ]
}
```

2. Crea una politica di autorizzazioni:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "OpenSearchDomain",
    "Effect": "Allow",
    "Action": ["es:ESHttp*"],
    "Resource": "*"
  }]
}
```

3. Crea il ruolo e allega le politiche:

```
aws iam create-role \
  --role-name NeoIdCAppRole \
  --assume-role-policy-document file://neoidc-trust-policy.json

aws iam put-role-policy \
  --role-name NeoIdCAppRole \
  --policy-name NeoIdCAppPermissions \
  --policy-document file://neoidc-permissions-policy.json
```

Fase 4: Creare l'applicazione OpenSearch UI con IAM Identity Center (account di origine)

```
aws opensearch create-application \
  --region region \
  --name "cross-account-idc-app" \
  --iam-identity-center-options '{
    "enabled":true,
```

```
    "iamIdentityCenterInstanceArn":"arn:aws:sso:::instance/ssoins-instance-id",
    "iamRoleForIdentityCenterApplicationArn":"arn:aws:iam::source-account-
id:role/NeoIdCAppRole"
  }' \
  --data-sources '[{
    "dataSourceArn":"arn:aws:es:region:target-account-id:domain/domain-name",
    "dataSourceDescription":"Cross-account domain",
    "iamRoleForDataSourceArn":"arn:aws:iam::target-account-
id:role/OpenSearchUIAccessRole"
  }]' \
  --app-configs '[{"key":"opensearchDashboards.dashboardAdmin.users","value":"[\"*
\"]"}]'
```

Fase 5: Creare e assegnare utenti e gruppi di IAM Identity Center

Crea un utente IAM Identity Center

Eseguire il seguente comando seguente. Sostituisci *placeholder values* con le informazioni appropriate.

```
aws identitystore create-user \
  --identity-store-id d-directory-id \
  --user-name user-email \
  --display-name "display-name" \
  --name Formatted=string,FamilyName=last-name,GivenName=first-name \
  --emails Value=user-email,Type=work,Primary=true
```

Crea un gruppo IAM Identity Center e aggiungi l'utente

Esegui i comandi seguenti:

```
aws identitystore create-group \
  --identity-store-id d-directory-id \
  --display-name "OpenSearchUsers" \
  --description "Users with OpenSearch access"

aws identitystore create-group-membership \
  --identity-store-id d-directory-id \
  --group-id group-id \
  --member-id UserId=user-id
```

Assegna l'utente o il gruppo all'applicazione

Esegui il comando seguente:

```
aws sso-admin create-application-assignment \
  --application-arn "arn:aws:sso:::source-account-id:application/ssoins-instance-id/
apl-application-id" \
  --principal-id user-id-or-group-id \
  --principal-type USER
```

Configura la mappatura dei ruoli di backend sul dominio di destinazione

Mappa il gruppo IAM Identity Center su un ruolo OpenSearch di sicurezza nel dominio di destinazione:

```
curl -XPUT "https://domain-endpoint/_plugins/_security/api/rolesmapping/all_access" \
  -u admin:master-password \
  -H 'Content-Type: application/json' \
  -d '{
    "backend_roles": ["group-id"],
    "hosts": [],
    "users": []
  }'
```

Fase 6: Verifica e accesso

```
aws opensearch get-application \
  --region region \
  --id application-id
```

- Vai all'URL dell'endpoint dell'applicazione.
- Accedi con le credenziali utente di IAM Identity Center.
- Le richieste di dati degli utenti di IAM Identity Center vengono firmate con il ruolo applicativo IAM Identity Center, non con il ruolo tra account.
- Le mappature dei ruoli di backend sul dominio controllano le autorizzazioni di accesso ai dati.

Scenario 3: utente IAM che accede a un dominio VPC

Fase 1: Creare il ruolo IAM per più account (account di destinazione)

Crea un ruolo IAM nell'account di destinazione che consenta all'account di origine di assumerlo per la convalida del dominio.

Per creare il ruolo tra account

1. Crea una politica di fiducia che consenta all'account di origine di assumere il ruolo:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account-id:root"
    },
    "Action": "sts:AssumeRole"
  }]
}
```

2. Crea il ruolo:

```
aws iam create-role \
  --role-name OpenSearchUIAccessRole \
  --assume-role-policy-document file://trust-policy.json
```

3. Crea una politica di autorizzazioni con la sola `es:DescribeDomain` azione:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": "es:DescribeDomain",
    "Resource": "arn:aws:es:region:target-account-id:domain/*"
  }]
}
```

4. Allega la politica delle autorizzazioni al ruolo:

```
aws iam put-role-policy \
  --role-name OpenSearchUIAccessRole \
  --policy-name ValidationOnly \
  --policy-document file://permissions-policy.json
```

Fase 2: Configurare il VPC (account di destinazione)

Salta questo passaggio se nell'account di destinazione esiste già un VPC.

```
# Create VPC
aws ec2 create-vpc \
  --cidr-block 10.0.0.0/16 \
  --region region

# Create subnet
aws ec2 create-subnet \
  --vpc-id vpc-id \
  --cidr-block 10.0.1.0/24 \
  --availability-zone regiona \
  --region region

# Create security group
aws ec2 create-security-group \
  --group-name opensearch-vpc-sg \
  --description "Security group for OpenSearch VPC domain" \
  --vpc-id vpc-id \
  --region region

# Allow inbound HTTPS
aws ec2 authorize-security-group-ingress \
  --group-id security-group-id \
  --protocol tcp \
  --port 443 \
  --cidr 10.0.0.0/16 \
  --region region
```

Scopri di più sulla creazione di domini VPC.

Fase 3: Creare il dominio VPC (account di destinazione)

```
aws opensearch create-domain \
  --domain-name vpc-domain-name \
  --engine-version OpenSearch_2.19 \
  --cluster-config InstanceType=m5.large.search,InstanceCount=1 \
  --ebs-options "EBSEnabled=true,VolumeType=gp3,VolumeSize=100" \
  --vpc-options "SubnetIds=subnet-id,SecurityGroupIds=security-group-id" \
  --advanced-security-options
 '{"Enabled":true,"InternalUserDatabaseEnabled":true,"MasterUserOptions":
{"MasterUserName":"admin","MasterUserPassword":"master-password"}}' \
  --node-to-node-encryption-options '{"Enabled":true}' \
  --encryption-at-rest-options '{"Enabled":true}' \
```

```
  --domain-endpoint-options '{"EnforceHTTPS":true,"TLSSecurityPolicy":"Policy-Min-
TLS-1-2-2019-07"}' \
  --access-policies '{"Version":"2012-10-17",       "Statement":
[{"Effect":"Allow","Principal":
{"AWS":"*"},"Action":"es:ESHttp*","Resource":"arn:aws:es:region:target-account-
id:domain/vpc-domain-name/*"}]}' \
  --region region
```

Attendi che lo stato del dominio diventi tale `Active` prima di procedere.

Passaggio 4: autorizzare l'endpoint VPC per il principale OpenSearch del servizio UI (account di destinazione)

> **Important**
>
> Si tratta di un passaggio fondamentale che riguarda esclusivamente i domini VPC. Il servizio OpenSearch UI deve essere esplicitamente autorizzato ad accedere all'endpoint VPC.

```
# Authorize the service principal
aws opensearch authorize-vpc-endpoint-access \
  --domain-name vpc-domain-name \
  --service "application.opensearchservice.amazonaws.com" \
  --region region

# Verify authorization
aws opensearch list-vpc-endpoint-access \
  --domain-name vpc-domain-name \
  --region region
```

Risposta prevista:

```
{
  "AuthorizedPrincipalList": [
    {
      "PrincipalType": "AWS_SERVICE",
      "Principal": "application.opensearchservice.amazonaws.com"
    }
  ]
}
```

Fase 5: Creare l'applicazione OpenSearch UI (account di origine)

```
aws opensearch create-application \
  --region region \
  --name "cross-account-vpc-iam-app" \
  --data-sources '[{
    "dataSourceArn":"arn:aws:es:region:target-account-id:domain/vpc-domain-name",
    "dataSourceDescription":"Cross-account VPC domain",
    "iamRoleForDataSourceArn":"arn:aws:iam::target-account-
id:role/OpenSearchUIAccessRole"
  }]' \
  --app-configs '[{"key":"opensearchDashboards.dashboardAdmin.users","value":"[\"*
\"]"}]'
```

Passaggio 6: verifica e accesso

Recupera i dettagli dell'applicazione per ottenere l'URL dell'endpoint:

```
aws opensearch get-application \
  --region region \
  --id application-id
```

- Passa all'URL dell'endpoint dell'applicazione dalla risposta.
- Accedi con le credenziali IAM.
- L'utente IAM firma le richieste del piano dati con le proprie credenziali.
- La policy di accesso al dominio di destinazione controlla a quali dati l'utente può accedere.

Scenario 4: utente IAM Identity Center che accede a un dominio VPC

Fase 1: Creare il ruolo IAM per più account (account di destinazione)

Crea un ruolo IAM nell'account di destinazione che consenta all'account di origine di assumerlo per la convalida del dominio.

Per creare il ruolo tra account

1. Crea una politica di fiducia che consenta all'account di origine di assumere il ruolo:

```
{
  "Version": "2012-10-17",
```

```
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account-id:root"
    },
    "Action": "sts:AssumeRole"
  }]
}
```

2. Crea il ruolo:

```
aws iam create-role \
  --role-name OpenSearchUIAccessRole \
  --assume-role-policy-document file://trust-policy.json
```

3. Crea una politica di autorizzazioni con la sola `es:DescribeDomain` azione:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": "es:DescribeDomain",
    "Resource": "arn:aws:es:region:target-account-id:domain/*"
  }]
}
```

4. Allega la politica delle autorizzazioni al ruolo:

```
aws iam put-role-policy \
  --role-name OpenSearchUIAccessRole \
  --policy-name ValidationOnly \
  --policy-document file://permissions-policy.json
```

Fase 2: Configurare il VPC (account di destinazione)

Salta questo passaggio se nell'account di destinazione esiste già un VPC.

```
# Create VPC
aws ec2 create-vpc \
  --cidr-block 10.0.0.0/16 \
  --region region
```

```
# Create subnet
aws ec2 create-subnet \
  --vpc-id vpc-id \
  --cidr-block 10.0.1.0/24 \
  --availability-zone regiona \
  --region region

# Create security group
aws ec2 create-security-group \
  --group-name opensearch-vpc-sg \
  --description "Security group for OpenSearch VPC domain" \
  --vpc-id vpc-id \
  --region region

# Allow inbound HTTPS
aws ec2 authorize-security-group-ingress \
  --group-id security-group-id \
  --protocol tcp \
  --port 443 \
  --cidr 10.0.0.0/16 \
  --region region
```

Scopri di più sulla creazione di domini VPC.

Fase 3: Creare il dominio VPC (account di destinazione)

Utilizza lo stesso comandothe section called “Fase 3: Creare il dominio VPC (account di destinazione)”, ma aggiorna la policy di accesso per consentire il ruolo dell'applicazione IAM Identity Center dall'account di origine:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account-id:role/NeoIdCAppRole"
    },
    "Action": "es:ESHttp*",
    "Resource": "arn:aws:es:region:target-account-id:domain/vpc-domain-name/*"
  }]
}
```

Attendi che lo stato del dominio diventi tale `Active` prima di procedere.

Passaggio 4: autorizzare l'endpoint VPC per il principale OpenSearch del servizio UI (account di destinazione)

> **Important**
>
> Si tratta di un passaggio fondamentale che riguarda esclusivamente i domini VPC. Il servizio OpenSearch UI deve essere esplicitamente autorizzato ad accedere all'endpoint VPC.

```
# Authorize the service principal
aws opensearch authorize-vpc-endpoint-access \
  --domain-name vpc-domain-name \
  --service "application.opensearchservice.amazonaws.com" \
  --region region

# Verify authorization
aws opensearch list-vpc-endpoint-access \
  --domain-name vpc-domain-name \
  --region region
```

Risposta prevista:

```
{
  "AuthorizedPrincipalList": [
    {
      "PrincipalType": "AWS_SERVICE",
      "Principal": "application.opensearchservice.amazonaws.com"
    }
  ]
}
```

Fase 5: Creare il ruolo IAM per l'applicazione IAM Identity Center (account di origine)

Crea un ruolo IAM nell'account di origine che l' OpenSearch interfaccia utente utilizza per l'accesso al piano dati degli utenti di IAM Identity Center.

Per creare il ruolo applicativo IAM Identity Center

1. Crea una politica di fiducia:

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "application.opensearchservice.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    },
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "application.opensearchservice.amazonaws.com"
      },
      "Action": "sts:SetContext",
      "Condition": {
        "ForAllValues:ArnEquals": {
          "sts:RequestContextProviders": "arn:aws:iam::source-account-id:oidc-
provider/portal.sso.region.amazonaws.com/apl/application-id"
        }
      }
    }
  ]
}
```

2. Crea una politica di autorizzazioni:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "OpenSearchDomain",
    "Effect": "Allow",
    "Action": ["es:ESHttp*"],
    "Resource": "*"
  }]
}
```

3. Crea il ruolo e allega le politiche:

```
aws iam create-role \
  --role-name NeoIdCAppRole \
  --assume-role-policy-document file://neoidc-trust-policy.json
```

```
aws iam put-role-policy \
  --role-name NeoIdCAppRole \
  --policy-name NeoIdCAppPermissions \
  --policy-document file://neoidc-permissions-policy.json
```

Fase 6: Creare l'applicazione OpenSearch UI con IAM Identity Center (account di origine)

```
aws opensearch create-application \
  --region region \
  --name "cross-account-vpc-idc-app" \
  --iam-identity-center-options '{
    "enabled":true,
    "iamIdentityCenterInstanceArn":"arn:aws:sso:::instance/ssoins-instance-id",
    "iamRoleForIdentityCenterApplicationArn":"arn:aws:iam::source-account-
id:role/NeoIdCAppRole"
  }' \
  --data-sources '[{
    "dataSourceArn":"arn:aws:es:region:target-account-id:domain/vpc-domain-name",
    "dataSourceDescription":"Cross-account VPC domain",
    "iamRoleForDataSourceArn":"arn:aws:iam::target-account-
id:role/OpenSearchUIAccessRole"
  }]' \
  --app-configs '[{"key":"opensearchDashboards.dashboardAdmin.users","value":"[\"*
\"]"}]'
```

Fase 7: Creare e assegnare utenti e gruppi di IAM Identity Center

Crea un utente IAM Identity Center

Eseguire il seguente comando seguente. Sostituisci *placeholder values* con le informazioni appropriate.

```
aws identitystore create-user \
  --identity-store-id d-directory-id \
  --user-name user-email \
  --display-name "display-name" \
  --name Formatted=string,FamilyName=last-name,GivenName=first-name \
  --emails Value=user-email,Type=work,Primary=true
```

Crea un gruppo IAM Identity Center e aggiungi l'utente

Esegui i comandi seguenti:

```
aws identitystore create-group \
  --identity-store-id d-directory-id \
  --display-name "OpenSearchUsers" \
  --description "Users with OpenSearch access"

aws identitystore create-group-membership \
  --identity-store-id d-directory-id \
  --group-id group-id \
  --member-id UserId=user-id
```

Assegna l'utente o il gruppo all'applicazione

Esegui il comando seguente:

```
aws sso-admin create-application-assignment \
  --application-arn "arn:aws:sso:::source-account-id:application/ssoins-instance-id/
apl-application-id" \
  --principal-id user-id-or-group-id \
  --principal-type USER
```

Configura la mappatura dei ruoli di backend sul dominio di destinazione

Mappa il gruppo IAM Identity Center su un ruolo OpenSearch di sicurezza nel dominio di destinazione:

```
curl -XPUT "https://domain-endpoint/_plugins/_security/api/rolesmapping/all_access" \
  -u admin:master-password \
  -H 'Content-Type: application/json' \
  -d '{
    "backend_roles": ["group-id"],
    "hosts": [],
    "users": []
  }'
```

Fase 8: Verifica e accesso

```
aws opensearch get-application \
  --region region \
  --id application-id
```

- Vai all'URL dell'endpoint dell'applicazione.

- Accedi con le credenziali utente di IAM Identity Center.
- Le richieste di dati degli utenti di IAM Identity Center vengono firmate con il ruolo applicativo IAM Identity Center, non con il ruolo tra account.
- Le mappature dei ruoli di backend sul dominio controllano le autorizzazioni di accesso ai dati.

Gestione delle applicazioni

Aggiorna un'applicazione con fonti di dati tra account diversi

Eseguire il seguente comando seguente. Sostituisci *placeholder values* con le informazioni appropriate.

```
aws opensearch update-application \
  --region region \
  --id application-id \
  --data-sources '[{
    "dataSourceArn":"arn:aws:es:region:target-account-id:domain/domain-1",
    "dataSourceDescription":"First cross-account domain",
    "iamRoleForDataSourceArn":"arn:aws:iam::target-account-
id:role/OpenSearchUIAccessRole"
  },{
    "dataSourceArn":"arn:aws:es:region:target-account-id:domain/domain-2",
    "dataSourceDescription":"Second cross-account domain",
    "iamRoleForDataSourceArn":"arn:aws:iam::target-account-
id:role/OpenSearchUIAccessRole"
  }]'
```

> **Important**
>
> L'operazione di aggiornamento sostituisce l'intero array di fonti di dati. Includi tutte le fonti di dati che desideri conservare.

Elenca le applicazioni

Esegui il comando seguente:

```
aws opensearch list-applications \
  --region region
```

Eliminazione di un'applicazione

Esegui il comando seguente:

```
aws opensearch delete-application \
  --region region \
  --id application-id
```

Revoca l'accesso agli endpoint VPC

Esegui il comando seguente:

```
aws opensearch revoke-vpc-endpoint-access \
  --domain-name vpc-domain-name \
  --service "application.opensearchservice.amazonaws.com" \
  --region region
```

Riferimento rapido

Le tabelle seguenti riassumono le principali differenze tra i tipi di dominio e i metodi di autenticazione.

Dominio pubblico rispetto al dominio VPC

Aspetto	Dominio pubblico	Dominio VPC
Autorizzazione degli endpoint VPC	Campo non obbligatorio	Obbligatoria: deve autorizzare `application.opensearchservice.amazonaws.com`
Configurazione della rete	Nessuno	VPC, sottorete, gruppo di sicurezza con HTTPS (443) in ingresso
Politica di accesso IAM	Richiesto	Richiesto
Ruolo tra account	Obbligatorio per più account	Obbligatorio per più account

Utente IAM rispetto all'utente IAM Identity Center

Aspetto	Utente IAM	Utente IAM Identity Center
Credenziali del piano dati	Credenziali IAM proprie dell'utente	Ruolo applicativo IAM Identity Center
Controllo accessi	Policy di accesso al dominio	Policy di accesso al dominio e mappatura dei ruoli di backend
Configurazione aggiuntiva	Nessuno	Ruolo dell'applicazione IAM Identity Center, user/group creazione, assegnazione delle applicazioni, mappatura dei ruoli di backend
OpenSearch Configurazione dell'applicazione UI	Nessuna opzione IAM Identity Center	`--iam-identity-center-options` obbligatorio

Note importanti

- `iamRoleForDataSourceArn`Deve essere nello stesso account del`dataSourceArn`.
- `iamRoleForDataSourceArn`È richiesto solo per le fonti di dati tra account diversi. Omettilo per le fonti di dati dello stesso account.
- Il ruolo tra account diversi richiede solo l'autorizzazione. `es:DescribeDomain` Non viene mai utilizzato per l'accesso al piano dati.
- Per i domini VPC, è necessario configurare sia la policy IAM che l'autorizzazione degli endpoint VPC.
- Versioni del motore supportate: OpenSearch 1.3 e successive.

Risoluzione dei problemi

Problema	Risoluzione
La creazione dell'applicazione non riesce con «Impossibile accedere al dominio»	Verifica che il ruolo interaccount disponga dell'`es:DescribeDomain` autorizzazione e che la politica di attendibilità consenta l'accesso all'account di origine.
L'associazione del dominio VPC non riesce	Assicurati che l'endpoint VPC sia autorizzato per. `application.opensearchservice.amazonaws.com`
Accesso al piano dati negato per l'utente IAM	Verifica che la policy di accesso al dominio di destinazione consenta l'utente o il responsabile del ruolo IAM.
Accesso al piano dati negato per l'utente IAM Identity Center	Verifica che la mappatura dei ruoli di backend includa l'ID del gruppo IAM Identity Center e che la policy del dominio consenta il ruolo dell'applicazione IAM Identity Center.
Errore di mancata corrispondenza dell'account	Assicurati che `iamRoleForDataSourceArn` si trovi nello stesso account del dominio in`dataSourceArn` .

Funzionalità di ricerca tra cluster

Utilizzando la ricerca tra cluster in Amazon OpenSearch Serverless, puoi eseguire query e aggregazioni su più domini connessi.

La ricerca tra cluster in Amazon OpenSearch Serverless utilizza i concetti di dominio di origine e dominio di destinazione. Una richiesta di ricerca tra cluster proviene da un dominio di origine. Il dominio di destinazione può appartenere a un dominio diverso Account AWS o Regione AWS (o entrambi) al dominio di origine da cui eseguire la query. Utilizzando la ricerca tra cluster, puoi configurare un dominio di origine da associare all' OpenSearch interfaccia utente nello stesso account e quindi creare connessioni ai domini di destinazione. Di conseguenza, puoi utilizzare l' OpenSearch interfaccia utente con i dati dei domini di destinazione anche se si trovano in un account o in una regione diversi.

Paghi AWS i costi standard di trasferimento dei dati per i dati trasferiti da e verso il OpenSearch Servizio Amazon. Non ti vengono addebitati costi per i dati trasferiti tra i nodi all'interno OpenSearch del tuo dominio Service. Per ulteriori informazioni sui costi di «ingresso» e «uscita» dei dati, consulta la pagina dei prezzi On-Demand di Amazon EC2 per il trasferimento dei dati.

Puoi utilizzare la ricerca tra cluster come meccanismo per associare l' OpenSearch interfaccia utente a cluster in un account o in una regione diversa. Per impostazione predefinita, le richieste tra domini sono crittografate in transito come parte della crittografia. node-to-node

> **Note**
>
> Lo OpenSearch strumento open source documenta anche la ricerca tra cluster. Tieni presente che la configurazione dello strumento open source differisce in modo significativo per i cluster open source rispetto ai domini Amazon OpenSearch Serverless gestiti.
> In particolare, in Amazon OpenSearch Serverless, si configurano connessioni tra cluster utilizzando le richieste Console di gestione AWS anziché utilizzando. `cURL` Il servizio gestito utilizza AWS Identity and Access Management (IAM) per l'autenticazione tra cluster oltre al controllo granulare degli accessi.
> Pertanto, consigliamo di utilizzare il contenuto di questo argomento per configurare la ricerca tra cluster per i domini anziché la documentazione open source. OpenSearch

Differenze funzionali quando si utilizza la ricerca tra cluster

Rispetto ai domini normali, i domini di destinazione creati utilizzando la ricerca tra cluster presentano le seguenti differenze e requisiti funzionali:

- Non è possibile scrivere o eseguire `PUT` comandi nel cluster remoto. L'accesso al cluster remoto è di sola lettura.
- Sia il dominio di origine che quello di destinazione devono essere OpenSearch domini. Non puoi connettere un dominio Elasticsearch o cluster OpenSearch /Elasticsearch autogestiti per l'interfaccia utente. OpenSearch
- Un dominio può avere un massimo di 20 connessioni ad altri domini. Ciò include sia le connessioni in uscita che quelle in entrata.
- Il dominio di origine deve trovarsi nella stessa versione o in una versione successiva del dominio OpenSearch di destinazione. Se desideri configurare connessioni bidirezionali tra due domini, i due domini devono avere la stessa versione. Ti consigliamo di aggiornare entrambi i domini alla

versione più recente prima di effettuare la connessione. Se devi aggiornare i domini dopo aver configurato la connessione bidirezionale, devi prima eliminare la connessione e poi ricrearla.

- Non è possibile utilizzare dizionari personalizzati o SQL con i cluster remoti.
- Non puoi utilizzarlo per CloudFormation connettere domini.
- Non è possibile utilizzare la ricerca tra cluster su istanze M3 o istanze espandibili (T2 e T3).
- La ricerca tra cluster non funziona per le raccolte Amazon OpenSearch Serverless.

Prerequisiti di ricerca tra cluster per l'interfaccia utente OpenSearch

Prima di configurare la ricerca tra cluster con due OpenSearch domini, assicurati che i domini soddisfino i seguenti requisiti:

- Il controllo granulare degli accessi è abilitato per entrambi i domini
- Node-to-node la crittografia è abilitata per entrambi i domini

Argomenti

- Configurazione delle autorizzazioni di accesso per l'accesso ai dati tra aree geografiche e tra account con la ricerca tra cluster
- Creazione di una connessione tra domini
- Test della configurazione di sicurezza per l'accesso ai dati tra regioni e account diversi con la ricerca tra cluster
- Eliminazione di una connessione

Configurazione delle autorizzazioni di accesso per l'accesso ai dati tra aree geografiche e tra account con la ricerca tra cluster

Quando invii una richiesta di ricerca tra cluster al dominio di origine, il dominio valuta tale richiesta in base alla propria politica di accesso al dominio. La ricerca tra cluster richiede un controllo granulare degli accessi. Di seguito è riportato un esempio con una politica di accesso aperto sul dominio di origine.

JSON

```
{
```

```
"Version":"2012-10-17",
"Statement": [
    {
        "Effect": "Allow",
        "Principal": {
          "AWS": [
            "*"
        ]
      },
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": "arn:aws:es:us-east-1:111222333444:domain/src-domain/*"
    }
  ]
}
```

> **Note**
>
> Se includi indici remoti nel percorso, devi codificare in formato URL l'URI nell'ARN del dominio.
> Ad esempio, utilizzate il seguente formato ARN: `:arn:aws:es:us-east-1:111222333444:domain/my-domain/local_index,dst%3Aremote_index`
> Non utilizzare il seguente formato ARN: `arn:aws:es:us-east-1:111222333444:domain/my-domain/local_index,dst:remote_index.`

Se si sceglie di utilizzare una politica di accesso restrittiva oltre al controllo granulare degli accessi, la politica deve almeno consentire l'accesso a. `es:ESHttpGet` Di seguito è riportato un esempio:

JSON

```
{
"Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
```

```
      "Principal": {
        "AWS": [
          "arn:aws:iam::111222333444:user/john-doe"
        ]
      },
      "Action": "es:ESHttpGet",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/my-domain/*"
    }
  ]
}
```

Il controllo granulare degli accessi sul dominio di origine valuta la richiesta per determinare se è firmata con credenziali di base IAM o HTTP valide. In caso affermativo, il controllo granulare degli accessi valuta successivamente se l'utente è autorizzato a eseguire la ricerca e accedere ai dati.

Di seguito sono riportati i requisiti di autorizzazione per le ricerche:

- Se la richiesta cerca solo dati nel dominio di destinazione (ad esempio`dest-alias:dest-index/_search`), le autorizzazioni sono richieste solo nel dominio di destinazione).
- Se la richiesta cerca dati su entrambi i domini (ad esempio`source-index,dest-alias:dest-index/_search`), le autorizzazioni sono necessarie su entrambi i domini).
- Per utilizzare un controllo granulare degli accessi, `indices:admin/shards/search_shards` è necessaria l'autorizzazione in aggiunta alle autorizzazioni standard di lettura o ricerca per gli indici pertinenti.

Il dominio di origine invia la richiesta al dominio di destinazione. Il dominio di destinazione valuta la richiesta in base alla policy di accesso al dominio. Per supportare tutte le funzionalità OpenSearch dell'interfaccia utente, come l'indicizzazione dei documenti e l'esecuzione di ricerche standard, è necessario impostare le autorizzazioni complete. Di seguito è riportato un esempio della nostra politica consigliata sul dominio di destinazione:

JSON

```
{
"Version":"2012-10-17",
 "Statement": [
    {
       "Effect": "Allow",
```

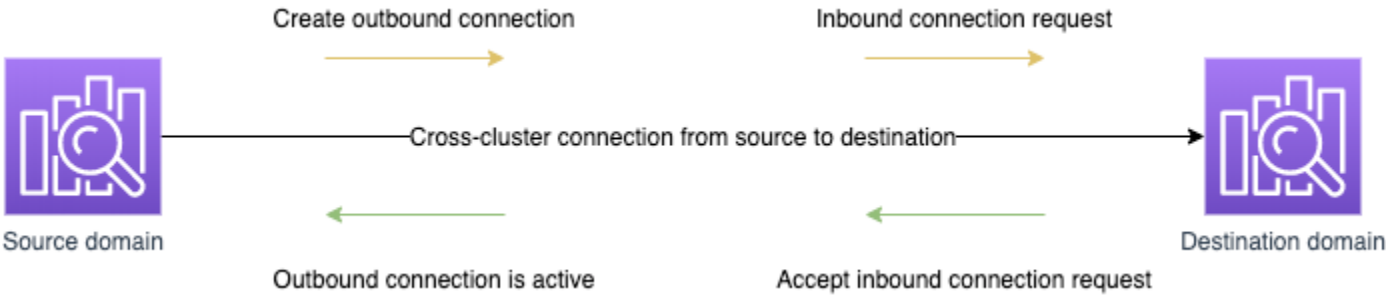
```
       "Principal": {
         "AWS": [
           "*"
         ]
       },
       "Action": [
         "es:ESHttp*"
       ],
       "Resource": "arn:aws:es:us-east-2:111222333444:domain/my-destination-
domain/*"
    },
    {
    "Effect": "Allow",
        "Principal": {
    "AWS": "*"
      },
      "Action": "es:ESCrossClusterGet",
      "Resource": "arn:aws:es:us-east-2:111222333444:domain/"
    }
  ]
}
```

Se desideri eseguire solo ricerche di base, il requisito minimo richiesto dalla policy è l'`es:ESCrossClusterGet`autorizzazione da applicare al dominio di destinazione senza il supporto dei caratteri jolly. Ad esempio, nella politica precedente, è necessario specificare il nome di dominio con e senza. */my-destination-domain /my-destination-domain/**

In questo caso, il dominio di destinazione esegue la ricerca e restituisce i risultati al dominio di origine. Il dominio di origine combina i propri risultati (se presenti) con i risultati del dominio di destinazione e li restituisce all'utente.

Creazione di una connessione tra domini

Una connessione di ricerca tra cluster è unidirezionale dal dominio di origine al dominio di destinazione. Ciò significa che i domini di destinazione (in un account o in una regione diversi) non possono interrogare il dominio di origine, che è locale all'interfaccia utente. OpenSearch Il dominio di origine crea una connessione in uscita al dominio di destinazione. Il dominio di destinazione riceve una richiesta di connessione in entrata dal dominio di origine.



Per creare una connessione tra domini

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nella barra di navigazione a sinistra, scegli Domini.
3. Scegli il nome di un dominio da utilizzare come dominio di origine, quindi scegli la scheda Connessioni.
4. Nell'area Connessioni in uscita, scegli Richiesta.
5. Per Alias di connessione, specifica un nome per la connessione. L'alias di connessione viene utilizzato nell' OpenSearch interfaccia utente per selezionare i domini di destinazione.
6. Per la modalità di connessione, scegli Direct per le ricerche o la replica tra cluster.
7. Per specificare che la connessione deve ignorare i cluster non disponibili durante una ricerca, seleziona la casella Ignora i cluster non disponibili. La scelta di questa opzione garantisce che le query tra cluster restituiscano risultati parziali indipendentemente dagli errori su uno o più cluster remoti.
8. Per Cluster di destinazione, scegli tra Connetti a un cluster in questo Account AWS e Connetti a un cluster in un altro Account AWS.
9. Per l'ARN del dominio remoto, inserisci l'Amazon Resource Name (ARN) per il cluster. L'ARN del dominio può essere posizionato nell'area Informazioni generali della pagina di dettaglio del dominio.

   Il dominio deve soddisfare i seguenti requisiti:

   - L'ARN deve essere nel formato. `arn:`*`partition`*`:es:`*`regionaccount-id`*`:`*`type`*`/`*`domain-id`* Esempio:

     ```
     arn:aws:es:us-east-2:111222333444:domain/my-domain
     ```
   - Il dominio deve essere configurato per utilizzare la OpenSearch versione 1.0 (o successiva) o la versione 6.7 di Elasticsearch (o successiva).

- Il controllo granulare degli accessi deve essere abilitato sul dominio.
- Il dominio deve essere in esecuzione. OpenSearch

10. Scegli Richiedi.

La ricerca tra cluster convalida innanzitutto la richiesta di connessione per assicurare che i prerequisiti siano soddisfatti. Se i domini sono incompatibili, la richiesta di connessione entra nello `Validation failed` stato.

Se la richiesta di connessione viene convalidata correttamente, viene inviata al dominio di destinazione, dove deve essere approvata. Fino a quando non viene concessa questa approvazione, la connessione rimane in uno `Pending acceptance` stato. Quando la richiesta di connessione viene accettata nel dominio di destinazione, lo stato cambia in `Active` e il dominio di destinazione diventa disponibile per le query.

La pagina del dominio mostra i dettagli generali dello stato del dominio e dello stato dell'istanza del dominio di destinazione. Solo i proprietari dei domini dispongono della flessibilità necessaria per creare, visualizzare, rimuovere e monitorare le connessioni da o verso i propri domini.

Una volta stabilita la connessione, tutto il traffico che scorre tra i nodi dei domini connessi viene crittografato. Quando connetti un dominio VPC a un dominio non VPC e il dominio non VPC è un endpoint pubblico in grado di ricevere traffico da Internet, il traffico intercluster tra i domini è ancora crittografato e sicuro.

Test della configurazione di sicurezza per l'accesso ai dati tra regioni e account diversi con la ricerca tra cluster

Dopo aver impostato le autorizzazioni di accesso per l'accesso ai dati tra aree geografiche e tra account con la ricerca tra cluster, ti consigliamo di testare la configurazione utilizzando una piattaforma di terze parti per lo sviluppo collaborativo di PostmanAPI.

Per impostare la configurazione di sicurezza utilizzando Postman

1. Nel dominio di destinazione, indicizza un documento. Di seguito è riportato un esempio di richiesta:

```
POST https://dst-domain.us-east-1.es.amazonaws.com/books/_doc/1
{
"Dracula": "Bram Stoker"
```

```
}
```

2. Per eseguire una query su questo indice dal dominio di origine, includere l'alias di connessione del dominio di destinazione all'interno della query. È possibile trovare l'alias di connessione nella scheda Connessioni nel pannello di controllo del dominio. Di seguito sono riportati un esempio di richiesta e una risposta troncata:

```
GET https://src-domain.us-east-1.es.amazonaws.com/connection_alias:books/_search
{
...
  "hits": [
{
"_index": "source-destination:books",
  "_type": "_doc",
  "_id": "1",
  "_score": 1,
  "_source": {
"Dracula": "Bram Stoker"
  }
}
  ]
}
```

3. (Facoltativo) È possibile creare una configurazione che includa più domini in un'unica ricerca. Ad esempio, supponiamo di aver impostato quanto segue:

   Una connessione tra `domain-a` a`domain-b`, con un alias di connessione denominato `cluster_b`

   Una connessione tra `domain-a` a`domain-c`, con un alias di connessione denominato `cluster_c`

   In questo caso, le ricerche includono il contenuto `domain-a`o`domain-b`, e. `domain-c` Di seguito sono riportati alcuni esempi di richiesta e risposta:

   Richiesta

```
GET https://src-domain.us-east-1.es.amazonaws.com/
local_index,cluster_b:b_index,cluster_c:c_index/_search
{
  "query": {
"match": {
```

```
  "user": "domino"
}
  }
}
```

Risposta:

```
{
"took": 150,
  "timed_out": false,
  "_shards": {
"total": 3,
"successful": 3,
"failed": 0,
"skipped": 0
  },
  "_clusters": {
"total": 3,
"successful": 3,
"skipped": 0
  },
  "hits": {
"total": 3,
"max_score": 1,
"hits": [
  {
"_index": "local_index",
    "_type": "_doc",
    "_id": "0",
    "_score": 1,
    "_source": {
"user": "domino",
      "message": "This is message 1",
      "likes": 0
    }
  },
  {
"_index": "cluster_b:b_index",
    "_type": "_doc",
    "_id": "0",
    "_score": 2,
    "_source": {
"user": "domino",
```

```
      "message": "This is message 2",
      "likes": 0
    }
  },
  {
"_index": "cluster_c:c_index",
    "_type": "_doc",
    "_id": "0",
    "_score": 3,
    "_source": {
"user": "domino",
      "message": "This is message 3",
      "likes": 0
    }
  }
]
  }
}
```

Se hai scelto di non ignorare i cluster non disponibili nella configurazione della connessione, tutti i cluster di destinazione che cerchi devono essere disponibili affinché la richiesta di ricerca venga eseguita correttamente. In caso contrario, l'intera richiesta avrà esito negativo: anche se uno dei domini non è disponibile non verrà restituito alcun risultato della ricerca.

Eliminazione di una connessione

L'eliminazione di una connessione interrompe tutte le operazioni di ricerca tra cluster nel dominio di destinazione.

È possibile eseguire la procedura seguente sul dominio di origine o di destinazione per rimuovere la connessione. Dopo aver rimosso la connessione, questa rimane visibile con uno stato `Deleted` di 15 giorni.

Non è possibile eliminare un dominio con connessioni tra cluster attive. Per eliminare un dominio, rimuovere innanzitutto tutte le connessioni in ingresso e in uscita da tale dominio. Questo per essere certi di considerare gli utenti del dominio tra cluster prima di eliminare il dominio.

Come eliminare una connessione

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.

2. Nella barra di navigazione a sinistra, scegli Domini.
3. Scegli il nome di un dominio da eliminare, quindi scegli la scheda Connessioni.
4. Seleziona il nome di una connessione da eliminare.
5. Scegli Elimina, quindi conferma l'eliminazione.

Gestione dell'accesso all' OpenSearch interfaccia utente da un endpoint VPC

Puoi creare una connessione privata tra il tuo VPC e l' OpenSearch interfaccia utente utilizzando. AWS PrivateLink Utilizzando questa connessione, puoi accedere alle applicazioni OpenSearch dell'interfaccia utente come se si trovassero nello stesso VPC. In questo modo, non è necessario configurare un gateway Internet, un dispositivo NAT, una connessione VPN o AWS Direct Connect stabilire la connessione. Le istanze nel tuo VPC non necessitano di indirizzi IP pubblici per accedere all' OpenSearch interfaccia utente.

Per stabilire questa connessione privata, devi prima creare un'interfaccia endpoint alimentata da. AWS PrivateLink Un'interfaccia di rete endpoint viene creata automaticamente in ogni sottorete specificata per l'endpoint di interfaccia. Si tratta di interfacce di rete gestite dai richiedenti che fungono da punto di ingresso per il traffico destinato alle applicazioni dell'interfaccia utente. OpenSearch

Creazione di una connessione privata tra un VPC e un'interfaccia utente OpenSearch

Puoi creare una connessione privata per accedere all' OpenSearch interfaccia utente da un VPC utilizzando o. Console di gestione AWS AWS CLI

Creazione di una connessione privata tra un VPC e un' OpenSearch interfaccia utente (console)

Per creare una connessione privata tra un VPC e l' OpenSearch interfaccia utente utilizzando la console

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Nella barra di navigazione a sinistra, in Serverless, scegli Endpoint VPC.

3. Scegli Create VPC endpoint (Crea endpoint VPC).
4. In Nome, inserisci un nome per l'endpoint.
5. Per VPC, seleziona il VPC da cui accederai alle OpenSearch applicazioni dell'interfaccia utente.
6. Per le sottoreti, seleziona una sottorete da cui accedere alle applicazioni dell'interfaccia utente. OpenSearch

> **Note**
>
> L'indirizzo IP e il tipo DNS di un endpoint si basano sul tipo di sottorete:
>
> - Dual-stack: se tutte le sottoreti hanno entrambi gli intervalli di indirizzi. IPv4 IPv6
> - IPv6: Se tutte le sottoreti sono solo sottoreti. IPv6
> - IPv4: Se tutte le sottoreti hanno intervalli di indirizzi. IPv4

7. Per i gruppi di sicurezza, selezionare uno o più gruppi di sicurezza da associare alle interfacce di rete degli endpoint.

> **Note**
>
> In questo passaggio, stai limitando le porte, i protocolli e le fonti per il traffico in entrata che stai autorizzando verso il tuo endpoint. Assicurati che le regole del gruppo di sicurezza consentano alle risorse che utilizzeranno l'endpoint VPC di comunicare con le applicazioni OpenSearch dell'interfaccia utente di comunicare anche con l'interfaccia di rete dell'endpoint.

8. 8. Seleziona Crea endpoint.

Creazione di una connessione privata tra un VPC e l' OpenSearch interfaccia utente ()AWS CLI

Per creare una connessione privata tra un VPC e l' OpenSearch interfaccia utente utilizzando AWS CLI

Eseguire il seguente comando seguente. Sostituisci *`placeholder values`* con le informazioni appropriate.

```
aws opensearchserverless create-vpc-endpoint \
```

```
    --region region \
    --endpoint endpoint \
    --name vpc_endpoint_name \
    --vpc-id vpc_id \
    --subnet-ids subnet_ids
```

Aggiornamento della policy degli endpoint VPC per consentire l'accesso all'applicazione UI OpenSearch

Dopo aver creato la connessione privata, aggiorna la policy degli endpoint VPC per consentire l'accesso all'applicazione OpenSearch UI nella policy degli endpoint VPC specificando l'ID dell'applicazione.

Per informazioni sull'aggiornamento di una policy per gli endpoint VPC, consulta Aggiornare una policy per gli endpoint VPC nella Guida.AWS PrivateLink

Assicurati che la policy degli endpoint VPC includa la seguente dichiarazione. Sostituisci *placeholder value* con le informazioni appropriate.

```
{
    "Statement": [{
        "Action": ["opensearch:*"],
        "Effect": "Allow",
        "Principal": "*",
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "opensearch:ApplicationId": ["opensearch-ui-application-id"]
            }
        }
    }]
}
```

Revoca dell'accesso all' OpenSearch interfaccia utente in una policy di endpoint VPC

OpenSearch L'interfaccia utente richiede un'autorizzazione esplicita nella policy degli endpoint VPC per consentire agli utenti di accedere all'applicazione dal VPC. Se non desideri più che gli utenti accedano all' OpenSearch interfaccia utente dal VPC, puoi rimuovere l'autorizzazione nella policy

degli endpoint. Successivamente, gli utenti visualizzano un messaggio di `403 forbidden` errore quando tentano di accedere all'interfaccia utente. OpenSearch

Per informazioni sull'aggiornamento di una policy per gli endpoint VPC, consulta Aggiornare una policy per gli endpoint VPC nella Guida.AWS PrivateLink

Di seguito è riportato un esempio di policy degli endpoint VPC che nega l'accesso alle applicazioni dell'interfaccia utente dal VPC:

```
{
    "Statement": [{
        "Action": ["opensearch:*"],
        "Effect": "Allow",
        "Principal": "*",
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "opensearch:ApplicationId": [""]
            }
        }
    }]
}
```

Migrazione degli oggetti salvati dalle dashboard all'interfaccia utente OpenSearch OpenSearch

Se disponi di dashboard, visualizzazioni, modelli di indice e altri oggetti salvati nelle OpenSearch dashboard, puoi migrarli e riutilizzarli nell'interfaccia utente. OpenSearch

Vantaggi della migrazione all'interfaccia utente: OpenSearch

- Alta disponibilità: l' OpenSearch interfaccia utente è ospitata nel dominio Cloud AWS e rimane disponibile durante gli aggiornamenti e la manutenzione del dominio, mentre OpenSearch le dashboard sono ospitate all'interno del dominio e non saranno temporaneamente disponibili.
- Fonte di dati multipla: l' OpenSearchinterfaccia utente può fornire un unico pannello di controllo consolidato su più fonti di dati, inclusi OpenSearch domini, raccolte serverless e connessioni dati con Amazon S3 e Amazon CloudWatch; mentre ogni OpenSearch dashboard può funzionare solo con un dominio o una raccolta.

- Nell'interfaccia utente sono disponibili funzionalità aggiuntive come AI Assistant e Workspaces. OpenSearch Per saperne di più:OpenSearch Interfaccia utente.

Argomenti

- Panoramica sulla migrazione
- Prerequisiti
- Fase 1: Esportazione degli oggetti salvati dai pannelli di controllo OpenSearch
- Passaggio 2: Importa gli oggetti salvati nell'interfaccia utente OpenSearch

Panoramica sulla migrazione

Il processo di migrazione prevede i seguenti passaggi di alto livello:

1. Esporta oggetti salvati dalle OpenSearch dashboard: utilizza l'interfaccia utente di gestione degli oggetti salvati di OpenSearch Dashboards o l'API di esportazione per scaricare dashboard, visualizzazioni, modelli di indice e altri oggetti come file JSON (NDJSON) delimitato da nuove righe.
2. Crea un'applicazione OpenSearch UI e un'area di lavoro: se non l'hai già fatto, crea un'applicazione UI e un'area di lavoro per ricevere gli oggetti importati. OpenSearch
3. Registra l'origine dati nell' OpenSearchinterfaccia utente: associa il tuo OpenSearch dominio all'applicazione OpenSearch UI e registralo come fonte di dati all'interno dell'area di lavoro. I modelli di indice negli oggetti importati fanno riferimento a questa fonte di dati.
4. Importa oggetti salvati nell' OpenSearch interfaccia utente: utilizza l' OpenSearch interfaccia utente di gestione degli oggetti salvati dell'interfaccia utente o l'API di importazione per caricare il file NDJSON nell'area di lavoro di destinazione.
5. Convalida gli oggetti importati: apri i dashboard e le visualizzazioni nell' OpenSearch interfaccia utente per verificare che vengano visualizzati correttamente e che i dati provengano dal dominio o dalla raccolta associati.

Prerequisiti

Prima della migrazione, verifica i seguenti prerequisiti:

- Disponi delle autorizzazioni IAM necessarie per chiamare Amazon OpenSearch Service e OpenSearch APIs. Per ulteriori informazioni, consulta the section called “Autorizzazioni richieste per la creazione di applicazioni Amazon OpenSearch Service”.
- Puoi accedere al dominio o alla raccolta e alle OpenSearch dashboard che desideri migrare.
- Hai creato un'applicazione OpenSearch UI. Per informazioni sulla creazione di un'applicazione e di un'area di lavoro, consultathe section called “Nozioni di base”.
- Hai associato lo stesso dominio o la stessa raccolta all'applicazione OpenSearch UI. Per informazioni sull'associazione delle fonti di dati, consultathe section called “Gestione delle associazioni delle fonti di dati e delle autorizzazioni di accesso VPC”.

> **Note**
>
> OpenSearch L'interfaccia utente supporta solo OpenSearch la versione 1.3 e successive. Verifica che OpenSearch sul tuo dominio sia in esecuzione la versione 1.3 o successiva prima di tentare di migrare gli oggetti salvati.

Fase 1: Esportazione degli oggetti salvati dai pannelli di controllo OpenSearch

Esporta gli oggetti salvati dalle OpenSearch dashboard utilizzando l'interfaccia utente di gestione o l'API di esportazione. L'esportazione produce un file JSON (NDJSON) delimitato da nuove righe che contiene tutti i tipi di oggetti salvati selezionati e le relative dipendenze.

Argomenti

- OpenSearch Esportazione manuale su dashboard
- Esportazione tramite API

OpenSearch Esportazione manuale su dashboard

Per esportare oggetti salvati utilizzando l'interfaccia utente di OpenSearch gestione delle dashboard

1. Apri l'istanza OpenSearch Dashboards.
2. Nel pannello di navigazione a sinistra, scegli Gestione.
3. In Gestione dashboard, scegli Oggetti salvati.

4. Seleziona gli oggetti salvati che desideri esportare. Per esportare tutti gli oggetti di un tipo specifico, filtra per tipo utilizzando la barra di ricerca. Per esportare tutti gli oggetti, selezionate la casella di controllo nell'intestazione della tabella.
5. Scegli Export (Esporta).
6. Nella finestra di dialogo Esporta oggetti salvati, assicuratevi che sia selezionata l'opzione Includi oggetti correlati. Questa opzione include tutti gli oggetti da cui dipendono gli oggetti salvati selezionati, come i modelli di indice a cui fanno riferimento le visualizzazioni. Deselezionate questa opzione solo se intendete gestire le dipendenze manualmente.
7. Scegliete Esporta per scaricare il `.ndjson` file sul computer locale.

> **Tip**
>
> Quando scegli Includi oggetti correlati, il file NDJSON esportato contiene tutti gli oggetti salvati necessari per il rendering delle dashboard e delle visualizzazioni selezionate, inclusi i modelli di indice, le visualizzazioni e gli oggetti di ricerca dipendenti. Ciò semplifica la fase di importazione ed evita errori di riferimento mancanti.

Esportazione tramite API

Puoi utilizzare l'API di esportazione degli oggetti salvati di OpenSearch Dashboards per esportare a livello di codice gli oggetti salvati. Ciò è utile per automatizzare le migrazioni o integrare la fase di esportazione in una pipeline. CI/CD

> **Note**
>
> Se il tuo OpenSearch dominio ha abilitato il controllo granulare degli accessi, devi fornire le credenziali di autenticazione con la richiesta di esportazione. Utilizza l'autenticazione di base HTTP aggiungendo il `-u` contrassegno con nome utente e password. Per ulteriori informazioni sulle opzioni di autenticazione, consulta the section called “Controllo granulare degli accessi”.

L'esempio seguente esporta tutti i dashboard con i relativi oggetti. Sostituisci *`placeholder values`* con le informazioni appropriate.

```
curl -X POST \
```

```
    "https://dashboards-endpoint/_dashboards/api/saved_objects/_export" \
    -u 'master-username:master-password' \
    -H "Content-Type: application/json" \
    -H "osd-xsrf: true" \
    -d '{
        "type": ["dashboard", "visualization", "index-pattern", "search"],
        "includeReferencesDeep": true,
        "excludeExportDetails": false
    }' \
    -o saved-objects-export.ndjson
```

Se nel tuo dominio non è abilitato il controllo granulare degli accessi, puoi omettere il flag. `-u`

Per esportare oggetti salvati specifici in base all'ID, utilizza il `objects` parametro anziché: `type`

```
curl -X POST \
    "https://dashboards-endpoint/_dashboards/api/saved_objects/_export" \
    -u 'master-username:master-password' \
    -H "Content-Type: application/json" \
    -H "osd-xsrf: true" \
    -d '{
        "objects": [
            {"type": "dashboard", "id": "dashboard-id"},
            {"type": "visualization", "id": "visualization-id"}
        ],
        "includeReferencesDeep": true
    }' \
    -o saved-objects-export.ndjson
```

Note

Per trovare oggetti salvati IDs, puoi utilizzare l'API Saved Objects per elencare tutti gli oggetti di un tipo specifico. L'esempio seguente elenca tutti i dashboard:

```
curl -X GET \
    "https://dashboards-endpoint/_dashboards/api/saved_objects/_find?
type=dashboard" \
    -u 'master-username:master-password'
```

La risposta include l'ID per ogni oggetto salvato. Puoi anche trovare l'ID nell'URL del browser quando visualizzi l'oggetto nelle OpenSearch dashboard.

Passaggio 2: Importa gli oggetti salvati nell'interfaccia utente OpenSearch

Dopo aver esportato gli oggetti salvati, puoi importare il file NDJSON nell' OpenSearch interfaccia utente manualmente o tramite API.

Argomenti

- Importazione manuale nell'interfaccia utente OpenSearch
- Importazione tramite API

Importazione manuale nell'interfaccia utente OpenSearch

Per importare oggetti salvati utilizzando l' OpenSearch interfaccia utente di gestione dell'interfaccia utente

1. Apri l'applicazione OpenSearch UI e accedi all'area di lavoro di destinazione.
2. Nell'area di lavoro, scegli Risorse dalla barra di navigazione in alto o vai alla pagina delle risorse dell'area di lavoro.
3. Scegliete Importa per aprire la finestra di dialogo Importa risorse.
4. Scegliete Seleziona file e selezionate il `.ndjson` file esportato dai OpenSearch pannelli di controllo.
5. Per la gestione dei conflitti, scegli una delle seguenti opzioni:
   - Crea nuove risorse con risorse uniche IDs (impostazione predefinita): genera nuove risorse IDs per tutti gli oggetti importati, evitando conflitti con le risorse esistenti.
   - Verifica la presenza di risorse esistenti: verifica la presenza di conflitti con gli oggetti esistenti. Se selezionata, scegliete una delle seguenti opzioni secondarie:
     - Sovrascrivi automaticamente i conflitti: le risorse esistenti con lo stesso ID vengono sostituite automaticamente.
     - Richiedi un'azione in caso di conflitto: ti viene richiesto di risolvere ogni conflitto singolarmente.
6. Scegli Importa.
7. Consulta il riepilogo delle importazioni.

Importazione tramite API

Per importare oggetti salvati utilizzando l'API con autenticazione AWS Signature Version 4, devi prima ottenere l'ID della fonte di dati, quindi utilizzarlo nella richiesta di importazione. Sostituisci *`placeholder values`* con le informazioni appropriate.

Passaggio 1: ottieni l'ID della fonte di dati per il tuo spazio di lavoro:

```
curl -X GET \
    "https://opensearch-ui-endpoint/w/workspace-id/api/saved_objects/_find?type=data-
source" \
    --aws-sigv4 "aws:amz:region:opensearch" \
    --user "$AWS_ACCESS_KEY_ID:$AWS_SECRET_ACCESS_KEY" \
    -H "x-amz-security-token: $AWS_SESSION_TOKEN" \
    -H "osd-xsrf: true"
```

> **Note**
>
> La risposta include l'ID della fonte di dati. Puoi anche trovare l'ID dell'origine dati nell'URL del browser quando visualizzi l'origine dati nell' OpenSearch interfaccia utente.

Passaggio 2: Importa gli oggetti salvati utilizzando l'ID dell'origine dati dal passaggio 1:

```
curl -X POST \
    "https://opensearch-ui-endpoint/w/workspace-id/api/saved_objects/_import?
overwrite=true&dataSourceId=data-source-id" \
    --aws-sigv4 "aws:amz:region:opensearch" \
    --user "$AWS_ACCESS_KEY_ID:$AWS_SECRET_ACCESS_KEY" \
    -H "x-amz-security-token: $AWS_SESSION_TOKEN" \
    -H "osd-xsrf: true" \
    -F "file=@saved-objects-export.ndjson"
```

> **Note**
>
> Questi esempi utilizzano l'`--aws-sigv4`opzione integrata di curl (disponibile in curl 7.75 o versioni successive) per firmare le richieste. Imposta AWS le tue credenziali come variabili di ambiente prima di eseguire i comandi:`AWS_ACCESS_KEY_ID`,`AWS_SECRET_ACCESS_KEY`, e `AWS_SESSION_TOKEN` (se usi credenziali temporanee).

OpenSearch Endpoint e quote dell'interfaccia utente

Per connettersi a livello di codice a un AWS servizio, si utilizza un endpoint. Le quote di servizio, a cui si fa riferimento anche come limiti, rappresentano il numero massimo di risorse di servizio o operazioni per l'account AWS .

Amazon OpenSearch UI è l'interfaccia OpenSearch utente di nuova generazione per le OpenSearch dashboard. Fornisce endpoint per accedere alle dashboard OpenSearch . Utilizza questo argomento per trovare gli endpoint e le quote di servizio per l'interfaccia utente. OpenSearch

Per ulteriori informazioni su altri OpenSearch servizi, consulta Service Endpoints and Quotas.

OpenSearch Endpoint dell'interfaccia utente

OpenSearch L'interfaccia utente è disponibile nelle seguenti regioni:

Nome Regione	Regione	Endpoint	Protocollo
Asia Pacifico (Mumbai)	ap-south-1	opensearch.ap-south-1.amazonaws.com es.ap-south-1.amazonaws.com	HTTPS HTTPS
Europa (Parigi)	eu-west-3	opensearch.eu-west-3.amazonaws.com es.eu-west-3.amazonaws.com	HTTPS HTTPS
Stati Uniti orientali (Ohio)	us-east-2	opensearch.us-east-2.amazonaws.com es.us-east-2.amazonaws.com	HTTPS HTTPS
Europa (Irlanda)	eu-west-1	opensearch.eu-west-1.amazonaws.com es.eu-west-1.amazonaws.com	HTTPS HTTPS
Europa (Francoforte)	eu-central-1	opensearch.eu-central-1.amazonaws.com	HTTPS HTTPS

Nome Regione	Regione	Endpoint	Protocollo
		es.eu-central-1.amazonaws.com	
Sud America (San Paolo)	sa-east-1	opensearch.sa-east-1.amazonaws.com es.sa-east-1.amazonaws.com	HTTPS HTTPS
Stati Uniti orientali (Virginia settentrionale)	us-east-1	opensearch.us-east-1.amazonaws.com es.us-east-1.amazonaws.com	HTTPS HTTPS
Europe (London)	eu-west-2	opensearch.eu-west-2.amazonaws.com es.eu-west-2.amazonaws.com	HTTPS HTTPS
Asia Pacifico (Tokyo)	ap-northeast-1	opensearch.ap-northeast-1.amazonaws.com es.ap-northeast-1.amazonaws.com	HTTPS HTTPS
Stati Uniti occidentali (Oregon)	us-west-2	opensearch.us-west-2.amazonaws.com es.us-west-2.amazonaws.com	HTTPS HTTPS
Asia Pacifico (Singapore)	ap-southeast-1	opensearch.ap-southeast-1.amazonaws.com es.ap-southeast-1.amazonaws.com	HTTPS HTTPS
Asia Pacifico (Sydney)	ap-southeast-2	opensearch.ap-southeast-2.amazonaws.com es.ap-southeast-2.amazonaws.com	HTTPS HTTPS

Nome Regione	Regione	Endpoint	Protocollo
Canada (Centrale)	ca-central-1	opensearch.ca-central-1.amazonaws.com es.ca-central-1.amazonaws.com	HTTPS HTTPS
Europa (Stoccolma)	eu-north-1	opensearch.eu-north-1.amazonaws.com es.eu-north-1.amazonaws.com	HTTPS HTTPS
Asia Pacific (Hong Kong)	ap-east-1	opensearch.ap-east-1.amazonaws.com es.ap-east-1.amazonaws.com	HTTPS HTTPS
Asia Pacific (Seoul)	ap-northeast-2	opensearch.ap-northeast-2.amazonaws.com es.ap-northeast-2.amazonaws.com	HTTPS HTTPS
Asia Pacifico (Osaka-Locale)	ap-northeast-3	opensearch.ap-northeast-3.amazonaws.com es.ap-northeast-3.amazonaws.com	HTTPS HTTPS
Asia Pacifico (Hyderabad)	ap-south-2	opensearch.ap-south-2.amazonaws.com es.ap-south-2.amazonaws.com	HTTPS HTTPS
Europa (Spagna)	eu-south-2	opensearch.eu-south-2.amazonaws.com es.eu-south-2.amazonaws.com	HTTPS HTTPS
Stati Uniti occidentali (California settentrionale)	us-west-1	opensearch.us-west-1.amazonaws.com es.us-west-1.amazonaws.com	HTTPS HTTPS

Nome Regione	Regione	Endpoint	Protocollo
Europa (Zurigo)	eu-central-2	opensearch.eu-central-2.amazonaws.com es.eu-central-2.amazonaws.com	HTTPS HTTPS
Europe (Milan)	eu-south-1	opensearch.eu-south-1.amazonaws.com es.eu-south-1.amazonaws.com	HTTPS HTTPS

OpenSearch quote di servizio UI

Il tuo AWS account ha le seguenti quote relative alle risorse OpenSearch dell'interfaccia utente.

Nome	Predefinita	Adattabile	Note
OpenSearch Applicazioni dell'interfaccia utente per account per regione	30	Sì	Il numero massimo di applicazioni OpenSearch UI che puoi creare per account per regione. Puoi aumentare il limite a 50 utilizzando la quota di servizio e farlo approvare automaticamente. Per richiedere un limite più alto, invia un ticket di assistenza.

Gestione degli indici in Amazon Service OpenSearch

Dopo aver aggiunto dati ad Amazon OpenSearch Service, spesso devi reindicizzarli, utilizzare gli alias dell'indice, spostare un indice su uno storage più conveniente o eliminarlo del tutto. Questo capitolo tratta le soluzioni di storage su più livelli, l'Index State Management e altre operazioni di indicizzazione. Per informazioni sull' OpenSearch indice APIs, consulta la OpenSearch documentazione.

Argomenti

- OpenSearch Istanze ottimizzate per i domini Amazon OpenSearch Service
- Storage a più livelli per Amazon Service OpenSearch
- UltraWarm spazio di archiviazione per Amazon OpenSearch Service
- Conservazione a freddo per Amazon OpenSearch Service
- Gestione dello stato dell'indice in Amazon OpenSearch Service
- Riepilogo degli indici in Amazon OpenSearch Service con indici cumulativi
- Trasformazione degli indici in Amazon Service OpenSearch
- Replica tra cluster per Amazon Service OpenSearch
- Migrazione degli indici di Amazon OpenSearch Service utilizzando la reindicizzazione remota
- Gestione dei dati di serie temporali in Amazon OpenSearch Service con flussi di dati

OpenSearch Istanze ottimizzate per i domini Amazon OpenSearch Service

La famiglia di istanze OpenSearch ottimizzata per Amazon OpenSearch Service è una soluzione conveniente per l'archiviazione di grandi volumi di dati. Un dominio con istanze OpenSearch ottimizzate utilizza lo storage locale come principale, con i dati copiati in modo sincrono su Amazon S3 non appena arrivano. Questa struttura di storage offre una maggiore velocità di indicizzazione con un'elevata durabilità. OR1 OR2, OM2 utilizza Amazon Elastic Block Store (Amazon EBSgp3) `io1` o volumi a livello locale, OI2 mentre le istanze utilizzano dischi locali. NVMe La famiglia di istanze OpenSearch ottimizzata supporta anche il ripristino automatico dei dati in caso di errore. Per informazioni sulle opzioni di tipo di istanza OpenSearch ottimizzate, consultathe section called “Tipi di istanza della generazione attuale”.

Se esegui l'indicizzazione di pesanti carichi di lavoro di analisi operativa come l'analisi dei log, l'osservabilità o l'analisi della sicurezza, puoi trarre vantaggio dalle migliori prestazioni ed efficienza di calcolo delle istanze ottimizzate. OpenSearch Inoltre, il ripristino automatico dei dati offerto dalle istanze OpenSearch ottimizzate migliora l'affidabilità complessiva del dominio.

OpenSearch Il servizio invia ad Amazon i parametri delle istanze OpenSearch ottimizzati relativi allo storage. CloudWatch Per un elenco di parametri disponibili, consulta ???.

OpenSearch le istanze ottimizzate sono disponibili su richiesta o con prezzi per istanze riservate, con una tariffa oraria per le istanze e lo storage forniti in Amazon EBS e Amazon S3.

Argomenti

- Limitazioni
- Ottimizzazione per una migliore velocità di ingestione
- In che modo le istanze OpenSearch ottimizzate differiscono dalle altre istanze
- In che modo le istanze OpenSearch ottimizzate differiscono dalle istanze UltraWarm
- Fornitura di istanze OpenSearch ottimizzate a un dominio

Limitazioni

Considera le seguenti limitazioni quando utilizzi istanze ottimizzate per il tuo dominio. OpenSearch

- I domini appena creati devono eseguire la OpenSearch versione 2.11 o successiva.
- I domini esistenti devono eseguire OpenSearch la versione 2.15 o successiva.
- Il dominio deve avere la crittografia a riposo abilitata. Per ulteriori informazioni, consulta ???.
- Se il tuo dominio utilizza nodi master dedicati, devono utilizzare istanze Graviton. Per ulteriori informazioni sui nodi master dedicati, consulta. ???
- L'intervallo di aggiornamento per gli indici sulle istanze OpenSearch ottimizzate deve essere pari o superiore a 10 secondi. L'intervallo di aggiornamento predefinito per le istanze ottimizzate è di 10 secondi. OpenSearch

Ottimizzazione per una migliore velocità di ingestione

Per ottenere la migliore velocità di indicizzazione dalle istanze OpenSearch ottimizzate, ti consigliamo di fare quanto segue:

- Utilizza grandi dimensioni in blocco per migliorare l'utilizzo del buffer. La dimensione consigliata è di 10 MB.
- Utilizza più client per migliorare le prestazioni di elaborazione parallela.
- Imposta il numero di shard primari attivi in modo che corrisponda al numero di nodi di dati per massimizzare l'utilizzo delle risorse.

In che modo le istanze OpenSearch ottimizzate differiscono dalle altre istanze

OpenSearch le istanze ottimizzate si differenziano dalle istanze non ottimizzate nei seguenti modi:

- Per le istanze OpenSearch ottimizzate, l'indicizzazione viene eseguita solo sugli shard primari.
- Se le istanze OpenSearch ottimizzate sono configurate con repliche, la velocità di indicizzazione potrebbe apparire inferiore a quella effettiva. Ad esempio, se sono presenti uno shard primario e uno shard di replica, la velocità di indicizzazione potrebbe mostrare una frequenza di 1000 quando la velocità di indicizzazione effettiva è 2000.
- OpenSearch le istanze ottimizzate eseguono operazioni di buffer prima dell'invio a una fonte remota. Ciò si traduce in latenze di ingestione più elevate.

  > **Note**
  >
  > La `IndexingLatency` metrica non è influenzata, in quanto non include il tempo di sincronizzazione del translog.

- I frammenti di replica possono trovarsi alcuni secondi indietro rispetto ai frammenti primari. Puoi monitorare il ritardo utilizzando la metrica di `ReplicationLagMaxTime` Amazon CloudWatch

In che modo le istanze OpenSearch ottimizzate differiscono dalle istanze UltraWarm

OpenSearch Il servizio fornisce UltraWarm istanze che rappresentano un modo conveniente per archiviare grandi quantità di dati di sola lettura. Sia le UltraWarm istanze OpenSearch ottimizzate che quelle archiviano i dati localmente in Amazon EBS e in remoto in Amazon S3. Tuttavia, le istanze e le UltraWarm istanze OpenSearch ottimizzate differiscono in diversi modi importanti:

- OpenSearch le istanze ottimizzate conservano una copia dei dati nell'archivio locale e remoto. In alcuni UltraWarm casi, i dati vengono conservati principalmente in un archivio remoto per ridurre i costi di archiviazione. A seconda dei modelli di utilizzo, i dati possono essere spostati nell'archiviazione locale.
- OpenSearch le istanze ottimizzate sono attive e possono accettare operazioni di lettura e scrittura, mentre i dati sulle UltraWarm istanze sono di sola lettura finché non vengono spostati manualmente nella memoria a caldo.
- UltraWarm si affida alle istantanee degli indici per la durabilità dei dati. OpenSearch le istanze ottimizzate, al confronto, eseguono la replica e il ripristino dietro le quinte. In caso di indice rosso, le istanze OpenSearch ottimizzate ripristineranno automaticamente gli shard mancanti dallo storage remoto in Amazon S3. Il tempo di ripristino varia a seconda del volume di dati da recuperare.

Per ulteriori informazioni sullo UltraWarm storage, vedere???.

Fornitura di istanze OpenSearch ottimizzate a un dominio

Puoi selezionare istanze OpenSearch ottimizzate per i tuoi nodi di dati quando crei un nuovo dominio con il Console di gestione AWS o il AWS Command Line Interface ().AWS CLIÈ quindi possibile indicizzare e interrogare i dati utilizzando gli strumenti esistenti.

Console

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Domains (Domini).
3. Scegli Crea dominio.
4. Nella sezione Numero di nodi dati, espandi il menu della famiglia di istanze e scegli OpenSearch ottimizzato.
5. Scegli il tipo di istanza e altre impostazioni di archiviazione.
6. Nella sezione Crittografia, assicurati che sia selezionata l'opzione Abilita la crittografia dei dati inattivi.
7. Configura il resto del dominio e scegli Crea.

AWS CLI

Per effettuare il provisioning di un dominio che utilizza lo storage OpenSearch ottimizzato utilizzando il AWS CLI, devi fornire il valore della dimensione del tipo di istanza specifico (ad esempio OR1 OR2 OM2,, o OI2) in`InstanceType`.

L'esempio seguente crea un dominio con OR1 istanze di dimensioni diverse `2xlarge` e abilita la crittografia a riposo.

```
aws opensearch create-domain \
  --domain-name test-domain \
  --engine-version OpenSearch_2.11 \
  --cluster-config
 "InstanceType=or1.2xlarge.search,InstanceCount=3,DedicatedMasterEnabled=true,DedicatedMasterTy
 \
  --ebs-options "EBSEnabled=true,VolumeType=gp3,VolumeSize=200" \
  --encryption-at-rest-options Enabled=true \
  --advanced-security-options
 "Enabled=true,InternalUserDatabaseEnabled=true,MasterUserOptions={MasterUserName=test-
user,MasterUserPassword=test-password}" \
  --node-to-node-encryption-options Enabled=true \
  --domain-endpoint-options EnforceHTTPS=true \
  --access-policies '{"Version": "2012-10-17",       "Statement":
[{"Effect":"Allow","Principal":{"AWS":"*"},"Action":"es:*","Resource":"arn:aws:es:us-
east-1:account-id:domain/test-domain/*"}]}'
```

L'esempio seguente crea un dominio con OI2 istanze di grandi dimensioni `large` e abilita la crittografia a riposo. Tieni presente che OI2 le istanze non richiedono la configurazione EBS poiché utilizzano l'archiviazione locale. NVMe

```
aws opensearch create-domain \
  --domain-name test-domain-oi2 \
  --engine-version OpenSearch_2.11 \
  --cluster-config
 "InstanceType=oi2.2xlarge.search,InstanceCount=3,DedicatedMasterEnabled=true,DedicatedMasterTy
 \
  --encryption-at-rest-options Enabled=true \
  --advanced-security-options
 "Enabled=true,InternalUserDatabaseEnabled=true,MasterUserOptions={MasterUserName=test-
user,MasterUserPassword=test-password}" \
  --node-to-node-encryption-options Enabled=true \
  --domain-endpoint-options EnforceHTTPS=true \
```

```
  --access-policies '{"Version": "2012-10-17",       "Statement":
[{"Effect":"Allow","Principal":{"AWS":"*"},"Action":"es:*","Resource":"arn:aws:es:us-
east-1:account-id:domain/test-domain-oi2/*"}]}'
```

Storage a più livelli per Amazon Service OpenSearch

Lo storage multilivello per Amazon OpenSearch Service è una soluzione intelligente di gestione dei dati che ottimizza prestazioni e costi gestendo i dati su diversi livelli di storage. Questa architettura consente alle organizzazioni di bilanciare in modo efficiente il compromesso tra prestazioni e costi conservando i dati a cui si accede di frequente in hot storage ad alte prestazioni e spostando i dati a cui si accede meno frequentemente verso sistemi di storage caldi più convenienti.

Amazon OpenSearch Service offre due opzioni di architettura per il livello hot/warm di storage:

- OpenSearch Architettura di storage a più livelli
  - Combina Amazon S3 con lo storage di istanze locali
  - Basato su istanze OpenSearch ottimizzate
  - Supporta le operazioni di scrittura in modalità warm tier
  - Supporta la migrazione dei dati senza interruzioni tra il livello caldo e quello caldo
  - Disponibile su OpenSearch 3.3 e versioni successive
  - Non supporta Cold Tier
- UltraWarmarchitettura basata
  - Combina Amazon S3 con lo storage di istanze locali
  - Basato su istanze UltraWarm
  - Ottimizzato per carichi di lavoro di sola lettura di livello caldo
  - Disponibile nella versione 6.8 e successive di Elasticsearch e in tutte le versioni OpenSearch
  - Supporta Cold Tier

> **Note**
>
> Questa documentazione si concentra solo sull'architettura a più livelli. Per l'architettura di storage Ultrawarm, vedi Ultrawarm and Cold Storage

Architettura di storage a più livelli

Argomenti

- Vantaggi principali
- Prerequisiti
- Limitazioni
- Da sapere
- Creazione di un dominio a più livelli
- Gestione delle migrazioni a più livelli
- Configurazione di sicurezza
- Best practice
- Monitoraggio delle metriche

Vantaggi principali

- Writable Warm: supporta le operazioni di scrittura su indici caldi
- Migrazione senza interruzioni: spostamento senza interruzioni dei dati tra i livelli di storage
- Ottimizzazione dei costi: riduci i costi di storage spostando i dati meno attivi in sistemi di storage a caldo a costi contenuti
- Miglioramento delle prestazioni: mantenimento di prestazioni elevate per i dati a cui si accede di frequente nella fascia più calda
- Gestione flessibile dei dati: scegli l'architettura più adatta ai requisiti del tuo carico di lavoro
- Gestione automatizzata: gestione semplificata del ciclo di vita dei dati su più livelli di storage

Prerequisiti

- Versione del motore: 3.3 o successiva OpenSearch
- Famiglie di istanze:
  - Hot nodes: OR1 OR2, OM2, o OI2
  - Nodi caldi: OI2
- Sicurezza: Node-to-node crittografia, crittografia a riposo, HTTPS applicato

Limitazioni

- Funziona su tutti i domini con istanze OpenSearch ottimizzate che non hanno già abilitato Ultrawarm
- Nessun supporto per la modalità cold tier

Da sapere

- Le migrazioni da «hot to warm» non attivano un'unione forzata in un'architettura a più livelli. Se necessario, l'utente può comunque orchestrare le unioni forzate utilizzando la politica di gestione degli indici.
- Oltre all'indicizzazione, i nodi caldi ora eseguono anche operazioni di unione in background (simili a quelle a caldo).
- Tutte le richieste di ricerca sugli indici caldi vengono indirizzate allo shard primario e la replica serve le letture solo quando lo shard primario è inattivo.
- Questa architettura supporta anche le istantanee automatizzate per gli indici caldi
- La replica tra cluster è supportata solo per gli hot index
- Indici APIs come Shrink, Split e Clone non funzionano con indici caldi.

Creazione di un dominio a più livelli

Fase 1: Creare il dominio

```
aws opensearch create-domain \
      --domain-name my-domain \
      --engine-version OpenSearch_3.3 \
      --cluster-config
 InstanceCount=3,InstanceType=or2.large.search,DedicatedMasterEnabled=true,DedicatedMasterType=
 \
      --ebs-options EBSEnabled=true,VolumeType=gp2,VolumeSize=11 \
      --node-to-node-encryption-options Enabled=true \
      --encryption-at-rest-options Enabled=true \
      --domain-endpoint-options EnforceHTTPS=true,TLSSecurityPolicy=Policy-Min-
TLS-1-2-2019-07 \
      --advanced-security-options
 '{"Enabled":true,"InternalUserDatabaseEnabled":true,"MasterUserOptions":
{"MasterUserName":"user_name","MasterUserPassword":"your_pass"}}' \
```

```
      --access-policies '{"Version": "2012-10-17",        "Statement":
[{"Effect":"Allow","Principal":"*","Action":"es:*","Resource":"*"}]}' \
      --region us-east-1
```

Fase 2: Verifica dei nodi caldi

```
aws opensearch describe-domain-nodes --domain-name my-domain --region us-east-1
```

Esempio di risposta (estratto):

```
{
      "NodeType": "Warm",
      "InstanceType": "oi2.large.search",
      "NodeStatus": "Active"
    }
```

Gestione delle migrazioni a più livelli

Supporto per domini multilivello:

- Nuovo tiering APIs per un'esperienza semplificata
- Legacy UltraWarm APIs per la compatibilità

Nuova suddivisione in più livelli APIs

Migra un indice su warm:

```
curl -XPOST 'https://localhost:9200/index-name/_tier/warm'
```

Risposta:

```
{"acknowledged": true}
```

Migrare un indice su hot:

```
curl -XPOST 'https://localhost:9200/index-name/_tier/hot'
```

Risposta:

```
{"acknowledged": true}
```

Verifica lo stato del tiering:

```
curl -XGET 'https://localhost:9200/index-name/_tier'
```

Risposta di esempio:

```
{
      "tiering_status": {
         "index": "index-name",
         "state": "RUNNING_SHARD_RELOCATION",
         "source": "HOT",
         "target": "WARM",
         "start_time": 1745836500563,
         "shard_level_status": {
           "running": 0,
           "total": 100,
           "pending": 100,
           "succeeded": 0
         }
      }
    }
```

Vista condivisa dettagliata:

```
curl 'https://localhost:9200/index1/_tier?detailed=true'
```

Elenca tutte le migrazioni in corso (testo):

```
curl 'https://localhost:9200/_tier/all'
```

Elenca tutte le migrazioni in corso (JSON):

```
curl 'https://localhost:9200/_tier/all?format=json'
```

Filtra per livello di destinazione:

```
curl 'https://localhost:9200/_tier/all?target=_warm'
```

Legacy UltraWarm APIs per la compatibilità

Migra a warm:

```
curl -XPOST localhost:9200/_ultrawarm/migration/index2/_warm
```

Migrazione a caldo:

```
curl -XPOST localhost:9200/_ultrawarm/migration/index2/_hot
```

Verifica lo stato:

```
curl -XGET localhost:9200/_ultrawarm/migration/index2/_status
```

Configurazione di sicurezza

Se abiliti lo storage a più livelli su un dominio Amazon OpenSearch Service preesistente, il `storage_tiering_manager` ruolo potrebbe non essere definito nel dominio. Gli utenti senza privilegi di amministratore devono essere mappati a questo ruolo in modo da gestire gli indici a caldo sui domini che utilizzano il controllo granulare degli accessi. Per creare manualmente il ruolo `storage_tiering_manager`, procedere nel seguente modo:

1. In OpenSearch Dashboards, vai su Sicurezza e scegli Autorizzazioni.
2. Scegliere Crea gruppo di operazioni e configurare i seguenti gruppi:

Nome del Gruppo	Permissions
storage_tiering_cluster	indici:admin/_tier/all
storage_tiering_index_read	indici:admin/_ tier/get, indices:admin/get
storage_tiering_index_write	indici:admin/_ _to_hot tier/hot_to_warm, indices:admin/_tier/warm

3. Scegliere Ruoli, quindi selezionare Crea ruolo.
4. Denomina il ruolo `storage_tiering_manager`.
5. Per Autorizzazioni cluster, selezionare `storage_tiering_cluster` e `cluster_monitor`.

6. Per Indice, digitare *.
7. Per le autorizzazioni relative all'indice, seleziona e. `storage_tiering_index_read` `storage_tiering_index_write` `indices_monitor`
8. Scegli Create (Crea).
9. Dopo aver creato il ruolo, associalo a qualsiasi ruolo utente o di backend che gestirà gli indici a più livelli.

Best practice

Quando implementi lo storage a più livelli nei tuoi domini Amazon OpenSearch Service, prendi in considerazione le seguenti best practice:

- Rivedi regolarmente i tuoi modelli di accesso ai dati per ottimizzare l'allocazione dei livelli
- Monitora le metriche delle prestazioni per garantire un utilizzo efficiente delle risorse
- Utilizza il nuovo tiering APIs per un controllo granulare sulla migrazione dei dati

Monitoraggio delle metriche

I domini di storage multilivello forniscono parametri aggiuntivi per il monitoraggio delle prestazioni dei livelli più caldi. Queste metriche includono sia metriche esistenti che nuove UltraWarm metriche specifiche dell'architettura Optimized Instances: OpenSearch

Nuovi parametri

Nome parametro	Statistiche a livello di nodo	Statistiche a livello di cluster	Granularity (Granularità)
WarmIndexingLatency	Media	Media	1 minuto
WarmIndexingRate	Media	Media, massima, somma	1 minuto
WarmThreadpoolIndexingQueue	Massimo	Somma, massimo, media	1 minuto
WarmThreadpoolIndexingRejected	Massimo	Somma	1 minuto

Nome parametro	Statistiche a livello di nodo	Statistiche a livello di cluster	Granularity (Granularità)
WarmThreadpoolIndexingThreads	Massimo	Somma, media	1 minuto

UltraWarm spazio di archiviazione per Amazon OpenSearch Service

UltraWarm offre un modo conveniente per archiviare grandi quantità di dati di sola lettura su Amazon Service. OpenSearch I nodi di dati standard utilizzano l'archiviazione "hot", che assume la forma di archivi di istanze o volumi Amazon EBS collegati a ciascun nodo. L'archiviazione a caldo offre le prestazioni più veloci possibili per l'indicizzazione e la ricerca di nuovi dati.

Invece dello storage collegato, UltraWarm i nodi utilizzano Amazon S3 e una sofisticata soluzione di caching per migliorare le prestazioni. Per gli indici su cui non scrivi attivamente, esegui query meno frequentemente e che non richiedono le stesse prestazioni, UltraWarm offre costi notevolmente inferiori per GiB di dati. Poiché gli indici caldi sono di sola lettura a meno che non vengano restituiti all'archiviazione a caldo, sono più adatti per dati immutabili, UltraWarm come i log.

In OpenSearch, gli indici caldi si comportano come qualsiasi altro indice. Puoi interrogarli utilizzando lo stesso APIs o usarli per creare visualizzazioni nelle dashboard. OpenSearch

Argomenti

- Prerequisiti
- UltraWarm requisiti di archiviazione e considerazioni sulle prestazioni
- UltraWarm prezzi
- Abilitazione UltraWarm
- Migrazione degli indici verso lo storage UltraWarm
- Automazione delle migrazioni
- Regolazione della migrazione
- Annullamento di migrazioni
- Elenco degli indici ad accesso frequente e degli indici a caldo
- Ritorno di indici a caldo all'archiviazione ad accesso frequente

- Ripristino degli indici caldi dalle istantanee
- Snapshot manuali di indici a caldo
- Migrazione degli indici a caldo all'archiviazione a freddo
- Procedure consigliate per gli indici KNN
- Disabilitazione UltraWarm

Prerequisiti

UltraWarm presenta alcuni prerequisiti importanti:

- UltraWarm richiede Elasticsearch 6.8 OpenSearch o versione successiva.
- Per utilizzare l'archiviazione warm, i domini devono disporre di nodi master dedicati.
- Quando si utilizza un dominio Multi-AZ con standby, il numero di nodi caldi deve essere un multiplo del numero di zone di disponibilità utilizzate.
- Se il dominio utilizza un tipo di istanza T2 o T3 per i nodi di dati, non è possibile utilizzare l'archiviazione a caldo.
- Se l'indice utilizza un valore approssimativo di k-NN (`"index.knn":true`), è possibile spostarlo nella memoria a caldo a partire dalla versione 2.17 e successive. I domini delle versioni precedenti alla 2.17 possono essere aggiornati alla 2.17 per utilizzare questa funzionalità, ma gli indici KNN creati nelle versioni precedenti alla 2.x non possono migrare verso. UltraWarm
- Se il dominio utilizza un controllo granulare degli accessi, gli utenti devono essere mappati al ruolo nelle dashboard per effettuare chiamate API. `ultrawarm_manager` OpenSearch UltraWarm

> **Note**
>
> Il `ultrawarm_manager` ruolo potrebbe non essere definito in alcuni domini di servizio preesistenti. OpenSearch Se il ruolo non è visualizzato in Dashboards, sarà necessario crearlo manualmente.

Configurazione delle autorizzazioni

Se si abilita UltraWarm su un dominio di OpenSearch servizio preesistente, il `ultrawarm_manager` ruolo potrebbe non essere definito nel dominio. Gli utenti senza privilegi di amministratore devono

essere mappati a questo ruolo in modo da gestire gli indici a caldo sui domini che utilizzano il controllo granulare degli accessi. Per creare manualmente il ruolo `ultrawarm_manager`, procedere nel seguente modo:

1. In OpenSearch Dashboard, vai su Sicurezza e scegli Autorizzazioni.
2. Scegliere Crea gruppo di operazioni e configurare i seguenti gruppi:

Group name (Nome gruppo)	Permissions
`ultrawarm_cluster`	• `cluster:admin/ultrawarm/migration/list` • `cluster:monitor/nodes/stats`
`ultrawarm_index_read`	• `indices:admin/ultrawarm/migration/get` • `indices:admin/get`
`ultrawarm_index_write`	• `indices:admin/ultrawarm/migration/warm` • `indices:admin/ultrawarm/migration/hot` • `indices:monitor/stats` • `indices:admin/ultrawarm/migration/cancel`

3. Scegliere Ruoli, quindi selezionare Crea ruolo.
4. Denominare il ruolo ultrawarm_manager.
5. Per Autorizzazioni cluster, selezionare `ultrawarm_cluster` e `cluster_monitor`.
6. Per Indice, digitare `*`.
7. Per Autorizzazioni indice, selezionare `ultrawarm_index_read`, `ultrawarm_index_write` e `indices_monitor`.
8. Scegli Create (Crea).
9. Dopo aver creato il ruolo, associalo a qualsiasi utente o ruolo di backend che gestirà gli indici. UltraWarm

UltraWarm requisiti di archiviazione e considerazioni sulle prestazioni

Come illustrato in precedenzathe section called “Calcolo dei requisiti di archiviazione”, i dati nell'archiviazione a caldo comportano un sovraccarico significativo: repliche, spazio riservato Linux

e OpenSearch spazio riservato ai servizi. Ad esempio, una partizione primaria da 20 GiB con una partizione di replica richiede circa 58 GiB di archiviazione ad accesso frequente.

Poiché utilizza Amazon S3, non UltraWarm comporta alcun sovraccarico. Nel calcolo dei requisiti UltraWarm di archiviazione, si considera solo la dimensione degli shard primari. La durata dei dati in S3 elimina la necessità di repliche e S3 elimina tutte le considerazioni relative al sistema operativo o al servizio. La stessa partizione da 20 GiB richiede 20 GiB di archiviazione a caldo. Se si esegue il provisioning di un'istanza `ultrawarm1.large.search`, è possibile utilizzare tutti i 20 TiB della relativa archiviazione massima per le partizioni primarie. Vedere the section called “UltraWarm quote di archiviazione” per un riepilogo dei tipi di istanza e la quantità massima di archiviazione che ciascuna può gestire.

Tuttavia UltraWarm, consigliamo comunque una dimensione massima dello shard di 50 GiB. Il numero di core della CPU e la quantità di RAM allocata a ciascun tipo di UltraWarm istanza ti danno un'idea del numero di shard che possono essere cercati contemporaneamente. Tieni presente che, sebbene solo gli shard primari contino ai fini UltraWarm dello storage in S3, in OpenSearch Dashboards le dimensioni UltraWarm dell'indice vengono `_cat/indices` comunque riportate come il totale di tutti gli shard primari e di replica.

Ad esempio, ogni istanza `ultrawarm1.medium.search` ha due core CPU e può indirizzare fino a 1,5 TiB di archiviazione su S3. Due di queste istanze hanno un'archiviazione combinata di 3 TiB, che corrisponde a circa 62 partizioni se ogni partizione è 50 GiB. Se una richiesta al cluster esegue la ricerca solo in quattro di queste partizioni, le prestazioni potrebbero essere eccellenti. Se la richiesta è ampia ed esegue la ricerca in tutte le 62 partizioni, i quattro core della CPU potrebbero faticare per eseguire l'operazione. Monitora le `WarmJVMMemoryPressure` UltraWarm metriche `WarmCPUUtilization` e per capire come le istanze gestiscono i carichi di lavoro.

Se le ricerche sono vaste o frequenti, valuta la possibilità di lasciare gli indici nell'archiviazione ad accesso frequente. Proprio come qualsiasi altro OpenSearch carico di lavoro, il passaggio più importante per determinare se UltraWarm soddisfa le proprie esigenze è eseguire test rappresentativi sui clienti utilizzando un set di dati realistico.

UltraWarm prezzi

Con l'archiviazione ad accesso frequente, si paga per ciò che si fornisce. Alcune istanze richiedono un volume Amazon EBS allegato, mentre altre includono un archivio di istanze. Se lo spazio di archiviazione è vuoto o pieno, si paga lo stesso prezzo.

Con UltraWarm lo spazio di archiviazione, paghi in base a ciò che usi. Un'istanza `ultrawarm1.large.search` può indirizzare fino a 20 TiB di archiviazione su S3, ma se si memorizza solo 1 TiB di dati, viene addebitato solo 1 TiB di dati. Come tutti gli altri tipi di nodi, paghi anche una tariffa oraria per ogni UltraWarm nodo. Per ulteriori informazioni, consulta the section called “Prezzi”.

Abilitazione UltraWarm

La console è il modo più semplice per creare un dominio che utilizza l'archiviazione calda. Durante la creazione del dominio, scegli Abilita nodi dati caldi e il numero di nodi caldi che desideri. Lo stesso processo di base funziona sui domini esistenti, a condizione che soddisfino i prerequisiti. Anche dopo che lo stato del dominio è cambiato da Processing a Active, UltraWarm potrebbe non essere disponibile all'uso per diverse ore.

Quando si utilizza un dominio Multi-AZ con standby, il numero di nodi caldi deve essere un multiplo del numero di zone di disponibilità. Per ulteriori informazioni, consulta the section called “Multi-AZ con Standby”.

Puoi anche utilizzare l'API AWS CLIo di configurazione per abilitare UltraWarm, in particolare `WarmEnabled` `WarmCount`, e `WarmType` le opzioni in. `ClusterConfig`

> **Note**
>
> I domini supportano un numero massimo di nodi a caldo. Per informazioni dettagliate, vedi the section called “Quote”.

Comando CLI di esempio

Il AWS CLI comando seguente crea un dominio con tre nodi di dati, tre nodi master dedicati, sei nodi caldi e un controllo granulare degli accessi abilitato:

```
aws opensearch create-domain \
  --domain-name my-domain \
  --engine-version Opensearch_1.0 \
  --cluster-config
 InstanceCount=3,InstanceType=r6g.large.search,DedicatedMasterEnabled=true,DedicatedMasterType=
 \
  --ebs-options EBSEnabled=true,VolumeType=gp2,VolumeSize=11 \
```

```
  --node-to-node-encryption-options Enabled=true \
  --encryption-at-rest-options Enabled=true \
  --domain-endpoint-options EnforceHTTPS=true,TLSSecurityPolicy=Policy-Min-
TLS-1-2-2019-07 \
  --advanced-security-options
 Enabled=true,InternalUserDatabaseEnabled=true,MasterUserOptions='{MasterUserName=master-
user,MasterUserPassword=master-password}' \
  --access-policies '{"Version": "2012-10-17",       "Statement":
[{"Effect":"Allow","Principal":{"AWS":["123456789012"]},"Action":
["es:*"],"Resource":"arn:aws:es:us-west-1:123456789012:domain/my-domain/*"}]}' \
  --region us-east-1
```

Per ulteriori informazioni, consultare Riferimento ai comandi AWS CLI.

Richiesta all'API di configurazione di esempio

La seguente richiesta all'API di configurazione crea un dominio con tre nodi di dati, tre nodi master dedicati e sei nodi a caldo con il controllo granulare degli accessi abilitato e una policy di accesso restrittiva:

```
POST https://es.us-east-2.amazonaws.com/2021-01-01/opensearch/domain
{
  "ClusterConfig": {
    "InstanceCount": 3,
    "InstanceType": "r6g.large.search",
    "DedicatedMasterEnabled": true,
    "DedicatedMasterType": "r6g.large.search",
    "DedicatedMasterCount": 3,
    "ZoneAwarenessEnabled": true,
    "ZoneAwarenessConfig": {
      "AvailabilityZoneCount": 3
    },
    "WarmEnabled": true,
    "WarmCount": 6,
    "WarmType": "ultrawarm1.medium.search"
  },
  "EBSOptions": {
    "EBSEnabled": true,
    "VolumeType": "gp2",
    "VolumeSize": 11
  },
  "EncryptionAtRestOptions": {
    "Enabled": true
```

```
  },
  "NodeToNodeEncryptionOptions": {
    "Enabled": true
  },
  "DomainEndpointOptions": {
    "EnforceHTTPS": true,
    "TLSSecurityPolicy": "Policy-Min-TLS-1-2-2019-07"
  },
   "AdvancedSecurityOptions": {
    "Enabled": true,
    "InternalUserDatabaseEnabled": true,
    "MasterUserOptions": {
      "MasterUserName": "master-user",
      "MasterUserPassword": "master-password"
    }
  },
  "EngineVersion": "Opensearch_1.0",
  "DomainName": "my-domain",
  "AccessPolicies": "{\"Version\":\"2012-10-17\",\"Statement\":[{\"Effect\":\"Allow
\",\"Principal\":{\"AWS\":[\"123456789012\"]},\"Action\":[\"es:*\"],\"Resource\":
\"arn:aws:es:us-east-1:123456789012:domain/my-domain/*\"}]}"
}
```

Per informazioni dettagliate, consulta l'Amazon OpenSearch Service API Reference.

Migrazione degli indici verso lo storage UltraWarm

Se hai finito di scrivere su un indice e non hai più bisogno delle prestazioni di ricerca più veloci possibili, esegui la migrazione da hot a: UltraWarm

```
POST _ultrawarm/migration/my-index/_warm
```

Quindi controlla lo stato della migrazione:

```
GET _ultrawarm/migration/my-index/_status

{
  "migration_status": {
    "index": "my-index",
    "state": "RUNNING_SHARD_RELOCATION",
    "migration_type": "HOT_TO_WARM",
    "shard_level_status": {
```

```
      "running": 0,
      "total": 5,
      "pending": 3,
      "failed": 0,
      "succeeded": 2
    }
  }
}
```

Per eseguire una migrazione, l'integrità dell'indice deve essere verde. Se si esegue la migrazione di diversi indici in rapida successione, è possibile ottenere un riepilogo di tutte le migrazioni in chiaro, simile all'API `_cat`:

```
GET _ultrawarm/migration/_status?v

index    migration_type state
my-index HOT_TO_WARM    RUNNING_SHARD_RELOCATION
```

OpenSearch Il servizio esegue la migrazione di un indice alla volta su. UltraWarm In coda possono essere presenti fino a 200 migrazioni. Qualsiasi richiesta che supera il limite verrà rifiutata. Per controllare il numero corrente di migrazioni nella coda, monitorare il parametro`HotToWarmMigrationQueueSize`. Gli indici rimangono disponibili durante tutto il processo di migrazione, senza tempi di inattività.

Il processo di migrazione ha i seguenti stati:

```
PENDING_INCREMENTAL_SNAPSHOT
RUNNING_INCREMENTAL_SNAPSHOT
FAILED_INCREMENTAL_SNAPSHOT
PENDING_FORCE_MERGE
RUNNING_FORCE_MERGE
FAILED_FORCE_MERGE
PENDING_FULL_SNAPSHOT
RUNNING_FULL_SNAPSHOT
FAILED_FULL_SNAPSHOT
PENDING_SHARD_RELOCATION
RUNNING_SHARD_RELOCATION
FINISHED_SHARD_RELOCATION
```

Come indicato da questi stati, le migrazioni potrebbero non riuscire durante gli snapshot, le traslocazioni di sezioni o le fusioni forzate. Gli errori durante gli snapshot o il trasferimento di

partizioni sono in genere dovuti a errori dei nodi o a problemi di connettività S3. La mancanza di spazio su disco è solitamente la causa sottostante degli errori di unioni forzate.

Al termine di una migrazione, la stessa richiesta `_status` restituisce un errore. Se controlli l'indice in quel momento, puoi visualizzare alcune impostazioni che sono univoche per gli indici a caldo:

```
GET my-index/_settings

{
  "my-index": {
    "settings": {
      "index": {
        "refresh_interval": "-1",
        "auto_expand_replicas": "false",
        "provided_name": "my-index",
        "creation_date": "1599241458998",
        "unassigned": {
          "node_left": {
            "delayed_timeout": "5m"
          }
        },
        "number_of_replicas": "1",
        "uuid": "GswyCdR0RSq0SJYmzsIpiw",
        "version": {
          "created": "7070099"
        },
        "routing": {
          "allocation": {
            "require": {
              "box_type": "warm"
            }
          }
        },
        "number_of_shards": "5",
        "merge": {
          "policy": {
            "max_merge_at_once_explicit": "50"
          }
        }
      }
    }
  }
}
```

- `number_of_replicas`, in questo caso, è il numero di repliche passive, che non consumano spazio su disco.
- `routing.allocation.require.box_type` specifica che l'indice deve utilizzare nodi caldi anziché nodi di dati standard.
- `merge.policy.max_merge_at_once_explicit` specifica il numero di segmenti da unire contemporaneamente durante la migrazione.

Gli indici nella memoria a caldo sono di sola lettura a meno che non vengano restituiti all'archiviazione a caldo, che li rende UltraWarm più adatti a dati immutabili, come i log. È possibile eseguire query sugli indici ed eliminarli, ma non è possibile aggiungere, aggiornare o eliminare singoli documenti. Se ci si prova, potrebbe essere restituito il seguente errore:

```
{
  "error" : {
    "root_cause" : [
      {
        "type" : "cluster_block_exception",
        "reason" : "index [indexname] blocked by: [TOO_MANY_REQUESTS/12/disk usage
 exceeded flood-stage watermark, index has read-only-allow-delete block];"
      }
    ],
    "type" : "cluster_block_exception",
    "reason" : "index [indexname] blocked by: [TOO_MANY_REQUESTS/12/disk usage exceeded
 flood-stage watermark, index has read-only-allow-delete block];"
  },
  "status" : 429
}
```

Automazione delle migrazioni

Si consiglia di utilizzare the section called “Index State Management” per automatizzare il processo di migrazione dopo che un indice raggiunge una certa età o soddisfa altre condizioni. Consultare la policy di esempio che illustra tale flusso di lavoro.

Regolazione della migrazione

Le migrazioni degli indici verso lo storage richiedono un'unione forzata. UltraWarm Ogni OpenSearch indice è composto da un certo numero di frammenti e ogni frammento è composto da un certo numero di segmenti di Lucene. L'operazione di unione forzata elimina i documenti contrassegnati per

l'eliminazione e consente di risparmiare spazio su disco. Per impostazione predefinita, UltraWarm unisce gli indici in un unico segmento, ad eccezione degli indici kNN, in cui viene utilizzato il valore predefinito di 20.

È possibile modificare questo valore fino a 1.000 segmenti utilizzando l'impostazione `index.ultrawarm.migration.force_merge.max_num_segments`. Valori più elevati velocizzano il processo di migrazione, ma aumentano la latenza delle query per l'indice a caldo al termine della migrazione. Per modificare l'impostazione, effettuare la seguente richiesta:

```
PUT my-index/_settings
{
  "index": {
    "ultrawarm": {
      "migration": {
        "force_merge": {
          "max_num_segments": 1
        }
      }
    }
  }
}
```

Per verificare quanto tempo richiede questa fase del processo di migrazione, monitorare il parametro `HotToWarmMigrationForceMergeLatency`.

Annullamento di migrazioni

UltraWarm gestisce le migrazioni in sequenza, in una coda. Se una migrazione è nella coda, ma non è ancora stata avviata, è possibile rimuoverla dalla coda utilizzando la seguente richiesta:

```
POST _ultrawarm/migration/_cancel/my-index
```

Se il dominio utilizza il controllo granulare degli accessi, per effettuare questa richiesta è necessaria l'autorizzazione `indices:admin/ultrawarm/migration/cancel`.

Elenco degli indici ad accesso frequente e degli indici a caldo

UltraWarm aggiunge due opzioni aggiuntive, simili a `_all`, per aiutare a gestire gli indici caldi e caldi. Per un elenco di tutti gli indici ad accesso frequente o a caldo, effettua le seguenti richieste:

```
GET _warm
GET _hot
```

Puoi utilizzare queste opzioni in altre richieste che specificano gli indici:

```
_cat/indices/_warm
_cluster/state/_all/_hot
```

Ritorno di indici a caldo all'archiviazione ad accesso frequente

Se è necessario scrivere nuovamente su un indice, eseguire la migrazione di nuovo all'archiviazione ad accesso frequente:

```
POST _ultrawarm/migration/my-index/_hot
```

È possibile avere fino a 10 migrazioni in coda dalla memorizzazione a caldo a quella a caldo alla volta. OpenSearch Il servizio elabora le richieste di migrazione una alla volta, nell'ordine in cui sono state messe in coda. Per controllare il numero corrente, monitorare il parametro `WarmToHotMigrationQueueSize`.

Al termine della migrazione, controllare le impostazioni dell'indice per assicurarti che soddisfino le tue esigenze. Gli indici tornano all'archiviazione ad accesso frequente con una replica.

Ripristino degli indici caldi dalle istantanee

Oltre al repository standard per le istantanee automatiche, UltraWarm aggiunge un secondo archivio per gli indici caldi,. `cs-ultrawarm` Ogni snapshot in questo repository contiene un solo indice. Se si elimina un indice a caldo, il relativo snapshot rimarrà nel repository `cs-ultrawarm` per altri 14 giorni, proprio come qualsiasi altro snapshot automatico.

Quando si ripristina uno snapshot da `cs-ultrawarm`, viene ripristinata l'archiviazione a caldo, non l'archiviazione ad accesso frequente. Gli snapshot nei repository `cs-automated` e `cs-automated-enc` ripristinano l'archiviazione ad accesso frequente.

Per ripristinare un'istantanea in una memoria calda UltraWarm

1. Identificare lo snapshot più recente che contiene l'indice che si desidera ripristinare:

   ```
   GET _snapshot/cs-ultrawarm/_all?verbose=false
   ```

```
{
  "snapshots": [{
    "snapshot": "snapshot-name",
    "version": "1.0",
    "indices": [
      "my-index"
    ]
  }]
}
```

> **Note**
>
> Per impostazione predefinita, l'`GET _snapshot/<repo>`operazione visualizza informazioni dettagliate sui dati come ora di inizio, ora di fine e durata per ogni istantanea all'interno di un repository. L'`GET _snapshot/<repo>`operazione recupera le informazioni dai file di ogni istantanea contenuta in un repository. Se non sono necessarie l'ora di inizio, l'ora di fine e la durata e si richiedono solo il nome e le informazioni sull'indice di un'istantanea, si consiglia di utilizzare il `verbose=false` parametro quando si elencano le istantanee per ridurre al minimo i tempi di elaborazione ed evitare il timeout.

2. Se l'indice esiste già, eliminalo:

```
DELETE my-index
```

Se non si desidera eliminare l'indice, riportarlo all'archiviazione ad accesso frequente e reindicizzarlo.

3. Ripristinare lo snapshot:

```
POST _snapshot/cs-ultrawarm/snapshot-name/_restore
```

UltraWarm ignora tutte le impostazioni dell'indice specificate in questa richiesta di ripristino, ma è possibile specificare opzioni come e. `rename_pattern` `rename_replacement` Per un riepilogo delle opzioni di ripristino delle OpenSearch istantanee, consulta la OpenSearch documentazione.

Snapshot manuali di indici a caldo

Puoi eseguire snapshot manuali di indici a caldo, ma non è consigliabile. Il repository `cs-ultrawarm` automatico contiene già uno snapshot per ogni indice a caldo, acquisito durante la migrazione, senza costi aggiuntivi.

Per impostazione predefinita, OpenSearch Service non include indici caldi nelle istantanee manuali. Ad esempio, la chiamata seguente include solo gli indici ad accesso frequente:

```
PUT _snapshot/my-repository/my-snapshot
```

Per eseguire snapshot manuali di indici a caldo, occorre fare delle considerazioni importanti.

- Non è possibile mescolare indici ad accesso frequente e indici a caldo. Ad esempio, la seguente richiesta ha esito negativo:

  ```
  PUT _snapshot/my-repository/my-snapshot
  {
    "indices": "warm-index-1,hot-index-1",
    "include_global_state": false
  }
  ```

  Se mescoli indici ad accesso frequente e indici a caldo, le istruzioni con caratteri jolly (*) non riescono.

- Puoi includere un solo indice a caldo per ogni snapshot. Ad esempio, la seguente richiesta ha esito negativo:

  ```
  PUT _snapshot/my-repository/my-snapshot
  {
    "indices": "warm-index-1,warm-index-2,other-warm-indices-*",
    "include_global_state": false
  }
  ```

  Questa richiesta ha esito positivo:

  ```
  PUT _snapshot/my-repository/my-snapshot
  {
    "indices": "warm-index-1",
    "include_global_state": false
  ```

```
}
```

- Gli snapshot manuali vengono sempre ripristinati nell'archiviazione ad accesso frequente, anche se originariamente includevano un indice a caldo.

Migrazione degli indici a caldo all'archiviazione a freddo

Se hai dati UltraWarm che richiedi raramente, valuta la possibilità di migrarli alla conservazione a freddo. L'archiviazione a freddo è pensata per i dati a cui si accede solo occasionalmente o che non sono più utilizzati. Non puoi leggere o scrivere su indici a freddo, ma puoi eseguirne di nuovo la migrazione allo storage a caldo senza alcun costo ogni volta che devi eseguire una query su tali indici. Per istruzioni, consulta Migrazione degli indici alla conservazione a freddo.

Procedure consigliate per gli indici KNN

- Il livello Ultrawarm/Cold è disponibile per tutti i tipi di motori a indice KNN. Lo consigliamo per gli indici KNN che utilizzano il motore Lucene e la ricerca vettoriale ottimizzata per disco, che non richiede il caricamento completo dei dati del grafico nella memoria off-heap. Durante l'utilizzo con motori nativi in memoria come FAISS e NMSLIB, è necessario tenere conto della dimensione del grafico degli shard su cui verrà eseguita la ricerca attiva ed effettuare il provisioning delle istanze, preferibilmente del tipo di istanza, di conseguenza. UltraWarm `uw.large` Ad esempio, se i clienti hanno 2 `uw.large` istanze configurate, ciascuna avrà a disposizione circa `knn.memory.circuit_breaker.limit * 61` GiB di memoria off-heap. Otterrete prestazioni ottimali se tutte le vostre «warm query» sono indirizzate a shard la cui dimensione cumulativa del grafico non supera la memoria off-heap disponibile. La latenza ne risente se la memoria disponibile è inferiore a quella necessaria per caricare il grafico a causa degli sfratti e dell'attesa che la memoria off-heap diventi disponibile. Ecco perché non consigliamo di utilizzare `uw.medium` istanze per casi d'uso in cui vengono utilizzati motori in memoria o per casi di throughput di ricerca più elevato, indipendentemente dai motori.
- Gli indici KNN verso cui verrà effettuata la migrazione non verranno uniti forzatamente in un singolo segmento UltraWarm . In questo modo si evita qualsiasi impatto sui nodi caldi e caldi che presentano problemi di OOM a causa della dimensione del grafico che diventa troppo grande per i motori in memoria. A causa dell'aumento del numero di segmenti per shard, ciò potrebbe comportare un maggiore consumo di spazio nella cache locale e consentire a un minor numero di indici di migrare al livello caldo. È possibile scegliere di unire forzatamente gli indici in un singolo segmento utilizzando l'impostazione esistente e sovrascrivendola prima di migrare gli indici al livello caldo. Per ulteriori informazioni, consulta the section called “Regolazione della migrazione”.

- Se hai un caso d'uso in cui gli indici vengono ricercati raramente e non soddisfano un carico di lavoro sensibile alla latenza, puoi scegliere di migrare tali indici al livello. UltraWarm Questo ti aiuterà a ridimensionare le istanze di calcolo di livello più elevato e a consentire a quest'ultimo di gestire la UltraWarm query su indici a priorità così bassa. Ciò può anche aiutare a garantire l'isolamento delle risorse consumate tra le query degli indici a bassa e alta priorità in modo che non si ripercuotano l'una sull'altra.

Disabilitazione UltraWarm

La console è il modo più semplice per UltraWarm disabilitare. Scegli il dominio, Operazioni quindi Modifica configurazione cluster. Deseleziona Abilita nodi dati caldi e scegli Salva modifiche. È inoltre possibile utilizzare l'opzione `WarmEnabled` nell'API di configurazione e in AWS CLI .

Prima di disabilitarli UltraWarm, devi eliminare tutti gli indici caldi o migrarli nuovamente all'hot storage. Una volta esaurita la memoria a caldo, attendi cinque minuti prima di provare a disattivarla. UltraWarm

Conservazione a freddo per Amazon OpenSearch Service

Lo storage a freddo ti consente di archiviare qualsiasi quantità di dati storici o ad accesso raro sul tuo dominio Amazon OpenSearch Service e di analizzarli su richiesta, a un costo inferiore rispetto ad altri livelli di storage. L'archiviazione a freddo è adatta se si ha bisogno di effettuare ricerche periodiche o analisi forensi sui dati più vecchi. Esempi pratici di dati adatti per l'archiviazione a freddo includono i log a cui si accede raramente, i dati che devono essere conservati per soddisfare i requisiti di conformità o i log che hanno un valore storico.

Analogamente allo UltraWarmstorage, lo storage a freddo è supportato da Amazon S3. Quando è necessario interrogare dati non aggiornati, è possibile collegarli in modo selettivo ai nodi esistenti UltraWarm. È possibile gestire la migrazione e il ciclo di vita dei dati a freddo manualmente o con le policy di Index State Management.

Argomenti

- Prerequisiti
- Requisiti di archiviazione a freddo e considerazioni sulle prestazioni
- Prezzi dell'archiviazione a freddo
- Abilitazione dell'archiviazione a freddo
- Gestione degli indici freddi nelle dashboard OpenSearch

- Migrazione degli indici all'archiviazione a freddo
- Automatizzazione delle migrazioni all'archiviazione a freddo
- Annullamento delle migrazioni all'archiviazione a freddo
- Elencare gli indici freddi
- Migrazione degli indici a freddo all'archiviazione a caldo
- Ripristino di indici a freddo da snapshot
- Annullamento delle migrazioni dall'archiviazione a freddo a quella a caldo
- Aggiornamento dei metadati dell'indice a freddo
- Eliminazione di indici freddi
- Disabilitazione dell'archiviazione a freddo

Prerequisiti

L'archiviazione a freddo ha i seguenti prerequisiti:

- La conservazione a freddo richiede Elasticsearch versione 7.9 OpenSearch o successiva.
- Per abilitare la conservazione a freddo su un dominio OpenSearch di servizio, è necessario abilitare anche la conservazione a caldo sullo stesso dominio.
- Per utilizzare l'archiviazione a freddo, i domini devono disporre di nodi principali dedicati.
- Se il dominio utilizza un tipo di istanza T2 o T3 per i nodi di dati, non è possibile utilizzare l'archiviazione a freddo.
- Se l'indice utilizza un valore approssimativo di k-NN (`"index.knn":true`), puoi spostarlo nella conservazione a freddo dalla versione 2.17 e successive. I domini delle versioni precedenti alla 2.17 possono essere aggiornati alla 2.17 per utilizzare questa funzionalità, ma gli indici KNN creati nelle versioni precedenti alla 2.x non possono migrare a Cold.
- Se il dominio utilizza un controllo granulare degli accessi, gli utenti non amministratori devono essere mappati al ruolo nelle dashboard per gestire gli indici freddi. `cold_manager` OpenSearch

> **Note**
>
> Il `cold_manager` ruolo potrebbe non esistere in alcuni domini di servizio preesistenti. OpenSearch Se il ruolo non è visualizzato in Dashboards, sarà necessario crearlo manualmente.

Configurazione delle autorizzazioni

Se abiliti la conservazione a freddo su un dominio di OpenSearch servizio preesistente, il `cold_manager` ruolo potrebbe non essere definito nel dominio. Se il dominio utilizza un controllo granulare degli accessi, gli utenti non amministratori devono essere mappati a questo ruolo per gestire gli indici freddi. Per creare manualmente il ruolo `cold_manager`, procedere nel seguente modo:

1. In OpenSearch Dashboard, vai su Sicurezza e scegli Autorizzazioni.
2. Scegliere Crea gruppo di operazioni e configurare i seguenti gruppi:

Group name (Nome gruppo)	Permissions
`cold_cluster`	• `cluster:monitor/nodes/stats` • `cluster:admin/ultrawarm*` • `cluster:admin/cold/*`
`cold_index`	• `indices:monitor/stats` • `indices:data/read/minmax` • `indices:admin/ultrawarm/migration/get` • `indices:admin/ultrawarm/migration/cancel`

3. Scegliere Ruoli, quindi selezionare Crea ruolo.
4. Denominare il ruolo cold_manager.
5. Per Autorizzazioni cluster, scegliere il gruppo `cold_cluster` creato.
6. Per Indice, immettere `*`.
7. Per Autorizzazioni indice, scegliere il gruppo `cold_index` creato.
8. Scegli Create (Crea).
9. Dopo aver creato il ruolo, associalo a qualsiasi ruolo utente o di backend che gestisce gli indici freddi.

Requisiti di archiviazione a freddo e considerazioni sulle prestazioni

Poiché il cold storage utilizza Amazon S3, non comporta alcun sovraccarico dello storage a caldo, ad esempio repliche, spazio riservato Linux e spazio riservato Service. OpenSearch L'archiviazione a freddo non ha tipi di istanza specifici perché non ha alcuna capacità di calcolo collegata. Con l'archiviazione a freddo è possibile archiviare qualsiasi quantità di dati. Monitora la `ColdStorageSpaceUtilization` metrica in Amazon CloudWatch per vedere quanto spazio di archiviazione a freddo stai utilizzando.

Prezzi dell'archiviazione a freddo

Analogamente all' UltraWarm archiviazione, con la conservazione a freddo paghi solo per l'archiviazione dei dati. Non è previsto alcun costo di calcolo per i dati a freddo e non verrà addebitato nulla se nello spazio di archiviazione a freddo non ci sono dati.

Non si applicano spese di trasferimento quando si spostano i dati tra l'archiviazione a freddo e quella a caldo. Mentre gli indici vengono migrati tra l'archiviazione a caldo e quella a freddo, si continua a pagare solo una copia dell'indice. Al termine della migrazione, l'indice viene fatturato in base al livello di archiviazione in cui è stata eseguita la migrazione. Per ulteriori informazioni sui prezzi delle celle frigorifere, consulta i prezzi OpenSearch di Amazon Service.

Abilitazione dell'archiviazione a freddo

La console è il modo più semplice per creare un dominio che utilizza l'archiviazione a freddo. Durante la creazione del dominio, scegli innanzitutto Enable warm data nodes, perché devi abilitare lo storage a caldo sullo stesso dominio. Quindi, scegli Enable cold storage.

Lo stesso processo di base funziona sui domini esistenti, a condizione che soddisfino i prerequisiti. Anche dopo che lo stato del dominio cambia da Elaborazione a Attivo, l'archiviazione a freddo potrebbe non essere disponibile per diverse ore.

Per abilitare l'archiviazione a freddo è possibile utilizzare anche la AWS CLI o l'API di configurazione.

Comando CLI di esempio

Il AWS CLI comando seguente crea un dominio con tre nodi di dati, tre nodi master dedicati, storage a freddo abilitato e controllo granulare degli accessi abilitato:

```
aws opensearch create-domain \
```

```
  --domain-name my-domain \
  --engine-version Opensearch_1.0 \
  --cluster-
config ColdStorageOptions={Enabled=true},WarmEnabled=true,WarmCount=4,WarmType=ultrawarm1.mediu
 \
  --ebs-options EBSEnabled=true,VolumeType=gp2,VolumeSize=11 \
  --node-to-node-encryption-options Enabled=true \
  --encryption-at-rest-options Enabled=true \
  --domain-endpoint-options EnforceHTTPS=true,TLSSecurityPolicy=Policy-Min-
TLS-1-2-2019-07 \
  --advanced-security-options
 Enabled=true,InternalUserDatabaseEnabled=true,MasterUserOptions='{MasterUserName=master-
user,MasterUserPassword=master-password}' \
  --region us-east-2
```

Per ulteriori informazioni, consultare Riferimento ai comandi AWS CLI.

Richiesta dell'API di configurazione di esempio

La seguente richiesta all'API di configurazione crea un dominio con tre nodi di dati, tre nodi principali dedicati, l'archiviazione a freddo abilitata e il controllo granulare degli accessi abilitato:

```
POST https://es.us-east-2.amazonaws.com/2021-01-01/opensearch/domain
{
  "ClusterConfig": {
    "InstanceCount": 3,
    "InstanceType": "r6g.large.search",
    "DedicatedMasterEnabled": true,
    "DedicatedMasterType": "r6g.large.search",
    "DedicatedMasterCount": 3,
    "ZoneAwarenessEnabled": true,
    "ZoneAwarenessConfig": {
      "AvailabilityZoneCount": 3
     },
    "WarmEnabled": true,
    "WarmCount": 4,
    "WarmType": "ultrawarm1.medium.search",
    "ColdStorageOptions": {
       "Enabled": true
     }
  },
  "EBSOptions": {
    "EBSEnabled": true,
```

```
    "VolumeType": "gp2",
    "VolumeSize": 11
  },
  "EncryptionAtRestOptions": {
    "Enabled": true
  },
  "NodeToNodeEncryptionOptions": {
    "Enabled": true
  },
  "DomainEndpointOptions": {
    "EnforceHTTPS": true,
    "TLSSecurityPolicy": "Policy-Min-TLS-1-2-2019-07"
  },
   "AdvancedSecurityOptions": {
    "Enabled": true,
    "InternalUserDatabaseEnabled": true,
    "MasterUserOptions": {
      "MasterUserName": "master-user",
      "MasterUserPassword": "master-password"
    }
  },
  "EngineVersion": "Opensearch_1.0",
  "DomainName": "my-domain"
}
```

Per informazioni dettagliate, consulta l'Amazon OpenSearch Service API Reference.

Gestione degli indici freddi nelle dashboard OpenSearch

Puoi gestire gli indici caldi, caldi e freddi con l'interfaccia Dashboards esistente nel tuo dominio di servizio. OpenSearch Dashboards consente di eseguire la migrazione degli indici tra l'archiviazione a caldo e quella a freddo e di monitorare lo stato della migrazione degli indici, senza utilizzare la CLI o l'API di configurazione. Per ulteriori informazioni, consulta Gestione degli indici nei dashboard. OpenSearch

Migrazione degli indici all'archiviazione a freddo

Quando si esegue la migrazione degli indici all'archiviazione a freddo, si fornisce un intervallo di tempo per i dati per semplificare l'individuazione. È possibile selezionare un campo timestamp in base ai dati dell'indice, fornire manualmente un timestamp di inizio e di fine oppure scegliere di non specificarne uno.

Parametro	Valore supportato	Description
`timestamp_field`	Il date/time campo della mappatura dell'indice.	I valori minimo e massimo del campo forniti vengono calcolati e archiviati come metadati `start_time` e `end_time` per l'indice a freddo.
`start_time` e `end_time`	Uno dei seguenti formati: • strict_date_optional_time. Ad esempio: `yyyy-MM-dd'T'HH:mm:ss.SSSZ` o `yyyy-MM-dd` • L'ora Epoch espressa in millisecondi	I valori forniti vengono calcolati e archiviati come metadati `start_time` e `end_time` per l'indice a freddo.

Se non si intendi specificare un timestamp, aggiungere `?ignore=timestamp` alla richiesta.

La richiesta seguente esegue la migrazione di un indice a caldo all'archiviazione a freddo e fornisce ora di inizio e fine per i dati in tale indice:

```
POST _ultrawarm/migration/my-index/_cold
  {
    "start_time": "2020-03-09",
    "end_time": "2020-03-09T23:00:00Z"
  }
```

Quindi controlla lo stato della migrazione:

```
GET _ultrawarm/migration/my-index/_status

{
  "migration_status": {
    "index": "my-index",
    "state": "RUNNING_METADATA_RELOCATION",
    "migration_type": "WARM_TO_COLD"
```

```
  }
}
```

OpenSearch Il servizio migra un indice alla volta nella conservazione a freddo. È possibile disporre di un massimo di 100 migrazioni nella coda. Qualsiasi richiesta che supera il limite verrà rifiutata. Per controllare il numero corrente di migrazioni nella coda, monitorare il parametro`WarmToColdMigrationQueueSize`. Il processo di migrazione ha i seguenti stati:

```
ACCEPTED_COLD_MIGRATION - Migration request is accepted and queued.
RUNNING_METADATA_MIGRATION - The migration request was selected for execution and
 metadata is migrating to cold storage.
FAILED_METADATA_MIGRATION - The attempt to add index metadata has failed and all
 retries are exhausted.
PENDING_INDEX_DETACH - Index metadata migration to cold storage is completed. Preparing
 to detach the warm index state from the local cluster.
RUNNING_INDEX_DETACH - Local warm index state from the cluster is being removed. Upon
 success, the migration request will be completed.
FAILED_INDEX_DETACH - The index detach process failed and all retries are exhausted.
```

Automatizzazione delle migrazioni all'archiviazione a freddo

Si consiglia di utilizzare Index State Management per automatizzare il processo di migrazione dopo che un indice raggiunge una certa età o soddisfa altre condizioni. Guarda la policy di esempio, che dimostra come migrare automaticamente gli indici dalla conservazione a caldo a quella a freddo. UltraWarm

> **Note**
>
> Un `timestamp_field` esplicito è necessario per spostare gli indici nell'archiviazione a freddo utilizzando una policy di Index State Management.

Annullamento delle migrazioni all'archiviazione a freddo

Se una migrazione all'archiviazione a freddo è in coda o in uno stato di errore, è possibile annullare la migrazione utilizzando la seguente richiesta:

```
POST _ultrawarm/migration/_cancel/my-index
```

```
{
  "acknowledged" : true
}
```

Se il dominio utilizza il controllo granulare degli accessi, per effettuare questa richiesta è necessaria l'autorizzazione `indices:admin/ultrawarm/migration/cancel`.

Elencare gli indici freddi

Prima di eseguire l'interrogazione, puoi elencare gli indici in cold storage per decidere a quali migrare per ulteriori analisi. UltraWarm La richiesta seguente elenca tutti gli indici freddi, ordinati per nome dell'indice:

```
GET _cold/indices/_search
```

Risposta di esempio

```
{
  "pagination_id" : "je7MtGbClwBF/2Zp9Utk/h3yCo8nvbEXAMPLEKEY",
  "total_results" : 3,
  "indices" : [
    {
      "index" : "my-index-1",
      "index_cold_uuid" : "hjEoh26mRRCFxRIMdgvLmg",
      "size" : 10339,
      "creation_date" : "2021-06-28T20:23:31.206Z",
      "start_time" : "2020-03-09T00:00Z",
      "end_time" : "2020-03-09T23:00Z"
    },
    {
      "index" : "my-index-2",
      "index_cold_uuid" : "0vIS2n-oROmOWDFmwFIgdw",
      "size" : 6068,
      "creation_date" : "2021-07-15T19:41:18.046Z",
      "start_time" : "2020-03-09T00:00Z",
      "end_time" : "2020-03-09T23:00Z"
    },
    {
      "index" : "my-index-3",
      "index_cold_uuid" : "EaeXOBodTLiDYcivKsXVLQ",
      "size" : 32403,
      "creation_date" : "2021-07-08T00:12:01.523Z",
```

```
      "start_time" : "2020-03-09T00:00Z",
      "end_time" : "2020-03-09T23:00Z"
    }
  ]
}
```

Filtraggio

È possibile filtrare gli indici a freddo in base a un modello di indice basato su prefisso e a offset di intervallo di tempo.

La seguente richiesta elenca gli indici che corrispondono al modello di prefisso di `event-*`:

```
GET _cold/indices/_search
 {
   "filters":{
      "index_pattern": "event-*"
   }
 }
```

Risposta di esempio

```
{
  "pagination_id" : "je7MtGbClwBF/2Zp9Utk/h3yCo8nvbEXAMPLEKEY",
  "total_results" : 1,
  "indices" : [
    {
      "index" : "events-index",
      "index_cold_uuid" : "4eFiab7rRfSvp3slrIsIKA",
      "size" : 32263273,
      "creation_date" : "2021-08-18T18:25:31.845Z",
      "start_time" : "2020-03-09T00:00Z",
      "end_time" : "2020-03-09T23:00Z"
    }
  ]
}
```

La seguente richiesta restituisce indici con campi di metadati `start_time` e `end_time` tra `2019-03-01` e `2020-03-01`:

```
GET _cold/indices/_search
```

```
{
  "filters": {
    "time_range": {
      "start_time": "2019-03-01",
      "end_time": "2020-03-01"
    }
  }
}
```

Risposta di esempio

```
{
  "pagination_id" : "je7MtGbClwBF/2Zp9Utk/h3yCo8nvbEXAMPLEKEY",
  "total_results" : 1,
  "indices" : [
    {
      "index" : "my-index",
      "index_cold_uuid" : "4eFiab7rRfSvp3slrIsIKA",
      "size" : 32263273,
      "creation_date" : "2021-08-18T18:25:31.845Z",
      "start_time" : "2019-05-09T00:00Z",
      "end_time" : "2019-09-09T23:00Z"
    }
  ]
}
```

Ordinamento

È possibile ordinare gli indici a freddo in base ai campi di metadati, ad esempio il nome dell'indice o la dimensione. La richiesta seguente elenca tutti gli indici ordinati per dimensione in ordine decrescente:

```
GET _cold/indices/_search
 {
 "sort_key": "size:desc"
 }
```

Risposta di esempio

```
{
  "pagination_id" : "je7MtGbClwBF/2Zp9Utk/h3yCo8nvbEXAMPLEKEY",
  "total_results" : 5,
```

```
  "indices" : [
    {
      "index" : "my-index-6",
      "index_cold_uuid" : "4eFiab7rRfSvp3slrIsIKA",
      "size" : 32263273,
      "creation_date" : "2021-08-18T18:25:31.845Z",
      "start_time" : "2020-03-09T00:00Z",
      "end_time" : "2020-03-09T23:00Z"
    },
    {
      "index" : "my-index-9",
      "index_cold_uuid" : "mbD3ZRVDRI6ONqgEOsJyUA",
      "size" : 57922,
      "creation_date" : "2021-07-07T23:41:35.640Z",
      "start_time" : "2020-03-09T00:00Z",
      "end_time" : "2020-03-09T23:00Z"
    },
    {
      "index" : "my-index-5",
      "index_cold_uuid" : "EaeXOBodTLiDYcivKsXVLQ",
      "size" : 32403,
      "creation_date" : "2021-07-08T00:12:01.523Z",
      "start_time" : "2020-03-09T00:00Z",
      "end_time" : "2020-03-09T23:00Z"
    }
  ]
}
```

Altre chiavi di ordinamento valide sono `start_time:asc/desc`, `end_time:asc/desc` e `index_name:asc/desc`.

Paginazione

È possibile impaginare un elenco di indici freddi. Configurare il numero di indici da restituire per pagina con il parametro `page_size` (il valore di default è 10). Ogni richiesta `_search` sugli indici a freddo restituisce un `pagination_id` che è possibile utilizzare per le chiamate successive.

La seguente richiesta impagina i risultati di una richiesta `_search` degli indici a freddo e visualizza i successivi 100 risultati:

```
GET _cold/indices/_search?page_size=100
{
"pagination_id": "je7MtGbClwBF/2Zp9Utk/h3yCo8nvbEXAMPLEKEY"
```

```
}
```

Migrazione degli indici a freddo all'archiviazione a caldo

Dopo aver ristretto l'elenco degli indici freddi con i criteri di filtro indicati nella sezione precedente, trasferiscili nuovamente UltraWarm dove puoi interrogare i dati e utilizzarli per creare visualizzazioni.

La richiesta seguente esegue la migrazione di due indici a freddo all'archiviazione a caldo:

```
POST _cold/migration/_warm
 {
 "indices": "my-index1,my-index2"
 }


{
  "acknowledged" : true
}
```

Per verificare lo stato della migrazione e recuperare l'ID di migrazione, inviare la seguente richiesta:

```
GET _cold/migration/_status
```

Risposta di esempio

```
{
  "cold_to_warm_migration_status" : [
    {
      "migration_id" : "tyLjXCA-S76zPQbPVHkOKA",
      "indices" : [
        "my-index1,my-index2"
      ],
      "state" : "RUNNING_INDEX_CREATION"
    }
  ]
}
```

Per ottenere informazioni sulla migrazione specifiche dell'indice, includere il nome dell'indice:

```
GET _cold/migration/my-index/_status
```

Anziché specificare un indice, è possibile elencare gli indici in base al relativo stato di migrazione corrente. I valori validi sono `_failed`, `_accepted` e `_all`.

Il comando seguente ottiene lo stato di tutti gli indici in una singola richiesta di migrazione:

```
GET _cold/migration/_status?migration_id=my-migration-id
```

Recuperare l'ID di migrazione utilizzando la richiesta di stato. Per informazioni dettagliate sulla migrazione, aggiungere `&verbose=true`.

È possibile migrare gli indici dall'archiviazione a freddo a quella a caldo in batch da 10 o meno, con un massimo di 100 indici simultanei. Qualsiasi richiesta che supera il limite verrà rifiutata. Per controllare il numero di migrazioni in atto, monitora il parametro `ColdToWarmMigrationQueueSize`. Il processo di migrazione ha i seguenti stati:

```
ACCEPTED_MIGRATION_REQUEST - Migration request is accepted and queued.
RUNNING_INDEX_CREATION - Migration request is picked up for processing and will create
 warm indexes in the cluster.
PENDING_COLD_METADATA_CLEANUP - Warm index is created and the migration service will
 attempt to clean up cold metadata.
RUNNING_COLD_METADATA_CLEANUP - Cleaning up cold metadata from the indexes migrated to
 warm storage.
FAILED_COLD_METADATA_CLEANUP - Failed to clean up metadata in the cold tier.
FAILED_INDEX_CREATION - Failed to create an index in the warm tier.
```

Ripristino di indici a freddo da snapshot

Se devi ripristinare un indice freddo eliminato, puoi ripristinarlo al livello caldo seguendo le istruzioni riportate the section called “Ripristino degli indici caldi dalle istantanee” e poi migrando nuovamente l'indice al livello freddo. Non è possibile ripristinare un indice di freddo eliminato direttamente nel livello freddo. OpenSearch Il servizio conserva gli indici freddi per 14 giorni dopo la loro eliminazione.

Annullamento delle migrazioni dall'archiviazione a freddo a quella a caldo

Se una migrazione dell'indice dall'archiviazione a freddo a quella a caldo è in coda o in uno stato di errore, è possibile annullare la migrazione utilizzando la seguente richiesta:

```
POST _cold/migration/my-index/_cancel
```

```
{
  "acknowledged" : true
}
```

Per annullare la migrazione per un batch di indici (massimo 10 alla volta), specificare l'ID di migrazione:

```
POST _cold/migration/_cancel?migration_id=my-migration-id

{
  "acknowledged" : true
}
```

Recuperare l'ID di migrazione utilizzando la richiesta di stato.

Aggiornamento dei metadati dell'indice a freddo

È possibile aggiornare i campi `start_time` e `end_time` per un indice a freddo:

```
PATCH _cold/my-index
 {
 "start_time": "2020-01-01",
 "end_time": "2020-02-01"
 }
```

Non è possibile aggiornare il `timestamp_field` di un indice nell'archiviazione a freddo.

> **Note**
>
> OpenSearch Dashboards non supporta il metodo PATCH. Per aggiornare i metadati a freddo, utilizzare curl, Postman o qualche altro metodo.

Eliminazione di indici freddi

Se non si utilizza una policy ISM, è possibile eliminare manualmente gli indici a freddo. La seguente richiesta elimina un indice a freddo:

```
DELETE _cold/my-index
```

```
{
  "acknowledged" : true
}
```

Disabilitazione dell'archiviazione a freddo

La console OpenSearch di servizio è il modo più semplice per disabilitare la conservazione a freddo. Seleziona il dominio e scegli Operazioni, Modifica configurazione cluster, quindi deseleziona Abilita archiviazione a freddo.

Per utilizzare la AWS CLI o l'API di configurazione, sotto`ColdStorageOptions`, imposta. `"Enabled"="false"`

Prima di disabilitare l'archiviazione a freddo, è necessario eliminare tutti gli indici a freddo o migrarli nuovamente all'archiviazione a caldo, altrimenti l'operazione di disabilitazione non riuscirà.

Gestione dello stato dell'indice in Amazon OpenSearch Service

Index State Management (ISM) in Amazon OpenSearch Service ti consente di definire politiche di gestione personalizzate che automatizzano le attività di routine e le applicano a indici e modelli di indice. Non è più necessario configurare e gestire processi esterni per eseguire le operazioni di indice.

Una policy contiene uno stato predefinito e un elenco di stati fra cui l'indice può transitare. All'interno di ciascuno stato, puoi definire un elenco di operazioni da eseguire e di condizioni che attivano queste transizioni. Un tipico caso d'uso è quello di eliminare periodicamente i vecchi indici dopo un certo periodo di tempo. Ad esempio, puoi definire una policy che sposta l'indice nello stato `read_only` dopo 30 giorni e successivamente lo elimina dopo 90 giorni.

Dopo aver collegato una policy a un indice, ISM crea un processo che viene eseguito ogni 5-8 minuti (o 30-48 minuti per i cluster precedenti alla versione 1.3) per eseguire le operazioni di policy, controllare le condizioni e passare l'indice in stati diversi. Il tempo di base per l'esecuzione di questo processo è ogni 5 minuti, più viene aggiunto un jitter casuale tra lo 0 e il 60% per evitare che le attività arrivino da tutti gli indici contemporaneamente. ISM non esegue processi se lo stato del cluster è rosso.

ISM richiede Elasticsearch OpenSearch 6.8 o versione successiva.

> **Note**
>
> Questa documentazione fornisce una breve panoramica di ISM e diverse politiche di esempio. Spiega inoltre in che modo ISM per i domini Amazon OpenSearch Service differisce da ISM sui cluster OpenSearch autogestiti. Per la documentazione completa di ISM, incluso un riferimento completo ai parametri, le descrizioni di ogni impostazione e un riferimento all'API, consulta Index State Management nella documentazione. OpenSearch

> **Important**
>
> Non puoi più utilizzare modelli di indice per applicare policy ISM agli indici appena creati. È possibile continuare a gestire automaticamente gli indici appena creati con il campo del modello ISM. Questo aggiornamento introduce una modifica sostanziale che influisce sui CloudFormation modelli esistenti che utilizzano questa impostazione.

Creazione di una policy ISM

Per iniziare a utilizzare la gestione degli stati degli indici

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Seleziona il dominio per cui creare una policy ISM.
3. Dalla dashboard del dominio, vai all'URL delle OpenSearch dashboard e accedi con il nome utente e la password principali. L'URL segue il seguente formato:

   ```
   domain-endpoint/_dashboards/
   ```

4. Apri il pannello di navigazione a sinistra all'interno di OpenSearch Dashboards e scegli Gestione degli indici, quindi Crea politica.
5. Utilizza il plug-in editor visivo o editor JSON per creare policy. Consigliamo di utilizzare l'editor visivo in quanto offre un modo più strutturato per definire le policy. Per assistenza nella creazione di policy, consulta le policy di esempio qui sotto.
6. Dopo aver creato una policy, puoi collegarla a uno o più indici:

   ```
   POST _plugins/_ism/add/my-index
   {
   ```

```
  "policy_id": "my-policy-id"
}
```

> **Note**
>
> Se il tuo dominio sta eseguendo una versione legacy di Elasticsearch, usa `_opendistro`invece di `_plugins`.

In alternativa, seleziona l'indice in OpenSearch Dashboard e scegli Applica politica.

Policy di esempio

Le policy di esempio riportate di seguito illustrano come automatizzare i casi d'uso comuni di ISM.

Archiviazione ad accesso frequente, a caldo, a freddo

Questa policy di esempio sposta un indice dalla memorizzazione a caldo a UltraWarm, e infine, a conservazione a freddo. Quindi, elimina l'indice.

L'indice si trova inizialmente nello stato `hot`. Dopo dieci giorni, ISM lo sposta allo stato `warm`. 80 giorni dopo, quando l'indice ha 90 giorni, lo sposta nello stato `cold`. Dopo un anno, il servizio invia una notifica ad una stanza Amazon Chime indicante che l'indice sta per essere eliminato, quindi lo elimina definitivamente.

Tenere presente che gli indici a freddo richiedono l'operazione `cold_delete` piuttosto che la normale operazione `delete`. Inoltre, per gestire gli indici a freddo con ISM nei dati è necessario un `timestamp_field` esplicito.

```
{
  "policy": {
    "description": "Demonstrate a hot-warm-cold-delete workflow.",
    "default_state": "hot",
    "schema_version": 1,
    "states": [{
        "name": "hot",
        "actions": [],
        "transitions": [{
          "state_name": "warm",
          "conditions": {
```

```
            "min_index_age": "10d"
          }
        }]
      },
      {
        "name": "warm",
        "actions": [{
          "warm_migration": {},
          "retry": {
            "count": 5,
            "delay": "1h"
          }
        }],
        "transitions": [{
          "state_name": "cold",
          "conditions": {
            "min_index_age": "90d"
          }
        }]
      },
      {
        "name": "cold",
        "actions": [{
            "cold_migration": {
              "timestamp_field": "<your timestamp field>"
            }
          }
        ],
        "transitions": [{
          "state_name": "delete",
          "conditions": {
             "min_index_age": "365d"
          }
        }]
      },
      {
        "name": "delete",
        "actions": [{
          "notification": {
            "destination": {
              "chime": {
                "url": "<URL>"
              }
            },
```

```
            "message_template": {
              "source": "The index {{ctx.index}} is being deleted."
            }
          }
        },
        {
          "cold_delete": {}
        }]
      }
    ]
  }
}
```

Riduzione del numero di repliche

Questa seconda policy di esempio riduce il numero di repliche a zero dopo sette giorni per risparmiare spazio su disco e quindi elimina l'indice dopo 21 giorni. Questa policy presuppone che l'indice non sia critico e che non riceva più richieste di scrittura; la presenza di repliche zero comporta un rischio di perdita di dati.

```
{
  "policy": {
    "description": "Changes replica count and deletes.",
    "schema_version": 1,
    "default_state": "current",
    "states": [{
        "name": "current",
        "actions": [],
        "transitions": [{
          "state_name": "old",
          "conditions": {
            "min_index_age": "7d"
          }
        }]
      },
      {
        "name": "old",
        "actions": [{
          "replica_count": {
            "number_of_replicas": 0
          }
        }],
        "transitions": [{
```

```
          "state_name": "delete",
          "conditions": {
            "min_index_age": "21d"
          }
        }]
      },
      {
        "name": "delete",
        "actions": [{
          "delete": {}
        }],
        "transitions": []
      }
    ]
  }
}
```

Acquisizione di uno snapshot dell'indice

Questa policy di esempio utilizza l'operazione `snapshot` per acquisire uno snapshot di un indice non appena contiene almeno un documento. `repository` è il nome del repository di snapshot manuali registrato in Amazon S3. `snapshot` è il nome dello snapshot. Per i prerequisiti di snapshot e i passaggi per registrare un repository, consultare the section called “Creazione di snapshot di indici”.

```
{
  "policy": {
    "description": "Takes an index snapshot.",
    "schema_version": 1,
    "default_state": "empty",
    "states": [{
        "name": "empty",
        "actions": [],
        "transitions": [{
          "state_name": "occupied",
          "conditions": {
            "min_doc_count": 1
          }
        }]
      },
      {
        "name": "occupied",
        "actions": [{
          "snapshot": {
```

```
            "repository": "<my-repository>",
            "snapshot": "<my-snapshot>"
            }
          }],
          "transitions": []
      }
    ]
  }
}
```

Modelli ISM

È possibile configurare un campo `ism_template` in una policy in modo che quando si crea un indice corrispondente allo schema del modello, la policy viene automaticamente associata a tale indice. In questo esempio, qualsiasi indice creato con un nome che inizia con "log" viene automaticamente abbinata alla policy ISM `my-policy-id`:

```
PUT _plugins/_ism/policies/my-policy-id
{
  "policy": {
    "description": "Example policy.",
    "default_state": "...",
    "states": [...],
    "ism_template": {
      "index_patterns": ["log*"],
      "priority": 100
    }
  }
}
```

Per un esempio più dettagliato, consultare Policy di esempio con modello ISM per il rollover automatico.

Differenze

Rispetto a OpenSearch Elasticsearch, ISM per Amazon OpenSearch Service presenta diverse differenze.

Operazioni ISM

- OpenSearch Il servizio supporta tre operazioni ISM uniche, `warm_migration`cold_migration, e: `cold_delete`
  - Se il dominio è UltraWarmabilitato, l'`warm_migration`azione trasferisce l'indice alla memorizzazione a caldo.
  - Se il tuo dominio ha l'archiviazione a freddo abilitata, l'operazione `cold_migration` transita l'indice all'archiviazione a freddo e l'operazione `cold_delete` elimina l'indice dall'archiviazione a freddo.

  Anche se una di queste operazioni non viene completata all'interno del periodo di timeout impostato, la migrazione o la cancellazione degli indici continua. L'impostazione di error_notification per una delle operazioni precedenti ti informerà che l'operazione non è riuscita se non è stata completata entro il periodo di timeout, ma la notifica è solo per tuo riferimento. L'effettiva operazione non ha alcun timeout intrinseco e continua a essere eseguita fino a quando riesce o non riesce.
- Se il tuo dominio esegue OpenSearch Elasticsearch 7.4 o versione successiva, OpenSearch Service supporta ISM e operazioni. `open` `close`
- Se il tuo dominio esegue Elasticsearch 7.7 OpenSearch o versione successiva, OpenSearch Service supporta l'operazione ISM. `snapshot`

Operazioni ISM di archiviazione a freddo

Per gli indici freddi, è necessario specificare un `?type=_cold` parametro quando si utilizza il seguente ISM: APIs

- Aggiungi policy
- Rimuovi policy
- Aggiorna policy
- Riprova indice non riuscito
- Spiega indice

Questi indici APIs per gli indici freddi presentano le seguenti differenze aggiuntive:

- Gli operatori con caratteri jolly non sono supportati tranne quando li si utilizza alla fine. Ad esempio, `_plugins/_ism/<add, remove, change_policy, retry, explain>/logstash-*` è supportato ma `_plugins/_ism/<add, remove, change_policy, retry, explain>/iad-*-prod` non lo è.
- Non sono supportati più schemi e nomi di indici. Ad esempio, `_plugins/_ism/<add, remove, change_policy, retry, explain>/app-logs` è supportato ma `_plugins/_ism/<add, remove, change_policy, retry, explain>/app-logs,sample-data` non lo è.

Impostazioni ISM

OpenSearch ed Elasticsearch consentono di modificare tutte le impostazioni ISM disponibili utilizzando l'API. `_cluster/settings` Su Amazon OpenSearch Service, puoi modificare solo le seguenti impostazioni ISM:

- Impostazioni a livello di cluster:
  - `plugins.index_state_management.enabled`
  - `plugins.index_state_management.history.enabled`
- Impostazioni a livello di indice:
  - `plugins.index_state_management.rollover_alias`

Tutorial: Automazione dei processi di Index State Management

Questo tutorial dimostra come implementare una policy PPL che automatizza le attività di routine di gestione degli indici e applicarle agli indici e ai modelli di indice.

Index State Management (ISM) in Amazon OpenSearch Service ti consente di automatizzare le attività ricorrenti di gestione degli indici, in modo da evitare l'uso di strumenti aggiuntivi per gestire i cicli di vita degli indici. Puoi creare una policy che automatizzi queste operazioni in base all'età, alle dimensioni dell'indice e ad altre condizioni, il tutto dall'interno del tuo dominio Amazon OpenSearch Service.

OpenSearch Il servizio supporta tre livelli di archiviazione: lo stato «caldo» predefinito per la scrittura attiva e l'analisi a bassa latenza, UltraWarm per dati di sola lettura fino a tre petabyte e la conservazione a freddo per l'archiviazione illimitata a lungo termine.

Questo tutorial presenta un esempio di caso d'uso della gestione dei dati di serie temporali negli indici giornalieri. In questo tutorial, imposti un criterio che acquisisce un'istantanea automatica di ogni indice allegato dopo 24 ore. Quindi migra l'indice dallo stato caldo predefinito allo UltraWarm storage dopo due giorni, allo storage a freddo dopo 30 giorni e infine elimina l'indice dopo 60 giorni.

Prerequisiti

- Il dominio OpenSearch di servizio deve eseguire Elasticsearch versione 6.8 o successiva.
- Il dominio deve avere una conservazione a freddo UltraWarmabilitata.
- Devi registrare un repository di snapshot manuali per il tuo dominio.
- Il tuo ruolo utente richiede autorizzazioni sufficienti per accedere alla console OpenSearch di servizio. Se necessario, convalida e configura l'accesso ai domini.

Fase 1: Configura la policy PPL

Innanzitutto, configura una politica ISM nei OpenSearch dashboard.

1. Dalla dashboard del dominio nella console di OpenSearch servizio, accedi all'URL delle OpenSearch dashboard e accedi con il nome utente e la password principali. L'URL segue il seguente formato: *domain-endpoint*/_dashboards/.
2. In OpenSearch Dashboard, scegli Aggiungi dati di esempio e aggiungi uno o più indici di esempio al tuo dominio.
3. Aprire il pannello di navigazione sinistro e scegliere Gestione indici, quindi scegliere Crea policy.
4. Assegnare un nome alla policy `ism-policy-example`.
5. Sostituisci la policy predefinita con la seguente policy:

```
{
  "policy": {
    "description": "Move indexes between storage tiers",
    "default_state": "hot",
    "states": [
      {
        "name": "hot",
        "actions": [],
        "transitions": [
          {
            "state_name": "snapshot",
            "conditions": {
```

```
            "min_index_age": "24h"
          }
        }
      ]
    },
    {
      "name": "snapshot",
      "actions": [
        {
          "retry": {
            "count": 5,
            "backoff": "exponential",
            "delay": "30m"
          },
          "snapshot": {
            "repository": "snapshot-repo",
            "snapshot": "ism-snapshot"
          }
        }
      ],
      "transitions": [
        {
          "state_name": "warm",
          "conditions": {
            "min_index_age": "2d"
          }
        }
      ]
    },
    {
      "name": "warm",
      "actions": [
        {
          "retry": {
            "count": 5,
            "backoff": "exponential",
            "delay": "1h"
          },
          "warm_migration": {}
        }
      ],
      "transitions": [
        {
          "state_name": "cold",
```

```
            "conditions": {
              "min_index_age": "30d"
            }
          }
        ]
      },
      {
        "name": "cold",
        "actions": [
          {
            "retry": {
              "count": 5,
              "backoff": "exponential",
              "delay": "1h"
            },
            "cold_migration": {
              "start_time": null,
              "end_time": null,
              "timestamp_field": "@timestamp",
              "ignore": "none"
            }
          }
        ],
        "transitions": [
          {
            "state_name": "delete",
            "conditions": {
              "min_index_age": "60d"
            }
          }
        ]
      },
      {
        "name": "delete",
        "actions": [
          {
            "cold_delete": {}
          }
        ],
        "transitions": []
      }
    ],
    "ism_template": [
      {
```

```
        "index_patterns": [
          "index-*"
        ],
        "priority": 100
      }
    ]
  }
}
```

> **Note**
>
> Il campo `ism_template` automaticamente il criterio a qualsiasi indice appena creato che corrisponde a uno dei `index_patterns` specificati. In questo caso, tutti gli indici che iniziano per `index-`. È possibile modificare questo campo in modo che corrisponda a un formato di indice nel proprio ambiente. Per ulteriori informazioni sui modelli, consulta i modelli ISM.

6. Nella sezione `snapshot` della policy, sostituire *`snapshot-repo`* con il nome del repository degli snapshot che hai registrato per il tuo dominio. Facoltativamente, è anche possibile sostituire *`ism-snapshot`*, che sarà il nome dello snapshot al momento della creazione.
7. Scegli Create (Crea). La policy dell'indice è ora visibile nella pagina Policy di gestione dello stato.

Fase 2: collega la policy a uno o più indici.

Dopo aver creato la policy, puoi collegarla a uno o più indici nel cluster.

1. Passa alla scheda Indici ad accesso frequente e cerca `opensearch_dashboards_sample`, che elenca tutti gli indici di esempio aggiunti nella fase 1.
2. Seleziona tutti gli indici e scegli Applica politica, quindi scegli la ism-policy-examplepolitica che hai appena creato.
3. Scegli Applica.

È possibile monitorare gli indici mentre si spostano attraverso i vari stati della pagina degli Indici gestiti dalle policy.

Riepilogo degli indici in Amazon OpenSearch Service con indici cumulativi

Gli indici cumulativi in Amazon OpenSearch Service consentono di ridurre i costi di storage raggruppando periodicamente i vecchi dati in indici riepilogati.

È possibile selezionare i campi di interesse e utilizzare un rollup di indice per creare un nuovo indice con solo i campi aggregati in periodi fissi più grossolani. È possibile archiviare mesi o anni di dati cronologici a una frazione del costo con le stesse prestazioni di query.

L'indicizzazione cumulativa richiede Elasticsearch 7.9 OpenSearch o versione successiva.

> **Note**
>
> Questa documentazione ti aiuta a iniziare a creare un job di indicizzazione cumulativa in Amazon OpenSearch Service. Per una documentazione completa, incluso un elenco di tutte le impostazioni disponibili e un riferimento completo all'API, consulta Index rollups nella OpenSearch documentazione.

Creazione di un processo di rollup dell'indice

Per iniziare, scegli Index Management in OpenSearch Dashboards. Selezionare Processi di rollup e scegliere Creazione di un processo di rollup.

Passaggio 1: configura gli indici

Impostare gli indici di origine e di destinazione. L'indice di origine è quello per cui si desidera eseguire il rollup. L'indice di destinazione è dove vengono salvati i risultati di rollup dell'indice.

Dopo aver creato un processo di rollup dell'indice, non sarà possibile modificarne le selezioni.

Fase 2: Definizione di aggregazioni e parametri

Selezionare gli attributi con le aggregazioni (termini e istogrammi) e i parametri (avg, sum, max, min e conteggio valore) per cui si desidera eseguire il rollup. Assicurarsi di non aggiungere troppi attributi altamente granulari perché si consumerà spazio.

Fase 3: Specifica delle pianificazioni

Specificare una pianificazione per eseguire il rollup degli indici mentre vengono importati. Il processo di rollup dell'indice è abilitato per impostazione predefinita.

Fase 4: Revisione e creazione

Rivedere la configurazione e selezionare Crea.

Fase 5: Ricerca nell'indice di destinazione

È possibile utilizzare l'API `_search` standard per cercare l'indice di destinazione. Non è possibile accedere alla struttura interna dei dati nell'indice di destinazione perché il plug-in riscrive automaticamente la query in background per adattarla all'indice di destinazione. Questo per assicurarsi di poter utilizzare la stessa query per l'indice di origine e di destinazione.

Per eseguire una query sull'indice di destinazione, impostare `size` su 0:

```
GET target_index/_search
{
  "size": 0,
  "query": {
    "match_all": {}
  },
  "aggs": {
    "avg_cpu": {
      "avg": {
        "field": "cpu_usage"
      }
    }
  }
}
```

> **Note**
>
> OpenSearch le versioni 2.2 e successive supportano la ricerca di più indici di rollup in un'unica richiesta. OpenSearch le versioni precedenti alla 2.2 e le versioni precedenti di Elasticsearch OSS supportano solo un indice di rollup per ricerca.

Trasformazione degli indici in Amazon Service OpenSearch

Mentre i processi di aggregazione degli indici consentono di ridurre la granularità dei dati raggruppando i vecchi dati in indici condensati, i processi di trasformazione consentono di creare una visualizzazione riepilogativa diversa dei dati incentrata su determinati campi, in modo da poterli visualizzare o analizzare in diversi modi.

Le trasformazioni degli indici dispongono di un'interfaccia utente Dashboards e di un'API REST. OpenSearch La funzionalità richiede la OpenSearch versione 1.0 o successiva.

> **Note**
>
> Questa documentazione fornisce una breve panoramica delle trasformazioni degli indici per aiutarti a iniziare a utilizzarli su un dominio Amazon OpenSearch Service. Per una documentazione completa e un riferimento all'API REST, consulta Index transforms nella documentazione open source OpenSearch .

Creazione di un processo di trasformazione dell'indice

Se non disponi di dati nel cluster, utilizza i dati di volo di esempio all'interno di OpenSearch Dashboards per provare Transform Jobs. Dopo aver aggiunto i dati, avvia OpenSearch Dashboards. Quindi scegliere Gestione degli indici, Processo di trasformazione e Crea processo di trasformazione.

Passaggio 1: scegli gli indici

Nella sezione Indici, selezionare l'indice di origine e di destinazione. È possibile selezionare un indice di destinazione esistente o crearne uno nuovo specificando un nome.

Se desideri trasformare solo un sottoinsieme dell'indice di origine, scegli Aggiungi filtro dati e usa la OpenSearch query DSL per specificare un sottoinsieme dell'indice di origine.

Fase 2: Scelta dei campi

Dopo aver scelto gli indici, scegli i campi che desideri utilizzare nel processo di trasformazione e se utilizzare raggruppamenti o aggregazioni.

- È possibile utilizzare i raggruppamenti per inserire i dati in bucket separati nell'indice trasformato. Ad esempio, se desideri raggruppare tutte le destinazioni aeroportuali all'interno dei dati di volo di esempio, raggruppa il `DestAirportID` campo in un campo di `DestAirportID_terms` campo

di destinazione e potrai trovare l'aeroporto raggruppato IDs nell'indice trasformato al termine del processo di trasformazione.

- D'altra parte, le aggregazioni consentono di eseguire calcoli semplici. Ad esempio, è possibile includere un'aggregazione nel processo di trasformazione per definire un nuovo campo di `sum_of_total_ticket_price` che calcola la somma di tutti i biglietti aerei. Quindi è possibile analizzare i nuovi dati nell'indice trasformato.

Fase 3: Specifica di una pianificazione

I processi di trasformazione sono abilitati per impostazione predefinita e sono eseguiti in base a pianificazioni. Per trasforma intervallo di esecuzione, specificare un intervallo in minuti, ore o giorni.

Fase 4: Revisione e monitoraggio

Rivedere la configurazione e selezionare Crea. Quindi monitorare la colonna Stato del processo di trasformazione.

Fase 5: Ricerca nell'indice di destinazione

Una volta terminato il processo, è possibile utilizzare l'API `_search` standard per cercare l'indice di destinazione.

Ad esempio, dopo aver eseguito un processo di trasformazione che trasforma i dati di volo in base al campo `DestAirportID`, è possibile eseguire la seguente richiesta per restituire tutti i campi che hanno un valore `SFO`:

```
GET target_index/_search
{
  "query": {
    "match": {
      "DestAirportID_terms" : "SFO"
    }
  }
}
```

Replica tra cluster per Amazon Service OpenSearch

Con la replica tra cluster in Amazon OpenSearch Service, puoi replicare indici utente, mappature e metadati da un dominio di servizio a un altro. OpenSearch L'utilizzo della replica tra cluster aiuta a

garantire il ripristino di emergenza in caso di interruzione e consente di replicare i dati su data center geograficamente distanti per ridurre la latenza. Paghi le tariffe standard di trasferimento AWS dei dati per i dati trasferiti tra domini.

La replica tra cluster segue un modello di replica attiva-passiva in cui l'indice locale o dei follower estrae i dati dall'indice remoto o leader. L'indice leader si riferisce all'origine dei dati o all'indice da cui si desidera replicare i dati. L'indice dei follower si riferisce alla destinazione dei dati o all'indice in cui si desidera replicare i dati.

La replica tra cluster è disponibile nei domini che eseguono Elasticsearch 7.10 o 1.1 o versioni successive. OpenSearch

> **Note**
>
> Questa documentazione descrive come configurare la replica tra cluster dal punto di vista di Amazon OpenSearch Service. Ciò include l'utilizzo Console di gestione AWS di per configurare connessioni tra cluster, cosa non possibile in un cluster autogestito. OpenSearch Per la documentazione completa, incluso un riferimento alle impostazioni e un riferimento completo all'API, consulta Replica tra cluster nella documentazione. OpenSearch

Argomenti

- Limitazioni
- Prerequisiti
- Requisiti per le autorizzazioni
- Configurazione di una connessione tra cluster
- Avvio della replica
- Conferma della replica
- Sospensione e ripristino della replica
- Arresto della replica
- Auto-follow
- Aggiornamento dei domini connessi

Limitazioni

La replica tra cluster ha le seguenti limitazioni:

- Non puoi replicare i dati tra domini Amazon OpenSearch Service e cluster autogestiti OpenSearch o Elasticsearch.
- Non puoi replicare un indice da un dominio follower a un altro dominio di follower. Se desideri replicare un indice su più domini di follower, puoi replicarlo solo dal singolo dominio leader.
- Un dominio può essere connesso, tramite una combinazione di connessioni in entrata e in uscita, a un massimo di altri 20 domini.
- Quando si configura inizialmente una connessione tra cluster, il dominio leader deve trovarsi nella stessa versione o in una versione successiva rispetto al dominio follower.
- Non è possibile utilizzarlo CloudFormation per connettere domini.
- Non è possibile utilizzare la replica tra cluster su istanze M3 o istanze espandibili (T2 e T3).
- Non è possibile replicare i dati tra indici freddi UltraWarm o tra diversi indici. Entrambi gli indici devono essere nell'archiviazione ad accesso frequente.
- Quando elimini un indice dal dominio leader, l'indice corrispondente nel dominio del follower non viene eliminato automaticamente.
- La replica tra cluster non è supportata tra default e opt-in. Entrambi i domini devono trovarsi nelle regioni predefinite o nelle regioni opt-in.

Prerequisiti

Prima di configurare la replica tra cluster, assicurati che i domini soddisfino i seguenti requisiti:

- Elasticsearch 7.10 o 1.1 o versioni successive OpenSearch
- Controllo granulare degli accessi abilitato
- Node-to-node crittografia abilitata
- Gli indici Leader devono essere `index.soft_deletes.enabled` impostati su. `true` Gli indici creati in Elasticsearch 7.0 o OpenSearch 1.0 e versioni successive hanno questa impostazione abilitata per impostazione predefinita. Tuttavia, gli indici creati in Elasticsearch 6.x e successivamente aggiornati mantengono. `soft_deletes=false` Per replicare tali indici, è necessario prima reindicizzarli.

  Per verificare se in un indice sono abilitate le eliminazioni automatiche:

```
GET <index-name>/_settings?
include_defaults=true&flat_settings=true&filter_path=*.settings.index.soft_deletes.enabled
```

In caso `soft_deletes` `false` affermativo, reindicizza i dati in un nuovo indice prima di iniziare la replica.

Requisiti per le autorizzazioni

Per avviare la replica, è necessario includere l'autorizzazione `es:ESCrossClusterGet` sul dominio remoto (principale). Consigliamo la seguente policy IAM sul dominio remoto. Questa policy consente anche di eseguire altre operazioni, come l'indicizzazione dei documenti e l'esecuzione di ricerche standard:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "*"
        ]
      },
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/leader-domain/*"
    },
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
      "Action": "es:ESCrossClusterGet",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/leader-domain"
    }
  ]
}
```

Assicurarsi che l'autorizzazione `es:ESCrossClusterGet` sia applicata per `/leader-domain` e non per `/leader-domain/*`.

Affinché gli utenti non amministratori possano eseguire attività di replica, devono essere mappati alle autorizzazioni appropriate. La maggior parte delle autorizzazioni corrisponde a operazioni REST API specifiche. Ad esempio, l'autorizzazione `indices:admin/plugins/replication/index/_resume` consente di riprendere la replica di un indice. Per un elenco completo delle autorizzazioni, consulta Autorizzazioni di replica nella documentazione. OpenSearch

> **Note**
>
> I comandi per avviare la replica e creare una regola di replica sono casi speciali. Poiché richiamano processi in background sui domini leader e follower, è necessario inserire un e nella richiesta. `leader_cluster_role` `follower_cluster_role` OpenSearch Il servizio utilizza questi ruoli in tutte le attività di replica del backend. Per informazioni sulla mappatura e l'utilizzo di questi ruoli, consulta Mappare i ruoli del cluster leader e follower nella documentazione. OpenSearch

Configurazione di una connessione tra cluster

Per replicare gli indici da un dominio all'altro, è necessario configurare una connessione tra cluster tra i domini. Il modo più semplice per connettere i domini è dalla scheda Connessioni del pannello di controllo del dominio. È inoltre possibile utilizzare l'API di configurazione o AWS CLI. Poiché la replica tra cluster segue un modello "pull" (ovvero di estrazione), le connessioni vengono avviate dal dominio follower.

> **Note**
>
> Se in precedenza sono stati collegati due domini per eseguire ricerche tra cluster, non sarà possibile utilizzare la stessa connessione per la replica. La connessione è contrassegnata come `SEARCH_ONLY` nella console. Per eseguire la replica tra due domini precedentemente connessi, è necessario eliminare la connessione e ricrearla. Una volta fatto questo, la connessione è disponibile sia per la ricerca tra cluster che per la replica tra cluster.

Per impostare una connessione

1. Nella console di Amazon OpenSearch Service, seleziona il dominio del follower, vai alla scheda Connessioni e scegli Richiesta.
2. Per Alias di connessione, specifica un nome per la connessione.
3. Scegli se connetterti a un dominio nella tua regione Account AWS e in un altro account o regione.

   - Per connetterti a un dominio nella tua regione Account AWS e nella tua regione, seleziona il dominio e scegli Richiedi.
   - Per connetterti a un dominio in un'altra regione Account AWS o in un'altra regione, specifica l'ARN del dominio remoto e scegli Richiesta.

OpenSearch Il servizio convalida la richiesta di connessione. Se i domini non sono compatibili, la connessione ha esito negativo. Se la convalida ha esito positivo, viene inviata al dominio di destinazione per l'approvazione. Una volta che dominio di destinazione ha approvato la richiesta, sarà possibile iniziare la replica.

La replica tra cluster supporta la replica bidirezionale. Ciò significa che è possibile creare una connessione in uscita dal dominio A al dominio B e un'altra connessione in uscita dal dominio B al dominio A. È quindi possibile impostare la replica in modo che il dominio A segua un indice nel dominio B e il dominio B segua un indice nel dominio A.

Avvio della replica

Dopo aver stabilito una connessione tra cluster, potrai iniziare a replicare i dati. Innanzitutto, crea un indice da replicare nel dominio principale:

```
PUT leader-01
```

Per replicare tale indice, invia questo comando al dominio follower:

```
PUT _plugins/_replication/follower-01/_start
{
   "leader_alias": "connection-alias",
   "leader_index": "leader-01",
   "use_roles":{
      "leader_cluster_role": "all_access",
      "follower_cluster_role": "all_access"
```

```
    }
}
```

È possibile trovare l'alias di connessione nella scheda Connessioni nel pannello di controllo del dominio.

In questo esempio si presuppone che un amministratore stia emettendo la richiesta e utilizzi `all_access` per il `leader_cluster_role` e `follower_cluster_role` per semplicità. Negli ambienti di produzione, tuttavia, si consiglia di creare utenti di replica sugli indici leader e follower e di mapparli di conseguenza. I nomi utente devono essere identici. Per informazioni su questi ruoli e su come mapparli, consulta Mappare i ruoli del cluster leader e follower nella OpenSearch documentazione.

Conferma della replica

Per confermare che la replica è in corso, ottieni lo stato di replica:

```
GET _plugins/_replication/follower-01/_status

{
  "status" : "SYNCING",
  "reason" : "User initiated",
  "leader_alias" : "connection-alias",
  "leader_index" : "leader-01",
  "follower_index" : "follower-01",
  "syncing_details" : {
    "leader_checkpoint" : -5,
    "follower_checkpoint" : -5,
    "seq_no" : 0
  }
}
```

I valori dei checkpoint principale e follower iniziano come numeri interi negativi e riflettono il numero di partizioni presenti (-1 per una partizione, -5 per cinque partizioni e così via). I valori aumentano e diventano numeri interi positivi con ogni modifica apportata. Se i valori sono uguali, significa che gli indici sono completamente sincronizzati. È possibile utilizzare questi valori di checkpoint per misurare la latenza di replica nei domini.

Per convalidare ulteriormente la replica, aggiungi un documento all'indice principale:

```
PUT leader-01/_doc/1
```

```
{
   "Doctor Sleep":"Stephen King"
}
```

Quindi conferma che viene visualizzato sull'indice follower:

```
GET follower-01/_search

{
    ...
    "max_score" : 1.0,
    "hits" : [
      {
        "_index" : "follower-01",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 1.0,
        "_source" : {
          "Doctor Sleep" : "Stephen King"
        }
      }
    ]
  }
}
```

Sospensione e ripristino della replica

È possibile sospendere temporaneamente la replica se è necessario risolvere problemi o ridurre il carico sul dominio principale. Invia questa richiesta al dominio follower. Assicurati di includere un corpo della richiesta vuoto:

```
POST _plugins/_replication/follower-01/_pause
{}
```

Quindi ottieni lo stato per assicurarti che la replica sia sospesa:

```
GET _plugins/_replication/follower-01/_status

{
  "status" : "PAUSED",
  "reason" : "User initiated",
  "leader_alias" : "connection-alias",
```

```
  "leader_index" : "leader-01",
  "follower_index" : "follower-01"
}
```

Una volta completate le modifiche, riprendi la replica. Invia questa richiesta al dominio follower. Assicurati di includere un corpo della richiesta vuoto:

```
POST _plugins/_replication/follower-01/_resume
{}
```

Non è possibile riprendere la replica dopo che è stata sospesa per più di 12 ore. È necessario interrompere la replica, eliminare l'indice di follower e riavviare la replica del leader.

Arresto della replica

Quando si interrompe completamente la replica, l'indice follower smette di seguire l'indice principale e diventa un indice standard. Non è possibile riavviare una replica dopo averla arrestata.

Interrompi la replica dal dominio follower. Assicurati di includere un corpo della richiesta vuoto:

```
POST _plugins/_replication/follower-01/_stop
{}
```

Auto-follow

È possibile definire un set di regole di replica in base a un singolo dominio principale che replicano automaticamente gli indici che corrispondono a uno schema specificato. Quando un indice nel dominio leader corrisponde a uno dei modelli (ad esempio,`books*`), viene creato un indice di follower corrispondente nel dominio del follower. OpenSearch Il servizio replica tutti gli indici esistenti che corrispondono al modello, nonché i nuovi indici creati dall'utente. Non replica gli indici già presenti nel dominio follower.

Per replicare tutti gli indici (ad eccezione degli indici creati dal sistema e di quelli già esistenti nel dominio follower), utilizzare un modello con carattere jolly (`*`).

Creazione di una regola di replica

Crea una regola di replica nel dominio follower e specifica il nome della connessione tra cluster:

```
POST _plugins/_replication/_autofollow
```

```
{
   "leader_alias" : "connection-alias",
   "name": "rule-name",
   "pattern": "books*",
   "use_roles":{
      "leader_cluster_role": "all_access",
      "follower_cluster_role": "all_access"
   }
}
```

È possibile trovare l'alias di connessione nella scheda Connessioni nel pannello di controllo del dominio.

In questo esempio si presuppone che un amministratore stia emettendo la richiesta e usa `all_access` come ruoli per i domini principale e follower per semplicità. Negli ambienti di produzione, tuttavia, si consiglia di creare utenti di replica sugli indici leader e follower e di mapparli di conseguenza. I nomi utente devono essere identici. Per informazioni su questi ruoli e su come mapparli, consulta Mappare i ruoli del cluster leader e follower nella OpenSearch documentazione.

Per recuperare un elenco di regole di replica esistenti in un dominio, utilizza l'operazione API auto-follow.

Per testare la regola, crea un indice che corrisponda al pattern sul dominio principale:

```
PUT books-are-fun
```

Quindi controlla che la replica venga visualizzata sul dominio follower:

```
GET _cat/indices

health status index          uuid                   pri rep docs.count docs.deleted
 store.size pri.store.size
green  open   books-are-fun  ldfHO78xYYdxRMULuiTvSQ     1   1          0            0
     208b           208b
```

Eliminazione di una regola di replica

Quando si elimina una regola di replica, OpenSearch Service interrompe la replica di nuovi indici che corrispondono al modello, ma continua l'attività di replica esistente finché non si interrompe la replica di tali indici.

Elimina le regole di replica dal dominio follower:

```
DELETE _plugins/_replication/_autofollow
{
   "leader_alias" : "connection-alias",
   "name": "rule-name"
}
```

Aggiornamento dei domini connessi

Per aggiornare la versione del motore di due domini con una connessione tra cluster, aggiorna prima il dominio follower e poi il dominio leader. Non eliminate la connessione tra di essi, altrimenti la replica si interrompe e non sarà possibile riprenderla.

Migrazione degli indici di Amazon OpenSearch Service utilizzando la reindicizzazione remota

La reindicizzazione remota consente di copiare gli indici da un dominio Amazon OpenSearch Service a un altro. Puoi migrare gli indici da qualsiasi dominio di OpenSearch servizio o da cluster autogestiti ed Elasticsearch. OpenSearch

Un dominio e un indice remoti si riferiscono alla fonte dei dati o al dominio e all'indice da cui si desidera copiare i dati. Un dominio e un indice locali si riferiscono alla destinazione dei dati o al dominio e all'indice in cui si desidera copiare i dati.

La reindicizzazione remota richiede OpenSearch 1.0 o versione successiva oppure Elasticsearch 6.7 o versione successiva, sul dominio locale. Il dominio remoto deve essere precedente o avere la stessa versione principale del dominio locale. Le versioni di Elasticsearch sono considerate precedenti alle OpenSearch versioni, il che significa che puoi reindicizzare i dati dai domini Elasticsearch ai domini. OpenSearch All'interno della stessa versione principale, il dominio remoto può essere una qualsiasi versione secondaria. Ad esempio, è supportata la reindicizzazione remota da Elasticsearch 7.10.x a 7.9, ma non OpenSearch da 1.0 a Elasticsearch 7.10.x.

> **Note**
>
> Questa documentazione descrive come reindicizzare i dati tra domini Amazon OpenSearch Service. Per la documentazione completa sull'`reindex`operazione, inclusi i passaggi

dettagliati e le opzioni supportate, consulta il documento Reindex nella documentazione. OpenSearch

Argomenti

- Prerequisiti
- Reindicizza i dati tra i domini Internet del Servizio OpenSearch
- Reindicizza i dati tra i domini OpenSearch di servizio quando il telecomando si trova in un VPC
- Reindicizza i dati tra domini non OpenSearch di servizio
- Reindicizzazione di set di dati di grandi dimensioni
- Impostazioni di reindicizzazione remota

Prerequisiti

La reindicizzazione remota ha i seguenti requisiti:

- Il dominio remoto deve essere accessibile dal dominio locale. Per un dominio remoto che risiede all'interno di un VPC, il dominio locale deve avere accesso al VPC. Questo processo varia in base alla configurazione di rete, ma probabilmente implica la connessione a una rete VPN o gestita o l'utilizzo della connessione endpoint VPC nativa. Per ulteriori informazioni, consulta the section called “Supporto per VPC”.
- La richiesta deve essere autorizzata dal dominio remoto come qualsiasi altra richiesta REST. Se il dominio remoto ha abilitato il controllo granulare degli accessi, è necessario disporre dell'autorizzazione per eseguire la reindicizzazione sul dominio remoto e leggere l'indice sul dominio locale. Per ulteriori considerazioni sulla sicurezza, consultare the section called “Controllo granulare degli accessi”.
- Ti consigliamo di creare un indice con l'impostazione desiderata sul tuo dominio locale prima di iniziare il processo di reindicizzazione.
- Se il dominio utilizza un tipo di istanza T2 o T3 per i nodi di dati, non è possibile utilizzare la reindicizzazione remota.

Reindicizza i dati tra i domini Internet del Servizio OpenSearch

Lo scenario più semplice è che l'indice remoto si trovi nello stesso Regione AWS dominio locale con un endpoint accessibile pubblicamente e che tu abbia firmato le credenziali IAM.

Dal dominio remoto, specifica l'indice remoto da cui reindicizzare e l'indice locale da reindicizzare su:

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443"
    },
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

È necessario aggiungere 443 alla fine dell'endpoint del dominio remoto per un controllo di convalida.

Per verificare che l'indice venga copiato nel dominio locale, invia questa richiesta al dominio locale:

```
GET local_index/_search
```

Se l'indice remoto si trova in una regione diversa dal dominio locale, inserisci il nome della regione, ad esempio in questa richiesta di esempio:

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443",
      "region": "eu-west-1"
    },
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

Nel caso di regioni isolate come regioni AWS GovCloud (US) o regioni della Cina, l'endpoint potrebbe non essere accessibile perché il tuo utente IAM non è riconosciuto in quelle regioni.

Se il dominio remoto è protetto con l'autenticazione di base, specifica il nome utente e la password:

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443",
      "username": "username",
      "password": "password"
    },
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

Reindicizza i dati tra i domini OpenSearch di servizio quando il telecomando si trova in un VPC

Ogni dominio di OpenSearch servizio è costituito dalla propria infrastruttura interna di cloud privato virtuale (VPC). Quando crei un nuovo dominio in un OpenSearch Service VPC esistente, viene creata un'interfaccia di rete elastica per ogni nodo di dati nel VPC.

Poiché l'operazione di reindicizzazione remota viene eseguita dal dominio del OpenSearch servizio remoto e quindi all'interno del relativo VPC privato, è necessario un modo per accedere al VPC del dominio locale. È possibile farlo utilizzando la funzionalità di connessione endpoint VPC integrata per stabilire una connessione o configurando un proxy. AWS PrivateLink

Se il tuo dominio locale utilizza OpenSearch la versione 1.0 o successiva, puoi utilizzare la console o la AWS CLI per creare una connessione. AWS PrivateLink Una AWS PrivateLink connessione consente alle risorse del VPC locale di connettersi privatamente alle risorse nel VPC remoto all'interno dello stesso. Regione AWS

Per creare una connessione endpoint VPC, il dominio di origine da reindicizzare deve trovarsi in un VPC locale e sia il dominio di origine che quello di destinazione devono trovarsi nello stesso dominio. Regione AWS

Reindicizza i dati con Console di gestione AWS

Puoi utilizzare la reindicizzazione remota con la console per copiare gli indici tra due domini che condividono una connessione endpoint VPC.

1. Accedi alla console di Amazon OpenSearch Service all'indirizzohttps://console.aws.amazon.com/aos/.
2. Nel riquadro di navigazione a sinistra, scegli Domains (Domini).
3. Seleziona il dominio locale o il dominio in cui desideri copiare i dati. Si apre la pagina dei dettagli del dominio. Scegli la scheda Connessioni sotto le informazioni generali e scegli Richiedi.
4. Nella pagina Richiedi connessione, seleziona VPC Endpoint Connection per la tua modalità di connessione e inserisci altri dettagli pertinenti. Questi dettagli includono il dominio remoto, che è il dominio da cui vuoi copiare i dati. Quindi, scegli Request (Richiesta).
5. Vai alla pagina dei dettagli del dominio remoto, scegli la scheda Connessioni e trova la tabella Connessioni in entrata. Seleziona la casella di controllo accanto al nome del dominio da cui hai appena creato la connessione (il dominio locale). Scegli Approve (Approva).
6. Torna al dominio locale, scegli la scheda Connections (Connessioni) e individua la tabella Outbound connections (Connessioni in uscita). Dopo che la connessione tra i due domini è attiva, un endpoint diventa disponibile nella colonna Endpoint della tabella. Copia l'endpoint.
7. Apri il pannello di controllo per il dominio locale e scegli Dev Tools (Strumenti di sviluppo) nella barra di navigazione a sinistra. Per confermare che l'indice del dominio remoto non esiste ancora nel tuo dominio locale, esegui la seguente richiesta GET. Sostituiscilo *remote-domain-index-name* con il tuo nome di indice.

   ```
   GET remote-domain-index-name/_search
   {
      "query":{
         "match_all":{}
      }
   }
   ```

   Nell'output, dovresti vedere un errore che indica che l'indice non è stato trovato.
8. Sotto la tua richiesta GET, crea una richiesta POST e usa il tuo endpoint come host remoto, come riportato di seguito.

   ```
   POST _reindex
   {
      "source":{
         "remote":{
            "host":"connection-endpoint",
            "username":"username",
            "password":"password"
   ```

```
      },
      "index":"remote-domain-index-name"
   },
   "dest":{
      "index":"local-domain-index-name"
   }
}
```

Esegui questa richiesta.

9. Esegui nuovamente la richiesta GET. L'output dovrebbe ora indicare che l'indice locale esiste. Puoi interrogare questo indice per verificare che siano OpenSearch stati copiati tutti i dati dall'indice remoto.

Reindicizza i dati con le operazioni dell'API OpenSearch di servizio

Puoi utilizzare la reindicizzazione remota con l'API per copiare gli indici tra due domini che condividono una connessione endpoint VPC.

1. Utilizza l'operazione CreateOutboundConnectionAPI per richiedere una nuova connessione dal dominio locale al dominio remoto.

```
POST https://es.region.amazonaws.com/2021-01-01/opensearch/cc/outboundConnection

{
   "ConnectionAlias": "remote-reindex-example",
   "ConnectionMode": "VPC_ENDPOINT",
   "LocalDomainInfo": {
      "AWSDomainInformation": {
         "DomainName": "local-domain-name",
         "OwnerId": "aws-account-id",
         "Region": "region"
      }
   },
   "RemoteDomainInfo": {
      "AWSDomainInformation": {
         "DomainName": "remote-domain-name",
         "OwnerId": "aws-account-id",
         "Region": "region"
      }
   }
}
```

Riceverai un `ConnectionId` messaggio nella risposta. Salva questo ID per utilizzarlo nel passaggio successivo.

2. Utilizza l'operazione AcceptInboundConnectionAPI con il tuo ID di connessione per approvare la richiesta dal dominio locale.

```
PUT https://es.region.amazonaws.com/2021-01-01/opensearch/cc/
inboundConnection/ConnectionId/accept
```

3. Usa l'operazione DescribeOutboundConnectionsAPI per recuperare l'endpoint per il tuo dominio remoto.

```
{
    "Connections": [
        {
            "ConnectionAlias": "remote-reindex-example",
            "ConnectionId": "connection-id",
            "ConnectionMode": "VPC_ENDPOINT",
            "ConnectionProperties": {
                "Endpoint": "connection-endpoint"
            },
            ...
        }
    ]
}
```

Salva il file *connection-endpoint* per utilizzarlo nel passaggio 5.

4. Per confermare che l'indice del dominio remoto non esiste ancora nel dominio locale, esegui la seguente richiesta GET. Sostituiscilo *remote-domain-index-name* con il tuo nome di indice.

```
GET local-domain-endpoint/remote-domain-index-name/_search
{
   "query":{
      "match_all":{}
   }
}
```

Nell'output, dovresti vedere un errore che indica che l'indice non è stato trovato.

5. Crea una richiesta POST e usa il tuo endpoint come host remoto, come segue.

```
POST local-domain-endpoint/_reindex
{
   "source":{
      "remote":{
         "host":"connection-endpoint",
         "username":"username",
         "password":"password"
      },
      "index":"remote-domain-index-name"
   },
   "dest":{
      "index":"local-domain-index-name"
   }
}
```

Esegui questa richiesta.

6. Esegui nuovamente la richiesta GET. L'output dovrebbe ora indicare che l'indice locale esiste. È possibile interrogare questo indice per verificare che siano OpenSearch stati copiati tutti i dati dall'indice remoto.

Se il dominio remoto è ospitato all'interno di un VPC e non desideri utilizzare la funzionalità di connessione agli endpoint VPC, devi configurare un proxy con un endpoint accessibile pubblicamente. In questo caso, OpenSearch Service richiede un endpoint pubblico perché non è in grado di inviare traffico al tuo VPC.

Quando esegui un dominio in modalità VPC, uno o più endpoint vengono inseriti nel tuo VPC. Tuttavia, questi endpoint servono solo per il traffico che entra nel dominio all'interno del VPC e non consentono il traffico verso il VPC stesso.

Il comando remote reindex viene eseguito dal dominio locale, quindi il traffico di origine non è in grado di utilizzare tali endpoint per accedere al dominio remoto. Ecco perché in questo caso d'uso è necessario un proxy. Il dominio proxy deve avere un certificato firmato di una certification authority (CA) pubblica. I certificati autofirmati o privati firmati dalla CA non sono supportati.

Reindicizza i dati tra domini non OpenSearch di servizio

Se l'indice remoto è ospitato all'esterno del OpenSearch servizio, ad esempio in un'istanza EC2 autogestita, imposta il parametro su: `external true`

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443",
      "username": "username",
      "password": "password",
      "external": true
    },
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

In questo caso, è supportata solo l'autenticazione di base con nome utente e password. Il dominio remoto deve avere un endpoint accessibile pubblicamente (anche se si trova nello stesso VPC del dominio di servizio OpenSearch locale) e un certificato firmato da una CA pubblica. I certificati autofirmati o privati firmati dalla CA non sono supportati.

Reindicizzazione di set di dati di grandi dimensioni

La reindicizzazione remota invia una richiesta di scorrimento al dominio remoto con i seguenti valori predefiniti:

- Contesto di ricerca di 5 minuti
- Timeout socket di 30 secondi
- Dimensione del batch di 1.000

Consigliamo di regolare questi parametri per adattarli ai dati. Per documenti di grandi dimensioni, considerate un timeout più and/or lungo per un batch di dimensioni inferiori. Per ulteriori informazioni, consulta Impaginare i risultati.

```
POST _reindex?pretty=true&scroll=10h&wait_for_completion=false
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443",
```

```
      "socket_timeout": "60m"
    },
    "size": 100,
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

Ti consigliamo inoltre di aggiungere le seguenti impostazioni all'indice locale per migliorare le prestazioni:

```
PUT local_index
{
  "settings": {
    "refresh_interval": -1,
    "number_of_replicas": 0
  }
}
```

Al termine del processo di reindicizzazione, è possibile impostare il numero di repliche desiderato e rimuovere l'impostazione dell'intervallo di aggiornamento.

Per reindicizzare solo un sottoinsieme di documenti selezionati tramite una query, invia questa richiesta al dominio locale:

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443"
    },
    "index": "remote_index",
    "query": {
      "match": {
        "field_name": "text"
      }
    }
  },
  "dest": {
    "index": "local_index"
```

```
  }
}
```

La reindicizzazione remota non supporta il partizionamento, quindi non è possibile eseguire più operazioni di scorrimento per la stessa richiesta in parallelo.

Impostazioni di reindicizzazione remota

Oltre alle opzioni di reindicizzazione standard, OpenSearch Service supporta le seguenti opzioni:

Opzioni	Valori validi	Description	Richiesto
external	Booleano	Se il dominio remoto non è un dominio OpenSearch di servizio o se stai reindicizzando tra due domini VPC, specifica come. `true`	No
region	Stringa	Se il dominio remoto si trova in una regione diversa, specifica il nome della regione.	No

Gestione dei dati di serie temporali in Amazon OpenSearch Service con flussi di dati

Un flusso di lavoro tipico per la gestione dei dati delle serie temporali comporta più passaggi, ad esempio la creazione di un alias di indice di rollover, la definizione di un indice di scrittura e la definizione di mappatura e impostazioni comuni per gli indici di supporto.

I flussi di dati in Amazon OpenSearch Service aiutano a semplificare questo processo di configurazione iniziale. I flussi di dati funzionano immediatamente per i dati basati sul tempo, ad esempio i log delle applicazioni che in genere sono di natura aggiuntiva.

I flussi di dati richiedono la OpenSearch versione 1.0 o successiva.

> **Note**
>
> Questa documentazione fornisce i passaggi di base per aiutarti a iniziare a utilizzare i flussi di dati su un dominio Amazon OpenSearch Service. Per una documentazione completa, consulta Data Streams nella OpenSearch documentazione.

Nozioni di base sui flussi di dati

Un flusso di dati è composto internamente da più indici di supporto. Le richieste di ricerca vengono instradate a tutti gli indici di supporto, mentre le richieste di indicizzazione vengono instradate all'indice di scrittura più recente.

Fase 1: Creazione di un modello di indice

Per creare un flusso dei dati, è necessario innanzitutto creare un modello di indice che configura un set di indici come flusso dei dati. L'oggetto `data_stream` indica che si tratta di un flusso di dati e non di un modello di indice normale. Il modello di indice corrisponde al nome del flusso di dati:

```
PUT _index_template/logs-template
{
  "index_patterns": [
    "my-data-stream",
    "logs-*"
  ],
  "data_stream": {},
  "priority": 100
}
```

In questo caso, ogni documento inserito deve avere un campo `@timestamp`. Puoi inoltre definire il tuo campo timestamp personalizzato come proprietà nell'oggetto `data_stream`:

```
PUT _index_template/logs-template
{
  "index_patterns": "my-data-stream",
  "data_stream": {
    "timestamp_field": {
      "name": "request_time"
    }
  }
```

```
}
```

Fase 2: Creazione di un flusso di dati

Dopo aver creato un modello di indice, è possibile avviare direttamente l'importazione di dati senza dover creare un flusso di dati.

Poiché abbiamo un modello di indice corrispondente a un `data_stream` oggetto, crea OpenSearch automaticamente il flusso di dati:

```
POST logs-staging/_doc
{
  "message": "login attempt failed",
  "@timestamp": "2013-03-01T00:00:00"
}
```

Fase 3: Importazione di dati nel flusso di dati

Per inserire dati in un flusso di dati, puoi utilizzare la normale indicizzazione APIs. Assicurarsi che ogni documento indicizzato abbia un campo di timestamp. Se si prova a inserire un documento che non dispone di un campo timestamp, viene visualizzato un errore.

```
POST logs-redis/_doc
{
  "message": "login attempt",
  "@timestamp": "2013-03-01T00:00:00"
}
```

Fase 4: Ricerca in un flusso di dati

È possibile eseguire una ricerca in un flusso di dati proprio come in un indice normale o in un alias di indice. L'operazione di ricerca si applica a tutti gli indici di supporto (tutti i dati presenti nel flusso).

```
GET logs-redis/_search
{
  "query": {
    "match": {
      "message": "login"
    }
  }
```

```
}
```

Fase 5: Rollover di un flusso di dati

È possibile configurare una policy Index State Management (ISM) per automatizzare il processo di rollover per il flusso di dati. La policy ISM viene applicata agli indici di supporto al momento della loro creazione. Quando si associa una policy a un flusso dei dati, questa influisce solo sugli indici di supporto futuri di tale flusso dei dati. Inoltre, non è necessario fornire l'impostazione `rollover_alias`, perché la policy ISM deduce queste informazioni dall'indice di supporto.

> **Note**
>
> Se esegui la migrazione di un indice di supporto alla conservazione a freddo, OpenSearch rimuove questo indice dal flusso di dati. Anche se si riporta l'indice in UltraWarm, l'indice rimane indipendente e non fa parte del flusso di dati originale. Dopo che un indice è stato rimosso dal flusso di dati, la ricerca all'interno del flusso non restituirà alcun dato dall'indice.

> **Warning**
>
> L'indice di scrittura per un flusso di dati non può essere migrato alla conservazione a freddo. Se desideri migrare i dati dal tuo flusso di dati alla cold storage, devi eseguire il rollover del flusso di dati prima della migrazione.

Fase 6: Gestisci i flussi di dati nelle dashboard OpenSearch

Per gestire i flussi di dati dai OpenSearch dashboard, apri OpenSearch Dashboard, scegli Gestione degli indici, seleziona Indici o Indici gestiti da policy.

Fase 7: Eliminazione di un flusso di dati

L'operazione di eliminazione elimina prima gli indici di supporto di un flusso dei dati e quindi elimina il flusso dei dati stesso.

Per eliminare un flusso di dati e tutti i relativi indici di supporto nascosti:

```
DELETE _data_stream/name_of_data_stream
```

Monitoraggio dei dati in Amazon OpenSearch Service

Monitora in modo proattivo i tuoi dati in Amazon OpenSearch Service con avvisi e rilevamento delle anomalie. Impostare gli avvisi per ricevere notifiche quando i dati superano determinate soglie. Il rilevamento delle anomalie utilizza il machine learning per rilevare automaticamente eventuali outlier nello streaming dei dati. È possibile associare il rilevamento delle anomalie con l'avviso per essere certi di ricevere una notifica non appena viene rilevata un'anomalia.

Argomenti

- Configurazione degli avvisi in Amazon Service OpenSearch
- Rilevamento di anomalie in Amazon Service OpenSearch

Configurazione degli avvisi in Amazon Service OpenSearch

Configura gli avvisi in Amazon OpenSearch Service per ricevere notifiche quando i dati di uno o più indici soddisfano determinate condizioni. Ad esempio, è possibile decidere di ricevere un'e-mail se l'applicazione registra più di cinque errori HTTP 503 in un'ora oppure di contattare uno sviluppatore se negli ultimi venti minuti non sono stati indicizzati nuovi documenti.

L'invio di avvisi richiede Elasticsearch OpenSearch 6.2 o versione successiva.

> **Note**
>
> Questa documentazione fornisce una breve panoramica degli avvisi ed evidenzia in che modo gli avvisi su un dominio Amazon OpenSearch Service differiscono dagli avvisi su un cluster open source. OpenSearch Per la documentazione completa sugli avvisi, tra cui un riferimento completo alle API, un elenco di campi di richiesta disponibili per i monitor compositi e le descrizioni delle variabili di attivazione e azione disponibili, consulta la sezione Avvisi nella documentazione. OpenSearch

Argomenti

- Autorizzazioni per gli avvisi
- Nozioni di base sugli avvisi
- Notifications
- Differenze

Autorizzazioni per gli avvisi

Gli avvisi supportano il controllo granulare degli accessi. Per i dettagli sulla combinazione e l'abbinamento delle autorizzazioni in base al caso d'uso, consulta Alerting security nella documentazione. OpenSearch

Per accedere alla pagina Avvisi all'interno delle OpenSearch dashboard, devi almeno essere mappato al ruolo `alerting_read_access` predefinito o disporre di autorizzazioni equivalenti. Questo ruolo concede le autorizzazioni per visualizzare avvisi, destinazioni e monitor, ma non per confermare gli avvisi o modificare destinazioni o monitor.

Nozioni di base sugli avvisi

Per creare un avviso, si configura un monitor, che è un processo che viene eseguito secondo una pianificazione definita e interroga gli indici. OpenSearch Inoltre, è necessario configurare uno o più trigger che definiscono le condizioni in base alle quali vengono generati gli eventi. Infine, si configurano le azioni, ossia ciò che accade dopo l'attivazione di un avviso.

Nozioni di base sugli avvisi

1. Scegli Avvisi dal menu principale delle OpenSearch dashboard e scegli Crea monitor.
2. Crea un monitor per query, per bucket, per cluster o per documento. Per istruzioni, consulta la sezione Create a monitor (Creazione di un monitor).
3. Per Trigger, crea uno o più trigger. Per istruzioni, consulta la sezione Creazione di trigger.
4. Per Actions (Azioni), imposta un notification channel (canale di notifica) per l'avviso. Scegliere tra Slack, Amazon Chime, un webhook personalizzato o Amazon SNS. Come si può immaginare, le notifiche richiedono connettività al canale. Ad esempio, per notificare un canale Slack o inviare un webhook personalizzato a un server di terze parti, il dominio del OpenSearch servizio necessita di una connettività di rete appropriata. Il webhook personalizzato deve avere un indirizzo IP pubblico per consentire a un dominio di OpenSearch servizio di inviargli avvisi.

> **Tip**
>
> Dopo che un'operazione ha inviato correttamente un messaggio, è responsabilità dell'utente proteggere l'accesso a tale messaggio (ad esempio, l'accesso a un canale Slack). Se il dominio contiene dati sensibili, considerare l'utilizzo di trigger senza operazioni e verificare periodicamente la presenza di avvisi in Dashboards.

Notifications

Alerting si integra con Notifications, un sistema unificato per le notifiche. OpenSearch Notifications consente di configurare il servizio di comunicazione che si desidera utilizzare e di visualizzare le statistiche pertinenti oltre alle informazioni sulla risoluzione dei problemi. Per una documentazione completa, consulta Notifiche nella documentazione. OpenSearch

Il tuo dominio deve avere la OpenSearch versione 2.3 o successiva per utilizzare le notifiche.

> **Note**
>
> OpenSearch le notifiche sono separate dalle notifiche OpenSearch di servizio, che forniscono dettagli sugli aggiornamenti del software di servizio, sui miglioramenti di Auto-Tune e altre importanti informazioni a livello di dominio. OpenSearch le notifiche sono specifiche del plug-in.

I canali di notifica hanno sostituito le destinazioni di avviso a partire OpenSearch dalla versione 2.0. Le destinazioni sono ufficialmente divenute obsolete e in futuro tutte le notifiche di avviso verranno gestite attraverso i canali.

Quando aggiorni i tuoi domini alla versione 2.3 o successiva (poiché il supporto del OpenSearch servizio per la versione 2.x inizia con la versione 2.3), le destinazioni esistenti vengono migrate automaticamente ai canali di notifica. Se una destinazione non riesce a migrare, il monitor continuerà a utilizzarla fino alla migrazione del monitor su un canale di notifica. Per ulteriori informazioni, consulta Domande sulle destinazioni nella documentazione. OpenSearch

Per iniziare a usare le notifiche, accedi alle OpenSearch dashboard e scegli Notifiche, Canali e Crea canale.

Amazon Simple Notification Service (Amazon SNS) è un tipo di canale supportato per le notifiche. Per autenticare gli utenti, è necessario fornire all'utente l'accesso completo ad Amazon SNS oppure lasciare che assuma un ruolo IAM munito delle autorizzazioni per accedere ad Amazon SNS. Per istruzioni, consulta la sezione Amazon SNS come tipo di canale.

Differenze

Rispetto alla versione open source di OpenSearch, gli avvisi in Amazon OpenSearch Service presentano alcune differenze notevoli.

Impostazioni degli avvisi

OpenSearch Il servizio consente di modificare le seguenti impostazioni di avviso:

- `plugins.scheduled_jobs.enabled`
- `plugins.alerting.alert_history_enabled`
- `plugins.alerting.alert_history_max_age`
- `plugins.alerting.alert_history_max_docs`
- `plugins.alerting.alert_history_retention_period`
- `plugins.alerting.alert_history_rollover_period`
- `plugins.alerting.filter_by_backend_roles`

Tutte le altre impostazioni utilizzano i valori di default che non è possibile modificare.

Per disattivare gli avvisi, inviare la seguente richiesta:

```
PUT _cluster/settings
{
  "persistent" : {
    "plugins.scheduled_jobs.enabled" : false
  }
}
```

La richiesta seguente configura gli avvisi per eliminare automaticamente gli indici della cronologia dopo sette giorni, anziché i 30 giorni predefiniti:

```
PUT _cluster/settings
{
  "persistent": {
    "plugins.alerting.alert_history_retention_period": "7d"
  }
}
```

Se hai creato dei monitor in precedenza e desideri interrompere la creazione di indici di avviso giornalieri, elimina tutti gli indici della cronologia degli avvisi:

```
DELETE .plugins-alerting-alert-history-*
```

Per ridurre il numero di frammenti per gli indici cronologici, crea un modello di indice. La richiesta seguente imposta gli indici della cronologia per gli avvisi a una sola partizione e una sola replica:

```
PUT _index_template/template-name
{
  "index_patterns": [".opendistro-alerting-alert-history-*"],
  "template": {
    "settings": {
      "number_of_shards": 1,
      "number_of_replicas": 1
    }
  }
}
```

A seconda della tolleranza per la perdita di dati, si potrebbe anche prendere in considerazione l'utilizzo di repliche zero. Per ulteriori informazioni sulla creazione e la gestione dei modelli di indice, consulta Modelli di indice nella documentazione. OpenSearch

Rilevamento di anomalie in Amazon Service OpenSearch

Il rilevamento delle anomalie in Amazon OpenSearch Service rileva automaticamente le anomalie nei OpenSearch dati quasi in tempo reale utilizzando l'algoritmo Random Cut Forest (RCF). L'RCF è un algoritmo di machine learning non supervisionato che modella un disegno del flusso di dati in ingresso. L'algoritmo calcola un `anomaly grade` e un valore `confidence score` per ogni punto di dati in ingresso. Il rilevamento delle anomalie utilizza questi valori per differenziare un'anomalia dalle normali variazioni dei dati.

Puoi associare il plug-in di rilevamento delle anomalie al plug-in Alerting per avvisarti non appena viene rilevata un'anomalia.

Il rilevamento delle anomalie è disponibile nei domini che eseguono qualsiasi OpenSearch versione di Elasticsearch 7.4 o versione successiva. Tutti i tipi di istanza, tranne `t2.micro` e `t2.small`, supportano il rilevamento delle anomalie.

> **Note**
>
> Questa documentazione fornisce una breve panoramica del rilevamento delle anomalie nel contesto di Amazon OpenSearch Service. Per una documentazione completa, che include passaggi dettagliati, un riferimento all'API, un riferimento a tutte le impostazioni disponibili e i

passaggi per creare visualizzazioni e dashboard, consulta il rilevamento delle anomalie nella documentazione open source. OpenSearch

Prerequisiti

Il rilevamento delle anomalie ha i seguenti requisiti preliminari:

- Il rilevamento delle anomalie richiede Elasticsearch 7.4 OpenSearch o versione successiva.
- Il rilevamento delle anomalie supporta il controllo granulare degli accessi solo sulle versioni 7.9 e successive di Elasticsearch e su tutte le versioni di. OpenSearch Prima di Elasticsearch 7.9, solo gli utenti amministratori potevano creare, visualizzare e gestire i rilevatori.
- Se il tuo dominio utilizza un controllo granulare degli accessi, gli utenti non amministratori devono essere mappati al `anomaly_read_access` ruolo nelle OpenSearch dashboard per visualizzare i rilevatori o per creare e gestire i rilevatori. `anomaly_full_access`

Nozioni di base sul rilevamento di anomalie

Per iniziare, scegli Anomaly Detection in Dashboards. OpenSearch

Fase 1: Creazione di un rilevatore

Un rivelatore è un'attività individuale di rilevamento delle anomalie. È possibile creare più rilevatori, e tutti possono essere eseguiti simultaneamente. Ogni rilevatore analizza i dati provenienti da origini diverse.

Fase 2: Aggiunta di caratteristiche al rilevatore

Una caratteristica è il campo nell'indice che viene controllato per la presenza di anomalie. Un rilevatore può rilevare anomalie in una o più caratteristiche. È necessario scegliere una delle seguenti aggregazioni per ogni funzionalità: `average()`, `sum()`, `count()`, `min()` o `max()`.

> **Note**
>
> Il metodo di `count()` aggregazione è disponibile solo in Elasticsearch OpenSearch 7.7 o versioni successive. Per Elasticsearch 7.4, utilizzare un'espressione personalizzata come la seguente:

```
{
  "aggregation_name": {
     "value_count": {
        "field": "field_name"
     }
  }
}
```

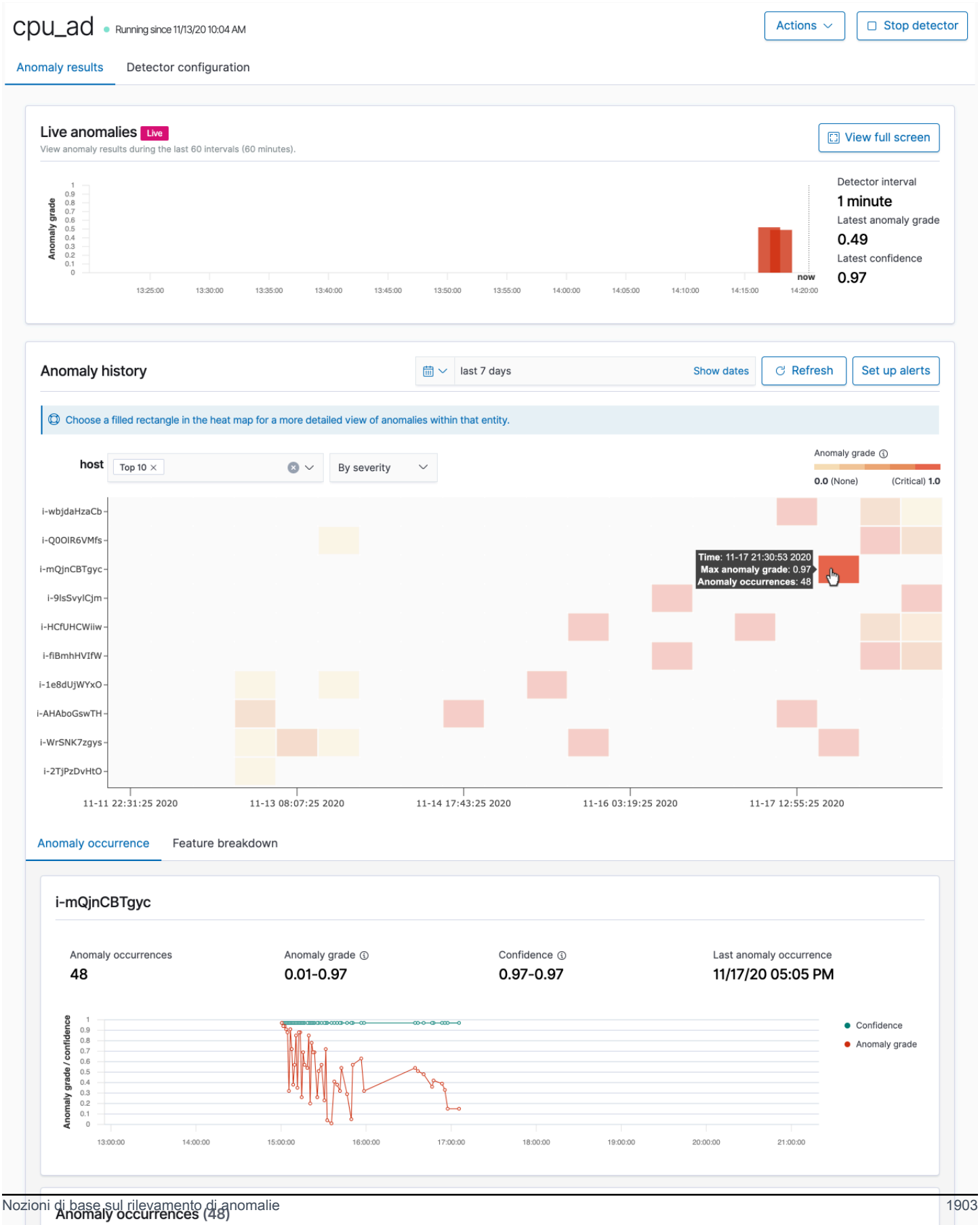
Il metodo di aggregazione determina ciò che costituisce un'anomalia. Ad esempio, se si sceglie `min()`, il rilevatore si concentra sulla ricerca di anomalie in base ai valori minimi della caratteristica. Se si sceglie `average()`, il rilevatore rileva anomalie in base ai valori medi della caratteristica. È possibile aggiungere un massimo di cinque caratteristiche per ogni rilevatore.

È possibile configurare le seguenti impostazioni facoltative (disponibili in Elasticsearch 7.7 e versioni successive):

- Campo Categoria: categorizzare o suddividere i dati con una dimensione come indirizzo IP, ID prodotto, codice paese e così via.
- Dimensione finestra: impostare il numero di intervalli di aggregazione dal flusso dei dati da considerare in una finestra di rilevamento.

Dopo aver configurato le funzioni, visualizzare in anteprima le anomalie dei campioni e, se necessario, regolare le impostazioni delle funzioni.

Fase 3: Osservazione dei risultati



- Anomalie in tempo reale visualizza i risultati delle anomalie in tempo reale per gli ultimi 60 intervalli. Ad esempio, se l'intervallo è impostato su 10, vengono visualizzati i risultati degli ultimi 600 minuti. Questo grafico viene aggiornato ogni 30 secondi.
- Cronologia delle anomalie traccia il grado di anomalia con il corrispondente livello di attendibilità.
- Il grafico Suddivisione delle caratteristiche traccia le caratteristiche in base al metodo di aggregazione. È possibile variare l'intervallo di data-ora del rilevatore.
- Occorrenze delle anomalie mostra `Start time`, `End time`, `Data confidence` e `Anomaly grade` per ogni anomalia rilevata.

  Se si imposta il campo della categoria, viene visualizzato una ulteriore mappa di calore che correla i risultati per le entità anomale. Scegliere un rettangolo pieno per visualizzare una vista più dettagliata dell'anomalia.

Fase 4: Configurazione degli avvisi

Per creare un monitoraggio per inviare notifiche quando vengono rilevate anomalie, scegliere Configura avvisi. Il plug-in reindirizza alla pagina Aggiungi monitor dove è possibile configurare un avviso.

Tutorial: Rileva un elevato utilizzo della CPU con il rilevamento delle anomalie

Questo tutorial dimostra come creare un rilevatore di anomalie in Amazon OpenSearch Service per rilevare un elevato utilizzo della CPU. Utilizzerai OpenSearch le dashboard per configurare un rilevatore per monitorare l'utilizzo della CPU e generare un avviso quando l'utilizzo della CPU supera una soglia specificata.

> **Note**
>
> Questi passaggi si applicano alla versione più recente di OpenSearch e potrebbero differire leggermente per le versioni precedenti.

Prerequisiti

- È necessario disporre di un dominio OpenSearch di servizio che esegue Elasticsearch 7.4 o versione successiva o qualsiasi versione. OpenSearch

- È necessario importare file di registro dell'applicazione nel cluster che contengono dati sull'utilizzo della CPU.

Fase 1: Creazione di un rilevatore

Innanzitutto, crea un rilevatore che identifichi le anomalie nei dati di utilizzo della CPU.

1. Apri il menu del pannello di sinistra in OpenSearch Dashboards e scegli Anomaly Detection, quindi scegli Crea rilevatore.
2. Assegna un nome al rilevatore **`high-cpu-usage`**.
3. Per l'origine dei dati, scegli il tuo indice che contiene i file di registro dell'utilizzo della CPU in cui desideri identificare le anomalie.
4. Seleziona Campo timestamp dai tuoi dati. Facoltativamente, puoi aggiungere un filtro dati. Questo filtro dati analizza solo un sottoinsieme dell'origine dati e riduce il rumore derivante da dati non rilevanti.
5. Impostazione della proprietà Intervallo del rilevatore a 2 minuti. Questo intervallo definisce il tempo (per intervallo di minuti) entro il quale il rilevatore raccoglie i dati.
6. Nello stato Ritardo di finestra, aggiungi un ritardo di 1 minuto. Questo ritardo aggiunge ulteriore tempo di elaborazione per garantire che tutti i dati all'interno della finestra siano presenti.
7. Scegli Next (Successivo). Nella dashboard di rilevamento delle anomalie, sotto il nome del rilevatore, selezionare Configura modello.
8. Per Nome caratteristica, digita **`max_cpu_usage`**. Per Stato della caratteristica, seleziona Attiva funzionalità.
9. Per Individuazione delle anomalie in base a , scegli Valore campo.
10. Per Metodo di aggregazione, scegli **`max()`**.
11. Per Campo, seleziona il campo nei dati per verificare la presenza di anomalie. Ad esempio, potrebbe essere chiamato `cpu_usage_percentage`.
12. Mantenere tutte le altre impostazioni come valori predefiniti e scegliere Successivo.
13. Ignorare la configurazione dei lavori del rilevatore e scegliere Successivo.
14. Nella finestra popup, scegliere quando avviare il rilevatore (automaticamente o manualmente), quindi scegliere Confermare.

Ora che il rilevatore è configurato, dopo l'inizializzazione, sarà possibile visualizzare i risultati in tempo reale dell'utilizzo della CPU nella sezione Risultati in tempo reale del rilevatore. La sezione

Anomalie dal vivo mostra tutte le anomalie che si verificano quando i dati vengono inseriti in tempo reale.

Fase 2: configura un avviso

Ora che hai creato un rilevatore, crea un monitor che richiami un avviso per inviare un messaggio a Slack quando rileva un utilizzo della CPU che soddisfa le condizioni specificate nelle impostazioni del rilevatore. Riceverai notifiche di Slack quando i dati di uno o più indici soddisfano le condizioni che richiamano l'avviso.

1. Apri il menu del pannello di sinistra in OpenSearch Dashboards e scegli Avvisi, quindi scegli Crea monitor.
2. Fornire un nome per il nome del monitor.
3. Per Tipo di monitoraggio, scegli Monitoraggio per query. Un monitor per query esegue una query specificata e definisce i trigger.
4. Per Metodo di definizione del monitor, scegli Rilevatore di anomalie, quindi selezionare il rilevatore creato nella sezione precedente dal menu a discesa Rilevatore.
5. Per Pianificazione, scegliere la frequenza con cui il monitor raccoglie i dati e la frequenza con cui si ricevono gli avvisi. Ai fini del presente tutorial, imposta la pianificazione per l'esecuzione di 7minuti.
6. Nella sezione Trigger, scegli Aggiungi trigger. Per Nome trigger, digita **`High CPU usage`**. Per questo tutorial, per Livello di gravità, scegli 1, che è il livello di gravità più alto.
7. Per Soglia di anomalie, scegli È SOPRA. Nel menu sottostante, scegli la soglia da applicare. Per questo tutorial, imposta il Grado di anomalie a 0.7.
8. Per Soglia di confidenza anomalie, scegli È SOPRA. Nel menu sottostante, inserisci lo stesso numero del tuo grado di anomalia. Per questo tutorial, imposta Soglia di confidenza anomalie a 0.7.
9. Nella sezione Operazioni, scegli Destinazione. Nel campo Nome, scegliere il nome della destinazione. Sul menu Tipo, scegli Margine di flessibilità. Nel campo Webhook URL, immettere un URL del webhook a cui ricevere gli avvisi. Per ulteriori informazioni, consulta Invio di messaggi tramite webhook in entrata.
10. Scegli Create (Crea).

Risorse correlate

- the section called “Avviso”

- the section called “Rilevamento anomalie”
- API di rilevamento delle anomalie

Assistente AI per Amazon OpenSearch Service

AI Assistant è un assistente generativo basato sull'intelligenza artificiale che aiuta gli sviluppatori e i professionisti IT a svolgere una serie di attività durante il ciclo di vita dello sviluppo del software. AI Assistant aiuta AWS i clienti a codificare, testare, implementare, risolvere i problemi e ottimizzare le applicazioni in esecuzione. AWS

L'integrazione di AI Assistant con Amazon OpenSearch Service offre le seguenti funzionalità generative:

- the section called “Genera visualizzazioni utilizzando il linguaggio naturale”
- the section called “Visualizza riepiloghi e approfondimenti sugli avvisi”
- the section called “Visualizza i riepiloghi dei risultati delle query generati da AI Assistant nella pagina Discover”
- the section called “Visualizza i rilevatori di anomalie consigliati”
- the section called “Accedi alla chat di AI Assistant per le domande OpenSearch di servizio”

Per accedere alle funzionalità di AI Assistant di OpenSearch Service pertinenti alla tua attività, cerca la seguente icona contestuale nelle pagine di visualizzazione Alerts, Discover e Create. Fai clic sull'icona per richiedere assistenza e utilizzare le funzionalità descritte in questa sezione.

Puoi anche chattare direttamente con AI Assistant facendo clic sull'icona nell'angolo in alto a destra della console di OpenSearch servizio. La chat di AI Assistant supporta risposte a domande sulle caratteristiche e funzionalità del OpenSearch Servizio.

Supportata Regioni AWS

Il supporto di AI Assistant per OpenSearch Service è disponibile nei seguenti paesi Regioni AWS:

- Stati Uniti orientali (Virginia settentrionale)
- Stati Uniti occidentali (Oregon)
- Asia Pacifico (Mumbai)
- Asia Pacifico (Sydney)
- Asia Pacifico (Tokyo)
- Canada (Centrale)

- Europa (Francoforte)
- Europa (Londra)
- Europa (Parigi)
- Sud America (San Paolo)

Configurazione di AI Assistant for OpenSearch Service

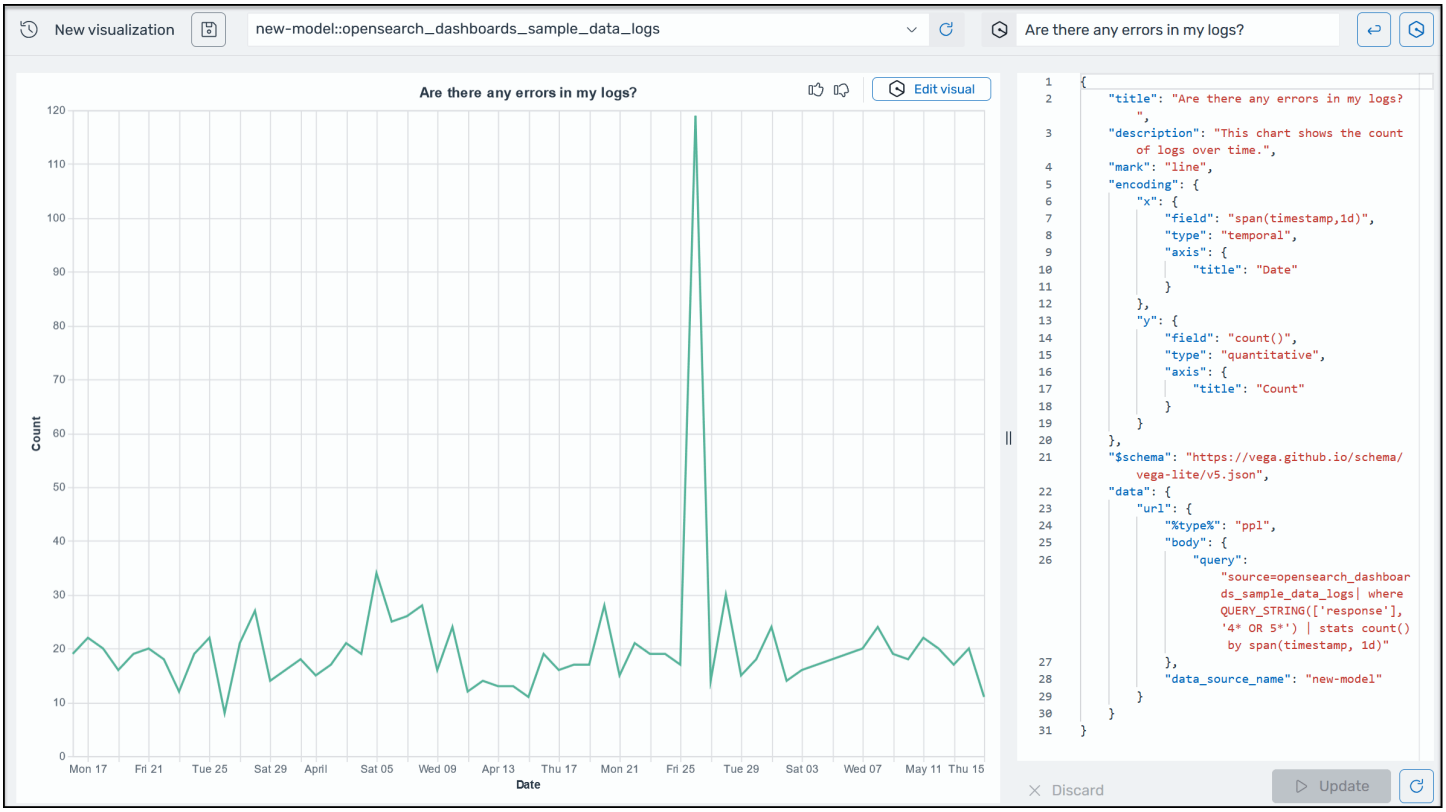
Completa i seguenti passaggi per configurare AI Assistant for OpenSearch Service.

1. Verifica di aver configurato il controllo granulare degli accessi di OpenSearch Service. Per ulteriori informazioni, consulta Controllo granulare degli accessi in Amazon Service OpenSearch.
2. Verifica che la tua origine dati sia nella OpenSearch versione 2.17 o successiva.
3. Verifica di aver selezionato la casella di controllo Abilita la generazione di query in linguaggio naturale nella sezione Artificial Intelligence (AI) e Machine Learning (ML) durante la creazione del dominio o modificando la configurazione del cluster.

> **Note**
>
> Le OpenSearch funzionalità di Amazon Q for Developer sono disponibili con il piano gratuito Q. Per ulteriori informazioni, consulta Comprendere i livelli di servizio nella Guida per l'utente di Amazon Q.

Genera visualizzazioni utilizzando il linguaggio naturale

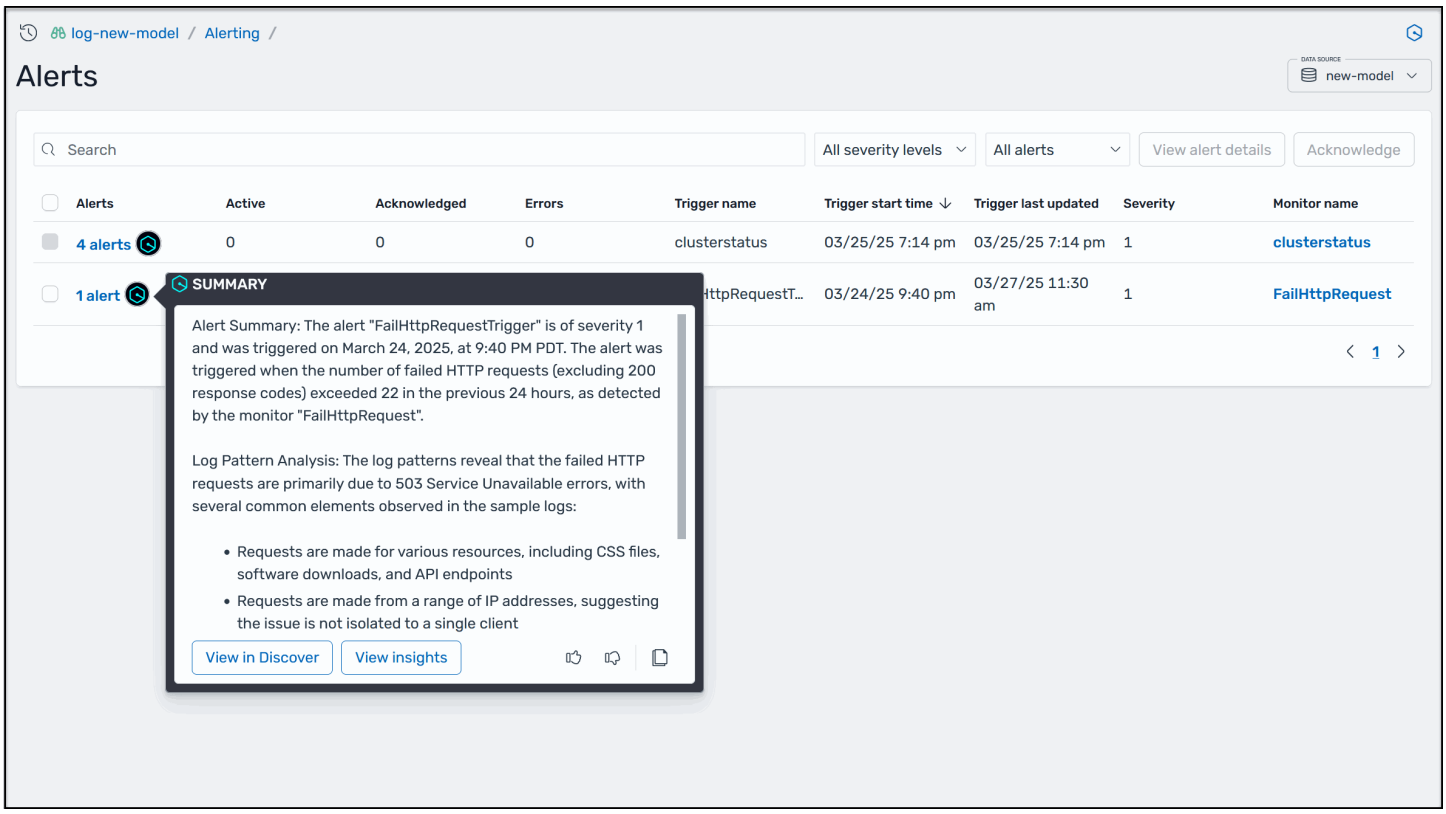
Per aiutarti a ottenere ulteriori informazioni sui tuoi dati operativi, AI Assistant for OpenSearch Service supporta l'utilizzo di istruzioni in linguaggio naturale per creare visualizzazioni. Prima di iniziare, verifica di aver configurato AI Assistant for Service. OpenSearch Una volta confermato, avvia l'applicazione OpenSearch Service UI, vai alla pagina Visualizzazioni, seleziona il pulsante + Crea visualizzazione e scegli il linguaggio naturale. Ora puoi specificare un'origine dati e inserire un requisito in linguaggio naturale per generare le visualizzazioni. Ecco un esempio.



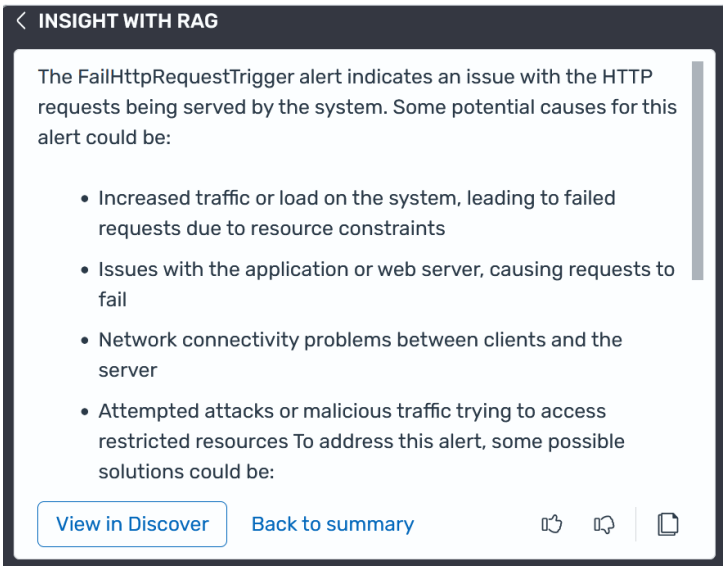
Le visualizzazioni possono accelerare la risoluzione dei problemi suddividendo l'analisi degli errori in base a diverse dimensioni. Le visualizzazioni possono anche aiutarti a identificare modelli e tendenze, accelerare il processo decisionale, rivelare relazioni e semplificare dati complessi. In alternativa, puoi anche creare visualizzazioni nella pagina Discover come parte dei risultati della query.

Visualizza riepiloghi e approfondimenti sugli avvisi

È possibile configurare OpenSearch Service per creare un monitoraggio degli avvisi quando i dati di uno o più indici soddisfano determinate condizioni. Per aiutarti a comprendere e risolvere rapidamente un avviso, puoi visualizzare un riepilogo degli avvisi facendo clic sull'icona dell'Assistente AI accanto a un avviso. Un riepilogo fornisce dettagli sul problema sottostante che ha attivato l'avviso e, se disponibile, un'analisi aggiuntiva per aiutarti a individuare la causa principale del problema. La schermata seguente mostra un esempio di riepilogo degli avvisi creato da AI Assistant.



Se colleghi una knowledge base per fornire un contesto aggiuntivo sull'ambiente, come descritto più avanti in questo argomento, AI Assistant crea informazioni dettagliate su un avviso. Insights fornisce dettagli e opzioni di risoluzione dei problemi per aiutarti a porre rimedio alla causa principale di un avviso. Per quanto riguarda l'avviso mostrato in precedenza, AI Assistant ha prodotto anche le seguenti informazioni.



> **Note**
>
> A seconda della natura dell'avviso e delle informazioni disponibili, AI Assistant può darti la possibilità di visualizzare i dati dell'avviso nella pagina Scopri delle OpenSearch dashboard. Se vedi il pulsante Visualizza in Discover nella parte inferiore del riepilogo di un avviso di AI Assistant, fai clic sul pulsante per aprire il set di dati corrispondente in Discover con un filtro attivo per i dati degli avvisi.

Argomenti

- Prima di iniziare
- Visualizzazione di riepiloghi e approfondimenti sugli avvisi

Prima di iniziare

Completa i seguenti passaggi per configurare una knowledge base di Amazon Bedrock in modo che AI Assistant possa creare approfondimenti per gli avvisi OpenSearch di servizio.

Fase 1: creazione del LambdaInvokeOpenSearch MLCommons ruolo IAM

Crea un nuovo ruolo denominato `LambdaInvokeOpenSearchMLCommonsRole` in AWS Identity and Access Management (IAM). OpenSearch Il servizio utilizza questo ruolo per creare un connettore AI OpenSearch che aiuta a produrre approfondimenti basati su articoli della knowledge base configurati. È necessario mappare questo ruolo al `ml_full_access` ruolo OpenSearch Servizio, come descritto nel passaggio 2.

Quando crei il nuovo ruolo, per il tipo di entità affidabile, scegli AWS account. Non è necessario specificare una politica di autorizzazione. Nella pagina Add permissions (Aggiungi autorizzazioni), scegli Next (Successivo). Per ulteriori informazioni sulla creazione di un nuovo ruolo, vedere Creazione di un ruolo per un AWS servizio (console).

Passaggio 2: mappare il LambdaInvokeOpenSearch MLCommons ruolo del ruolo al ruolo OpenSearch Service ml_full_access

Utilizzare la procedura seguente per mappare il `LambdaInvokeOpenSearchMLCommonsRole` ruolo al ruolo di servizio. OpenSearch `ml_full_access` Questa mappatura aiuta anche OpenSearch Service a creare il connettore AI.

Per mappare il ruolo IAM richiesto al ruolo OpenSearch Service ml_full_access

1. Apri la pagina di amministrazione dei dati OpenSearch di Service Dashboard.
2. In Accesso ai dati e utente, scegli Ruoli.
3. Usa la casella di ricerca per individuare il `ml_full_access` ruolo.
4. Nella pagina ml_full_access, scegli la scheda Utenti mappati.
5. Scegliete Utenti della mappa.
6. Nel campo Backend roles, incolla l'Amazon Resource Name (ARN) **`LambdaInvokeOpenSearchMLCommonsRole`** del ruolo, quindi scegli Map.

Passaggio 3: configura una knowledge base OpenSearch del servizio utilizzando CloudFormation

Utilizza la seguente procedura per configurare una knowledge base OpenSearch del servizio AWS CloudFormation in modo che AI Assistant possa generare approfondimenti.

Per configurare una knowledge base per approfondimenti

1. Accedi alla https://console.aws.amazon.com/aos/home page della console OpenSearch di Amazon Service in un formato supportato Regione AWS. Per ulteriori informazioni, consulta Supportata Regioni AWS.
2. Nel riquadro di navigazione, scegli Integrazioni.
3. Nella sezione Modelli di integrazione, scegli il modello Integrazione con la knowledge base tramite Amazon Bedrock. Se non vedi questo modello, verifica di trovarti in una regione supportata.
4. Nel riquadro Integrazione con la knowledge base tramite Amazon Bedrock, scegli Configura dominio, quindi scegli una delle opzioni disponibili. OpenSearch Il servizio apre il modello di CloudFormation stack con i campi obbligatori precompilati. Lo CloudFormation stack supporta l'integrazione per domini pubblici e VPC.
5. Seleziona Crea stack. Dopo aver CloudFormation creato le risorse, il servizio visualizza l'agente AgentIdAmazon Bedrock e ModelId. ConnectorId

Se applicabile, AI Assistant ora crea approfondimenti per gli avvisi OpenSearch di servizio.

Visualizzazione di riepiloghi e approfondimenti sugli avvisi

Utilizzare la seguente procedura per visualizzare i riepiloghi e gli approfondimenti degli avvisi in OpenSearch Service.

Visualizzazione dei riepiloghi e degli approfondimenti sugli avvisi

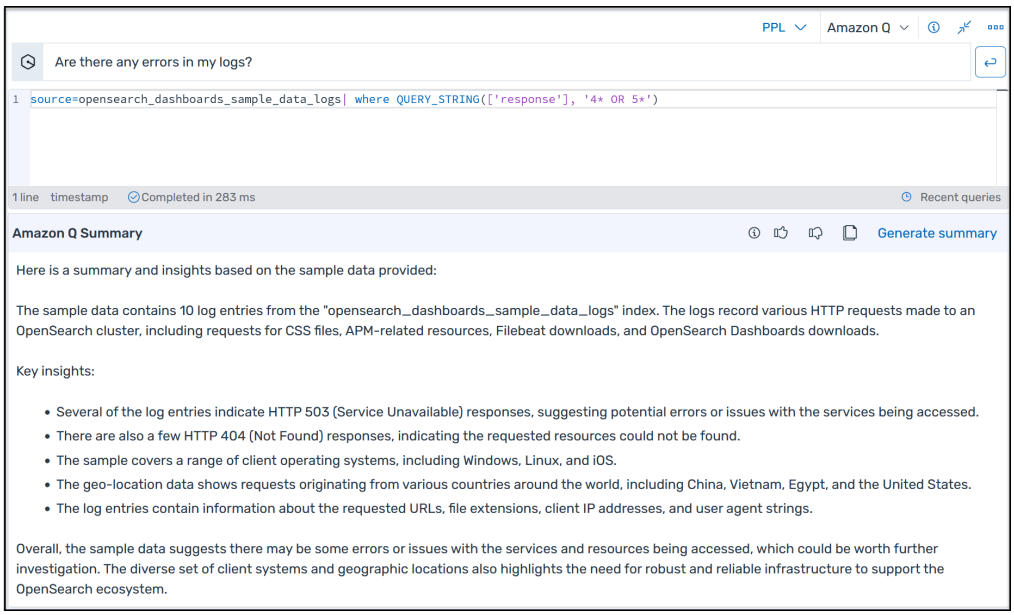
1. Verifica di aver configurato AI Assistant for OpenSearch Service.
2. Verifica di aver impostato gli avvisi per OpenSearch Service.
3. Nel menu principale delle OpenSearch dashboard, scegli Avvisi, quindi scegli Avvisi.
4. Scegli l'icona dell'Assistente AI accanto a un avviso. AI Assistant può impiegare fino a 10 secondi per generare il riepilogo.
5. Se presente nel riepilogo dell'avviso, scegli Visualizza approfondimenti per visualizzare maggiori dettagli sull'avviso in base alla knowledge base configurata.
6. Se presente nel riepilogo degli avvisi, scegli Visualizza in Discover per visualizzare i dati degli avvisi nella pagina Scopri delle OpenSearch dashboard.

Visualizza i riepiloghi dei risultati delle query generati da AI Assistant nella pagina Discover

OpenSearch Il servizio consente di interrogare i dati con istruzioni in linguaggio naturale utilizzando il linguaggio di query PPL (Piped Processing Language) nella pagina Discover. Ad esempio, puoi scrivere query come le seguenti:

- Ci sono errori nei miei registri degli errori?
- Qual è la dimensione media delle richieste per settimana?
- Quante richieste ci sono state raggruppate per codice di risposta la settimana scorsa?

In risposta, AI Assistant genera riepiloghi in linguaggio naturale dei risultati delle query sulla base dei primi dieci record, come i seguenti:



La combinazione della generazione di query in linguaggio naturale e dei riepiloghi delle query può aggiungere un ulteriore livello di indagine durante la risoluzione dei problemi di un avviso o un mezzo intuitivo per comprendere i dati senza dover scrivere query complesse.

Per visualizzare i riepiloghi dei risultati delle query generate da AI Assistant nella pagina Discover

1. Verifica di aver configurato AI Assistant for OpenSearch Service.
2. Nel menu principale OpenSearch delle dashboard, scegli Scopri.
3. Nell'elenco a discesa della lingua di interrogazione, scegli PPL.
4. Nella casella di testo AI Assistant, inserisci una richiesta, quindi fai clic sul pulsante accanto alla casella di testo per eseguire la query. AI Assistant può impiegare fino a 10 secondi per restituire un riepilogo. Dopo la query iniziale, devi scegliere Genera riepilogo per i riepiloghi successivi.

> **Note**
>
> Puoi disattivare la generazione di riepiloghi dall'elenco a discesa AI Assistant in Discover.

Visualizza i rilevatori di anomalie consigliati

Il rilevamento delle anomalie in Amazon OpenSearch Service rileva automaticamente le anomalie nei OpenSearch dati quasi in tempo reale utilizzando l'algoritmo Random Cut Forest (RCF). L'RCF

è un algoritmo di machine learning non supervisionato che modella un disegno del flusso di dati in ingresso. L'algoritmo calcola un `anomaly grade` e un valore `confidence score` per ogni punto di dati in ingresso. Il rilevamento delle anomalie utilizza questi valori per differenziare un'anomalia dalle normali variazioni dei dati.

Per semplificare il processo di creazione dei rilevatori di anomalie, AI Assistant può generare rilevatori suggeriti in base alla fonte di dati selezionata nella pagina Discover. AI Assistant supporta i rilevatori di anomalie suggeriti per qualsiasi lingua.

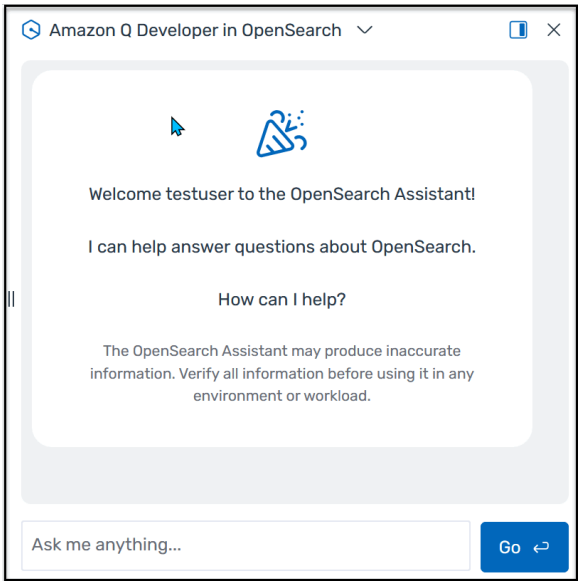
Per visualizzare i rilevatori di anomalie consigliati da AI Assistant

1. Verifica di aver configurato AI Assistant for OpenSearch Service.
2. Nel menu principale dell' OpenSearch interfaccia utente, scegli la pagina Scopri, quindi scegli un'origine dati.
3. Dal menu Assistente AI, scegli Suggerisci un rilevatore di anomalie, come mostrato nella schermata seguente.

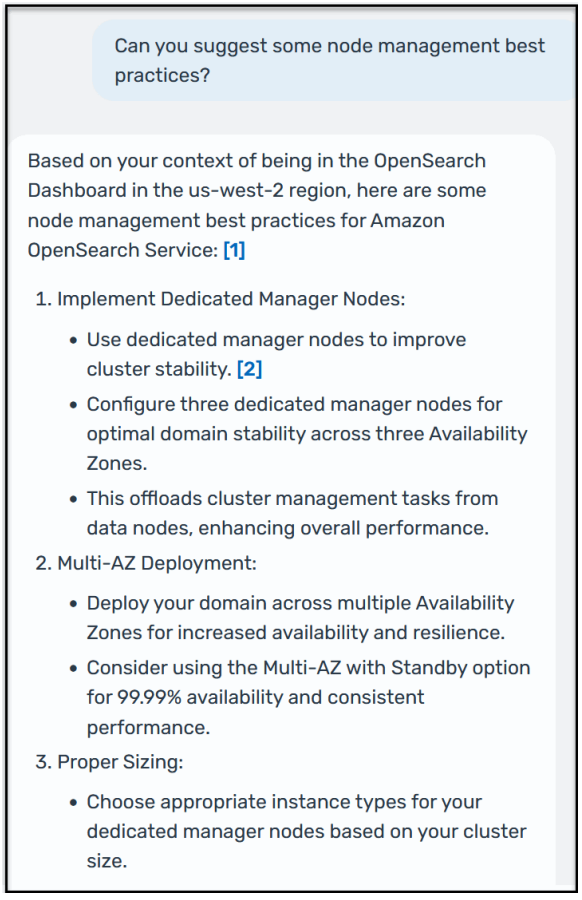
   AI Assistant può impiegare alcuni secondi per generare le funzionalità del rilevatore.
4. Scegli Crea rilevatore.

Accedi alla chat di AI Assistant per domande OpenSearch correlate

Se hai domande su Amazon OpenSearch Service, incluse domande concettuali o procedurali relative alle caratteristiche o funzionalità del OpenSearch Servizio, puoi chiedere informazioni a AI Assistant. Il chatbot supporta una chat conversazionale che preserva il contesto della discussione. Salva anche una cronologia delle conversazioni per consultarle in un secondo momento. Per accedere al chatbot, fai clic sull'icona dell'Assistente AI nell'angolo in alto a destra di qualsiasi pagina di OpenSearch servizio. Digita una domanda nella casella di testo e fai clic su Vai. AI Assistant può impiegare fino a 30 secondi per generare una risposta.



Ecco un esempio di risposta di AI Assistant.



Apprendimento automatico per Amazon OpenSearch Service

ML Commons è un OpenSearch plug-in che fornisce una serie di algoritmi di machine learning (ML) comuni tramite il trasporto e le chiamate API REST. Queste chiamate scelgono i nodi e le risorse giusti per ogni richiesta ML e monitorano le attività ML per garantire l'operatività. Ciò consente di sfruttare gli algoritmi ML open source esistenti e di ridurre lo sforzo richiesto per sviluppare nuove funzionalità di machine learning. Per ulteriori informazioni sul plug-in, consulta Machine learning nella documentazione. OpenSearch Questo capitolo spiega come utilizzare il plug-in con Amazon OpenSearch Service.

Argomenti

- Connettori Amazon OpenSearch Service ML per Servizi AWS
- Connettori Amazon OpenSearch Service ML per piattaforme di terze parti
- Utilizzo CloudFormation per configurare l'inferenza remota per la ricerca semantica
- Impostazioni ML Commons non supportate
- OpenSearch Modelli di framework di flussi di servizio

Connettori Amazon OpenSearch Service ML per Servizi AWS

Quando utilizzi connettori di apprendimento automatico (ML) di Amazon OpenSearch Service con un altro Servizio AWS, devi configurare un ruolo IAM per connettere in modo sicuro OpenSearch Service a quel servizio. Servizi AWS che puoi configurare un connettore per includere Amazon SageMaker AI e Amazon Bedrock. In questo tutorial, spieghiamo come creare un connettore da OpenSearch Service a SageMaker Runtime. Per ulteriori informazioni sui connettori, consulta Connettori supportati.

Argomenti

- Prerequisiti
- Crea un connettore di servizio OpenSearch

Prerequisiti

Per creare un connettore, devi disporre di un endpoint di dominio Amazon SageMaker AI e di un ruolo IAM che garantisca l'accesso OpenSearch al servizio.

Configura un dominio Amazon SageMaker AI

Consulta Deploy a Model in Amazon SageMaker AI nella Amazon SageMaker AI Developer Guide per implementare il tuo modello di machine learning. Prendi nota dell'URL dell'endpoint per il tuo modello, di cui hai bisogno per creare un connettore AI.

Creazione di un ruolo IAM

Imposta un ruolo IAM per delegare le autorizzazioni SageMaker di Runtime a Service. OpenSearch Per creare un nuovo ruolo, consulta Creating an IAM role (console) nella IAM User Guide. Facoltativamente, puoi utilizzare un ruolo esistente purché abbia lo stesso set di privilegi. Se crei un nuovo ruolo invece di usare un ruolo AWS gestito, sostituiscilo `opensearch-sagemaker-role` in questo tutorial con il nome del tuo ruolo.

1. Allega la seguente policy IAM gestita al tuo nuovo ruolo per consentire a OpenSearch Service di accedere al tuo endpoint SageMaker AI. Per allegare una policy a un ruolo, consulta Aggiungere i permessi di identità IAM.

   JSON

   ```
   {
       "Version":"2012-10-17",
       "Statement": [
           {
               "Action": [
                   "sagemaker:InvokeEndpointAsync",
                   "sagemaker:InvokeEndpoint"
               ],
               "Effect": "Allow",
               "Resource": "*"
           }
       ]
   }
   ```

2. Segui le istruzioni riportate in Modifica della politica di fiducia di un ruolo per modificare la relazione di trust del ruolo. Nella seguente politica, sostituiscila *`service-principal`* con uno dei seguenti principali di servizio per OpenSearch Service o OpenSearch Serverless:

Per Service OpenSearch

```
opensearchservice.amazonaws.com
```

Per OpenSearch Serverless

```
ml.opensearchservice.amazonaws.com
```

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
                "sts:AssumeRole"
            ],
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "ml.opensearchservice.amazonaws.com"
                ]
            }
        }
    ]
}
```

Ti consigliamo di utilizzare i tasti `aws:SourceAccount` e `aws:SourceArn` condition per limitare l'accesso a un dominio specifico. `SourceAccount`è l' Account AWS ID che appartiene al proprietario del dominio e il `SourceArn` è l'ARN del dominio. Ad esempio, puoi aggiungere il seguente blocco di condizioni alla politica di fiducia:

```
"Condition": {
    "StringEquals": {
        "aws:SourceAccount": "account-id"
    },
    "ArnLike": {
        "aws:SourceArn": "arn:aws:es:region:account-id:domain/domain-name"
    }
}
```

Configurazione delle autorizzazioni

Per creare il connettore, è necessaria l'autorizzazione per passare il ruolo IAM a OpenSearch Service. Devi inoltre disporre dell'accesso all'operazione `es:ESHttpPost`. Per concedere entrambe queste autorizzazioni, collega la policy seguente al ruolo IAM le cui credenziali vengono utilizzate per firmare la richiesta:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": "arn:aws:iam::111122223333:role/opensearch-sagemaker-
role"
        },
        {
            "Effect": "Allow",
            "Action": "es:ESHttpPost",
            "Resource": "arn:aws:es:us-east-1:111122223333:domain/domain-name/*"
        }
    ]
}
```

Se il tuo utente o ruolo non dispone `iam:PassRole` delle autorizzazioni per trasferire il ruolo, potresti riscontrare un errore di autorizzazione quando tenti di registrare un repository nel passaggio successivo.

Mappa il ruolo ML nelle OpenSearch dashboard (se utilizzi un controllo di accesso granulare)

Il controllo granulare degli accessi introduce un passaggio aggiuntivo per la configurazione di un connettore. Anche se si utilizza l'autenticazione di base HTTP per tutti gli altri scopi, è necessario mappare il ruolo `ml_full_access` al ruolo IAM che dispone delle autorizzazioni `iam:PassRole` per inviare `opensearch-sagemaker-role`.

1. Vai al plug-in OpenSearch Dashboards per il tuo dominio di servizio. OpenSearch Puoi trovare l'endpoint Dashboards nella dashboard del tuo dominio nella OpenSearch console di servizio.
2. Dal menu principale scegli Sicurezza, Ruoli e seleziona il ruolo ml_full_access.
3. Scegliere Utenti mappati, Gestisci mappatura.
4. In Ruoli di backend, aggiungi l'ARN del ruolo che ha le autorizzazioni da passare. `opensearch-sagemaker-role`

   ```
   arn:aws:iam::account-id:role/role-name
   ```

5. Selezionare Mappa e confermare che l'utente o il ruolo venga visualizzato in Utenti mappati.

Crea un connettore di servizio OpenSearch

Per creare un connettore, invia una `POST` richiesta all'endpoint del dominio di OpenSearch servizio. Puoi usare curl, il client Python di esempio, Postman o un altro metodo per inviare una richiesta firmata. Nota che non puoi usare una `POST` richiesta nella console Kibana. La richiesta ha il seguente formato:

```
POST domain-endpoint/_plugins/_ml/connectors/_create
{
   "name": "sagemaker: embedding",
   "description": "Test connector for Sagemaker embedding model",
   "version": 1,
   "protocol": "aws_sigv4",
   "credential": {
      "roleArn": "arn:aws:iam::account-id:role/opensearch-sagemaker-role"
   },
   "parameters": {
      "region": "region",
      "service_name": "sagemaker"
   },
   "actions": [
      {
         "action_type": "predict",
         "method": "POST",
         "headers": {
            "content-type": "application/json"
         },
         "url": "https://runtime.sagemaker.region.amazonaws.com/endpoints/endpoint-id/
invocations",
```

```
        "request_body": "{ \"inputs\": { \"question\": \"${parameters.question}\",
 \"context\": \"${parameters.context}\" } }"
      }
   ]
}
```

Se il dominio risiede all'interno di un cloud privato virtuale (VPC), il computer deve essere connesso al VPC affinché la richiesta crei correttamente il connettore AI. L'accesso a un VPC varia in base alla configurazione di rete, ma di solito comporta la connessione a una VPN o a una rete aziendale. Per verificare di poter accedere al dominio del OpenSearch servizio, accedi a `https://your-vpc-domain.region.es.amazonaws.com` In un browser Web e verifica di ricevere la risposta JSON predefinita.

Client Python di esempio

Il client Python è più semplice da automatizzare rispetto a una richiesta HTTP e ha una migliore riusabilità. Per creare il connettore AI con il client Python, salva il seguente codice di esempio in un file Python. Il client richiede i pacchetti AWS SDK per Python (Boto3)`requests`, e `requests-aws4auth`.

```
import boto3
import requests
from requests_aws4auth import AWS4Auth

host = 'domain-endpoint/'
region = 'region'
service = 'es'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

# Register repository
path = '_plugins/_ml/connectors/_create'
url = host + path

payload = {
   "name": "sagemaker: embedding",
   "description": "Test connector for Sagemaker embedding model",
   "version": 1,
   "protocol": "aws_sigv4",
   "credential": {
      "roleArn": "arn:aws:iam::account-id:role/opensearch-sagemaker-role"
```

```
   },
   "parameters": {
      "region": "region",
      "service_name": "sagemaker"
   },
   "actions": [
      {
         "action_type": "predict",
         "method": "POST",
         "headers": {
            "content-type": "application/json"
         },
         "url": "https://runtime.sagemaker.region.amazonaws.com/endpoints/endpoint-id/
invocations",
         "request_body": "{ \"inputs\": { \"question\": \"${parameters.question}\",
 \"context\": \"${parameters.context}\" } }"
      }
   ]
}
headers = {"Content-Type": "application/json"}

r = requests.post(url, auth=awsauth, json=payload, headers=headers)
print(r.status_code)
print(r.text)
```

Connettori Amazon OpenSearch Service ML per piattaforme di terze parti

In questo tutorial, spieghiamo come creare un connettore da OpenSearch Service a Cohere. Per ulteriori informazioni sui connettori, consulta Connettori supportati.

Quando utilizzi un connettore di apprendimento automatico (ML) di Amazon OpenSearch Service con un modello remoto esterno, devi memorizzare le tue credenziali di autorizzazione specifiche in Gestione dei segreti AWS. Potrebbe trattarsi di una chiave API o di una combinazione di nome utente e password. Ciò significa che devi anche creare un ruolo IAM che consenta l'accesso al OpenSearch servizio per la lettura da Secrets Manager.

Argomenti

- Prerequisiti
- Crea un connettore di servizio OpenSearch

Prerequisiti

Per creare un connettore per Cohere o per qualsiasi provider esterno con OpenSearch Service, devi disporre di un ruolo IAM che conceda l'accesso al OpenSearch Servizio Gestione dei segreti AWS, a cui archiviare le tue credenziali. È inoltre necessario memorizzare le credenziali in Secrets Manager.

Creazione di un ruolo IAM

Imposta un ruolo IAM per delegare le autorizzazioni di Secrets Manager a OpenSearch Service. Puoi anche utilizzare il ruolo esistente`SecretManagerReadWrite`. Per creare un nuovo ruolo, consulta Creating an IAM role (console) nella IAM User Guide. Se crei un nuovo ruolo invece di utilizzare un ruolo AWS gestito, `opensearch-secretmanager-role` sostituiscilo in questo tutorial con il nome del tuo ruolo.

1. Allega la seguente policy IAM gestita al tuo nuovo ruolo per consentire a OpenSearch Service di accedere ai tuoi valori di Secrets Manager. Per allegare una policy a un ruolo, consulta Aggiungere autorizzazioni di identità IAM.

   JSON

   ```
   {
       "Version":"2012-10-17",
       "Statement": [
           {
               "Action": [
                   "secretsmanager:GetSecretValue"
               ],
               "Effect": "Allow",
               "Resource": "*"
           }
       ]
   }
   ```

2. Segui le istruzioni riportate in Modifica della politica di fiducia di un ruolo per modificare la relazione di trust del ruolo. Nella seguente politica, sostituiscila *`service-principal`* con uno dei seguenti principali di servizio per OpenSearch Service o OpenSearch Serverless:

   Per Service OpenSearch

   `opensearchservice.amazonaws.com`

Per OpenSearch Serverless

`ml.opensearchservice.amazonaws.com`

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
                "sts:AssumeRole"
            ],
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "service-principle"
                ]
            }
        }
    ]
}
```

Ti consigliamo di utilizzare i tasti `aws:SourceAccount` e `aws:SourceArn` condition per limitare l'accesso a un dominio specifico. `SourceAccount`è l' Account AWS ID che appartiene al proprietario del dominio e il `SourceArn` è l'ARN del dominio. Ad esempio, puoi aggiungere il seguente blocco di condizioni alla politica di fiducia:

```
"Condition": {
    "StringEquals": {
        "aws:SourceAccount": "account-id"
    },
    "ArnLike": {
        "aws:SourceArn": "arn:aws:es:region:account-id:domain/domain-name"
    }
}
```

Configurazione delle autorizzazioni

Per creare il connettore, è necessaria l'autorizzazione per passare il ruolo IAM a OpenSearch Service. Devi inoltre disporre dell'accesso all'operazione `es:ESHttpPost`. Per concedere entrambe

queste autorizzazioni, collega la policy seguente al ruolo IAM le cui credenziali vengono utilizzate per firmare la richiesta:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "iam:PassRole",
      "Resource": "arn:aws:iam::111122223333:role/opensearch-secretmanager-role"
    },
    {
      "Effect": "Allow",
      "Action": "es:ESHttpPost",
      "Resource": "arn:aws:es:us-east-1:111122223333:domain/domain-name/*"
    }
  ]
}
```

Se il tuo utente o ruolo non dispone `iam:PassRole` delle autorizzazioni per trasferire il ruolo, potresti riscontrare un errore di autorizzazione quando tenti di registrare un repository nel passaggio successivo.

Configurare Gestione dei segreti AWS

Per memorizzare le credenziali di autorizzazione in Secrets Manager, consulta Creare un Gestione dei segreti AWS segreto nella Guida per l'Gestione dei segreti AWS utente.

Dopo che Secrets Manager ha accettato la coppia chiave-valore come segreta, riceverai un ARN con il formato:. `arn:aws:secretsmanager:us-west-2:123456789012:secret:MySecret-a1b2c3` Tieni traccia di questo ARN, così come lo usi, e della tua chiave quando crei un connettore nel passaggio successivo.

Mappa il ruolo ML nelle OpenSearch dashboard (se utilizzi un controllo di accesso granulare)

Il controllo granulare degli accessi introduce un passaggio aggiuntivo per la configurazione di un connettore. Anche se si utilizza l'autenticazione di base HTTP per tutti gli altri scopi, è necessario

mappare il ruolo `ml_full_access` al ruolo IAM che dispone delle autorizzazioni `iam:PassRole` per inviare `opensearch-sagemaker-role`.

1. Vai al plug-in OpenSearch Dashboards per il tuo dominio di servizio. OpenSearch Puoi trovare l'endpoint Dashboards nella dashboard del tuo dominio nella OpenSearch console di servizio.
2. Dal menu principale scegli Sicurezza, Ruoli e seleziona il ruolo ml_full_access.
3. Scegliere Utenti mappati, Gestisci mappatura.
4. In Ruoli di backend, aggiungi l'ARN del ruolo che ha le autorizzazioni da passare. `opensearch-sagemaker-role`

   ```
   arn:aws:iam::account-id:role/role-name
   ```

5. Selezionare Mappa e confermare che l'utente o il ruolo venga visualizzato in Utenti mappati.

Crea un connettore di servizio OpenSearch

Per creare un connettore, invia una `POST` richiesta all'endpoint del dominio di OpenSearch servizio. Puoi usare curl, il client Python di esempio, Postman o un altro metodo per inviare una richiesta firmata. Nota che non puoi usare una `POST` richiesta nella console Kibana. La richiesta ha il seguente formato:

```
POST domain-endpoint/_plugins/_ml/connectors/_create
{
    "name": "Cohere Connector: embedding",
    "description": "The connector to cohere embedding model",
    "version": 1,
    "protocol": "http",
    "credential": {
        "secretArn": "arn:aws:secretsmanager:region:account-id:secret:cohere-key-id",
        "roleArn": "arn:aws:iam::account-id:role/opensearch-secretmanager-role"
    },
    "actions": [
        {
            "action_type": "predict",
            "method": "POST",
            "url": "https://api.cohere.ai/v1/embed",
            "headers": {
                "Authorization": "Bearer ${credential.secretArn.cohere-key-used-in-
secrets-manager}"
            },
```

```
            "request_body": "{ \"texts\": ${parameters.texts}, \"truncate\": \"END\" }"
        }
    ]
}
```

Il corpo della richiesta per questa richiesta è diverso da quello di una richiesta di connettore open source in due modi. All'interno del `credential` campo, si passa l'ARN per il ruolo IAM che consente a OpenSearch Service di leggere da Secrets Manager, insieme all'ARN per il segreto what. Nel `headers` campo, si fa riferimento al segreto utilizzando la chiave segreta e al fatto che proviene da un ARN.

Se il dominio risiede all'interno di un cloud privato virtuale (VPC), il computer deve essere connesso al VPC affinché la richiesta crei correttamente il connettore AI. L'accesso a un VPC varia in base alla configurazione di rete, ma di solito comporta la connessione a una VPN o a una rete aziendale. Per verificare di poter accedere al dominio del OpenSearch servizio, accedi a `https://your-vpc-domain.region.es.amazonaws.com` In un browser Web e verifica di ricevere la risposta JSON predefinita.

Client Python di esempio

Il client Python è più semplice da automatizzare rispetto a una richiesta HTTP e ha una migliore riusabilità. Per creare il connettore AI con il client Python, salva il seguente codice di esempio in un file Python. Il client richiede i pacchetti AWS SDK per Python (Boto3)`requests`, e `requests-aws4auth`.

```
import boto3
import requests
from requests_aws4auth import AWS4Auth

host = 'domain-endpoint/'
region = 'region'
service = 'es'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

path = '_plugins/_ml/connectors/_create'
url = host + path

payload = {
    "name": "Cohere Connector: embedding",
```

```
    "description": "The connector to cohere embedding model",
    "version": 1,
    "protocol": "http",
    "credential": {
        "secretArn": "arn:aws:secretsmanager:region:account-id:secret:cohere-key-id",
        "roleArn": "arn:aws:iam::account-id:role/opensearch-secretmanager-role"
    },
    "actions": [
        {
            "action_type": "predict",
            "method": "POST",
            "url": "https://api.cohere.ai/v1/embed",
            "headers": {
                "Authorization": "Bearer ${credential.secretArn.cohere-key-used-in-
secrets-manager}"
            },
            "request_body": "{ \"texts\": ${parameters.texts}, \"truncate\": \"END\" }"
        }
    ]
}

headers = {"Content-Type": "application/json"}

r = requests.post(url, auth=awsauth, json=payload, headers=headers)
print(r.status_code)
print(r.text)
```

Utilizzo CloudFormation per configurare l'inferenza remota per la ricerca semantica

A partire dalla OpenSearch versione 2.9, puoi utilizzare l'inferenza remota con la ricerca semantica per ospitare i tuoi modelli di machine learning (ML). L'inferenza remota utilizza il plug-in ML Commons.

Con l'inferenza remota, puoi ospitare le inferenze dei tuoi modelli in remoto su servizi ML, come Amazon AI SageMaker e Amazon Bedrock, e collegarli ad Amazon OpenSearch Service con connettori ML.

Per facilitare la configurazione dell'inferenza remota, Amazon OpenSearch Service fornisce un AWS CloudFormationmodello nella console. CloudFormation è un ambiente Servizio AWS in cui è possibile effettuare il provisioning AWS e gestire risorse di terze parti trattando l'infrastruttura come codice.

Il OpenSearch CloudFormation modello automatizza il processo di provisioning del modello per te, in modo che tu possa creare facilmente un modello nel tuo dominio di OpenSearch servizio e quindi utilizzare l'ID del modello per importare dati ed eseguire query di ricerca neurali.

Quando utilizzi codificatori neurali sparsi con la versione 2.12 e successive del OpenSearch servizio, ti consigliamo di utilizzare il modello tokenizer localmente anziché distribuirlo in remoto. Per ulteriori informazioni, consulta i modelli di codifica Sparse nella documentazione. OpenSearch

Argomenti

- Modelli disponibili CloudFormation
- Prerequisiti
- Modelli Amazon Bedrock
- Configurazione di Agentic Search con Bedrock Claude
- Modelli di integrazione con server MCP
- SageMaker Modelli Amazon
- Inferenza remota per modelli di evidenziazione semantica

Modelli disponibili CloudFormation

I seguenti modelli di AWS CloudFormation machine learning (ML) sono disponibili per l'uso:

Modelli Amazon Bedrock

Integrazione con Amazon Titan Text Embeddings

Si connette ai modelli ML ospitati di Amazon Bedrock, elimina la necessità di implementare modelli separati e utilizza endpoint Amazon Bedrock predeterminati. Per ulteriori informazioni, consulta Amazon Titan Text Embeddings nella Guida per l'utente di Amazon Bedrock.

Integrazione con Cohere Embed

Fornisce l'accesso ai modelli Cohere Embed ed è ottimizzato per flussi di lavoro di elaborazione del testo specifici. Per ulteriori informazioni, consulta Incorpora nel sito Web di Cohere docs.

Incorporamenti multimodali di Amazon Titan

Supporta l'incorporamento di testo e immagini e abilita funzionalità di ricerca multimodale. Per ulteriori informazioni, consulta Amazon Titan Multimodal Embeddings nella Amazon Bedrock User Guide.

Modelli di integrazione con server MCP

Integrazione con server MCP

Implementa un Amazon Bedrock AgentCore Runtime, fornisce un endpoint agente, gestisce l'autenticazione in entrata e in uscita e supporta l'autenticazione aziendale. OAuth

SageMaker Modelli Amazon

Integrazione con modelli di incorporamento del testo tramite Amazon SageMaker

Implementa modelli di incorporamento di testo in Amazon SageMaker Runtime, crea ruoli IAM per l'accesso agli artefatti del modello e stabilisce connettori ML per la ricerca semantica.

Integrazione con Sparse Encoders tramite SageMaker

Imposta modelli di codifica sparsi per la ricerca neurale, crea AWS Lambda funzioni per la gestione dei connettori e restituisce il modello per l'uso immediato. IDs

Prerequisiti

Per utilizzare un CloudFormation modello con OpenSearch Service, completa i seguenti prerequisiti.

Configura un dominio di OpenSearch servizio

Prima di poter utilizzare un CloudFormation modello, devi configurare un dominio Amazon OpenSearch Service con la versione 2.9 o successiva e il controllo granulare degli accessi abilitato. Crea un ruolo OpenSearch di backend del servizio per autorizzare il plug-in ML Commons a creare il connettore per te.

Il CloudFormation modello crea per te un ruolo Lambda IAM con il nome predefinito`LambdaInvokeOpenSearchMLCommonsRole`, che puoi sostituire se desideri scegliere un nome diverso. Dopo che il modello ha creato questo ruolo IAM, devi autorizzare la funzione Lambda a chiamare il tuo dominio di OpenSearch servizio. A tale scopo, associa il ruolo denominato `ml_full_access` al tuo ruolo OpenSearch di backend Service con i seguenti passaggi:

1. Vai al plug-in OpenSearch Dashboards per il tuo dominio di OpenSearch servizio. Puoi trovare l'endpoint Dashboards nella dashboard del tuo dominio nella OpenSearch console di servizio.
2. Dal menu principale scegli Sicurezza, Ruoli e seleziona il ruolo ml_full_access.
3. Scegliere Utenti mappati, Gestisci mappatura.

4. In Ruoli di backend, aggiungi l'ARN del ruolo Lambda che richiede l'autorizzazione per chiamare il tuo dominio.

```
arn:aws:iam::account-id:role/role-name
```

5. Selezionare Mappa e confermare che l'utente o il ruolo venga visualizzato in Utenti mappati.

Dopo aver mappato il ruolo, vai alla configurazione di sicurezza del tuo dominio e aggiungi il ruolo Lambda IAM alla OpenSearch tua policy di accesso al servizio.

Abilita le autorizzazioni sul tuo Account AWS

Account AWS Devi avere l'autorizzazione per accedere CloudFormation e Lambda, insieme a qualsiasi modello Servizio AWS tu scelga per il tuo modello, che sia Runtime SageMaker o Amazon Bedrock.

Se utilizzi Amazon Bedrock, devi anche registrare il tuo modello. Consulta Model access nella Amazon Bedrock User Guide per registrare il tuo modello.

Se utilizzi il tuo bucket Amazon S3 per fornire artefatti del modello, devi aggiungere il ruolo CloudFormation IAM alla tua policy di accesso S3. Per ulteriori informazioni, consulta Aggiunta e rimozione di autorizzazioni per identità IAM nella Guida per l'utente di IAM .

Modelli Amazon Bedrock

I CloudFormation modelli Amazon Bedrock forniscono le AWS risorse necessarie per creare connettori tra OpenSearch Service e Amazon Bedrock.

Innanzitutto, il modello crea un ruolo IAM che consente alla futura funzione Lambda di accedere al dominio del OpenSearch servizio. Il modello crea quindi la funzione Lambda, che consente al dominio di creare un connettore utilizzando il plug-in ML Commons. Dopo che OpenSearch Service ha creato il connettore, la configurazione dell'inferenza remota è terminata e puoi eseguire ricerche semantiche utilizzando le operazioni dell'API Amazon Bedrock.

> **Note**
>
> Poiché Amazon Bedrock ospita i propri modelli di machine learning, non è necessario distribuire un modello in Runtime. SageMaker Al contrario, il modello utilizza un endpoint predeterminato per Amazon Bedrock e salta le fasi di fornitura degli endpoint.

Per utilizzare il modello Amazon Bedrock CloudFormation

1. Apri la console Amazon OpenSearch Service.
2. Nel riquadro di navigazione a sinistra, scegli Integrazioni.
3. Nella sezione Integrazione con il modello Amazon Titan Text Embeddings tramite Amazon Bedrock, scegli Configura dominio, Configura dominio pubblico.
4. Segui le istruzioni per configurare il tuo modello.

> **Note**
>
> OpenSearch Il servizio fornisce anche un modello separato per configurare un dominio Amazon VPC. Se utilizzi questo modello, devi fornire l'ID Amazon VPC per la funzione Lambda.

Inoltre, OpenSearch Service fornisce i seguenti modelli Amazon Bedrock per connettersi al modello Cohere e al modello Amazon Titan Multimodal Embeddings:

- `Integration with Cohere Embed through Amazon Bedrock`
- `Integrate with Amazon Bedrock Titan Multi-modal`

Configurazione di Agentic Search con Bedrock Claude

Agentic search sfrutta agenti autonomi per eseguire ricerche complesse per conto dell'utente, comprendendo le intenzioni degli utenti, orchestrando gli strumenti giusti, generando query ottimizzate e fornendo riepiloghi trasparenti delle loro decisioni attraverso un'interfaccia in linguaggio naturale. Questi agenti si basano su modelli di ragionamento, come Bedrock Claude.

Segui i passaggi seguenti per aprire ed eseguire un CloudFormation modello che configuri automaticamente i modelli Bedrock Claude per la ricerca agentica e come configurare e creare i tuoi agenti nel plug-in AI Search Flows su Dashboards. OpenSearch

Attivazione di Bedrock Claude Access

1. Prerequisito: se il tuo dominio utilizza un controllo degli accessi dettagliato, esegui il mapping `arn:aws:iam::`*`your-account-id`*`:role/`*`LambdaInvokeOpenSearchMLCommonsRole`*

come ruolo di backend al ruolo prima di eseguire il modello. `ml_full_access` Questo ruolo IAM verrà creato automaticamente CloudFormation se non esiste già. Per ulteriori informazioni su come configurare la mappatura, consultaMappa il ruolo ML nelle OpenSearch dashboard (se utilizzi un controllo di accesso granulare).

2. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
3. Nella barra di navigazione a sinistra, scegli Integrazioni.
4. In Integrazione con Bedrock Claude for Agentic Search, scegli Configura dominio. Assicurati che il tuo dominio sia nella versione 3.3 o successiva.
5. Nel CloudFormation modello, inserisci l'endpoint OpenSearch del dominio di servizio e seleziona un modello. I campi rimanenti sono facoltativi o precompilati. Fai clic su Crea stack e attendi il completamento del provisioning.
6. Dalla console di Amazon OpenSearch Service, seleziona Domini e seleziona il tuo dominio. Fai clic sull'URL delle OpenSearch dashboard per accedere OpenSearch alle dashboard.

Creazione di agenti ed esecuzione di Agentic Search

1. Da OpenSearch Dashboards, apri il menu sul lato sinistro. Seleziona OpenSearch Plugin > AI Search Flows per accedere al plugin.
2. Nella pagina Flussi di lavoro, seleziona la scheda Nuovo flusso di lavoro e, nella scheda Agentic Search, fai clic su Crea.
3. Fornisci un nome univoco per la configurazione di ricerca e fai clic su Crea.
4. In Configura agente, fai clic su Crea nuovo agente. Seleziona il modello Bedrock Claude appena creato, quindi fai clic su Crea agente. Se il pulsante è disabilitato, seleziona Impostazioni avanzate > Interfaccia LLM e assicurati che sia selezionata un'interfaccia valida. Tutti i modelli CloudFormation saranno modelli Bedrock Claude, quindi puoi selezionare Bedrock Claude, se non lo è già, quindi fare clic su Crea agente.
5. In Test flow, prova a eseguire ricerche agentiche. Fornisci una query di ricerca in linguaggio naturale e fai clic su Cerca.

Per la documentazione completa del plug-in AI Search Flows, consulta Configuring Agentic Search nella documentazione. OpenSearch

Per ulteriori informazioni su come funziona Agentic Search, consulta Agentic Search nella documentazione. OpenSearch

Modelli di integrazione con server MCP

Con i modelli di server Model Context Protocol (MCP), puoi implementare un server MCP OpenSearch ospitato su Amazon Bedrock AgentCore, riducendo la complessità di integrazione tra agenti e strumenti di intelligenza artificiale. OpenSearch Per ulteriori informazioni, consulta Cos'è Amazon Bedrock AgentCore? .

Caratteristiche del modello

Questo modello include le seguenti funzionalità chiave per la distribuzione e la gestione del server MCP.

Implementazione gestita del server MCP

Esegue la distribuzione opensearch-mcp-server-pyutilizzando Amazon Bedrock AgentCore Runtime e fornisce un endpoint agente che invia tramite proxy le richieste al server MCP sottostante. Per ulteriori informazioni, consulta on. opensearch-mcp-server-pyGitHub

Autenticazione e sicurezza

Gestisce sia l'autenticazione in entrata (dagli utenti al server MCP) che l'autenticazione in uscita (dal server MCP a OpenSearch) e supporta OAuth l'autenticazione aziendale.

> **Note**
>
> Il modello di server MCP è disponibile solo nei seguenti casi: Regioni AWS
>
> - Stati Uniti orientali (Virginia settentrionale)
> - Stati Uniti occidentali (Oregon)
> - Europa (Francoforte)
> - Asia Pacifico (Sydney)

Per utilizzare il modello di server MCP

Segui questi passaggi per distribuire il modello di server MCP e collegarlo al tuo OpenSearch dominio.

1. Apri la console Amazon OpenSearch Service.

2. Nel riquadro di navigazione a sinistra, scegli Integrazioni.
3. Individua il modello di integrazione del server MCP.
4. Scegli Configura dominio. Quindi, inserisci l'endpoint OpenSearch del tuo dominio.

Il modello crea un AgentCore Runtime e i seguenti componenti, se non vengono specificati i parametri opzionali corrispondenti:

- Repository Amazon ECR
- Un pool di utenti Amazon Cognito come autorizzatore OAuth
- Un ruolo di esecuzione utilizzato dal Runtime AgentCore

Dopo aver completato questa procedura, è necessario seguire questi passaggi successivi alla creazione:

1. Per Amazon OpenSearch Service: associa il tuo ruolo di esecuzione ARN a un ruolo di OpenSearch backend per controllare l'accesso al tuo dominio.

   Per Amazon OpenSearch Serverless: crea una policy di accesso ai dati che consenta al tuo ruolo esecutivo di accedere alla tua raccolta.
2. Ottieni un token di OAuth accesso dal tuo autorizzatore. Quindi usa questo token per accedere al server MCP all'URL elencato nell'output CloudFormation dello stack.

Per ulteriori informazioni, consulta Azioni politiche per Serverless OpenSearch.

Integrazione con agenti AI

Dopo l'implementazione, è possibile integrare il server MCP con qualsiasi agente compatibile con MCP. Per ulteriori informazioni, consulta Richiamare il server MCP distribuito nella Amazon Bedrock Developer Guide.

Integrazione con gli sviluppatori

È possibile aggiungere l'endpoint del server MCP alla configurazione dell'agente. Puoi anche usarlo con Amazon Q Developer CLI, agenti personalizzati o altri agenti compatibili con MCP.

Distribuzione aziendale

Gli agenti ospitati centralmente possono connettersi a più servizi OpenSearch utilizzando un unico componente. Questo agente supporta tutti i sistemi di autenticazione aziendali OAuth ed è scalabile per supportare più utenti e casi d'uso.

Esempio di utilizzo del framework Strands Agents

```
import os
import requests
from strands import Agent
from strands.tools.mcp import MCPClient
from mcp.client.streamable_http import streamablehttp_client

def get_bearer_token(discovery_url: str, client_id: str, client_secret: str):
    response = requests.get(discovery_url)
    discovery_data = response.json()
    token_endpoint = discovery_data['token_endpoint']

    data = {
        'grant_type': 'client_credentials',
        'client_id': client_id,
        'client_secret': client_secret
    }
    headers = {
        'Content-Type': 'application/x-www-form-urlencoded'
    }

    response = requests.post(token_endpoint, data=data, headers=headers)
    token_data = response.json()
    return token_data['access_token']

if __name__ == "__main__":
    discovery_url = os.environ["DISCOVERY_URL"]
    client_id = os.environ["CLIENT_ID"]
    client_secret = os.environ["CLIENT_SECRET"]
    mcp_url = os.environ["MCP_URL"]

    bearer_token = get_bearer_token(discovery_url, client_id, client_secret)

    opensearch_mcp_client = MCPClient(lambda: streamablehttp_client(mcp_url, {
        "authorization": f"Bearer {bearer_token}",
        "Content-Type": "application/json"
```

```
    }))

    with opensearch_mcp_client:
        tools = opensearch_mcp_client.list_tools_sync()
        agent = Agent(tools=tools)
        agent("list indices")
```

Per ulteriori informazioni, consulta Hosting OpenSearch MCP Server with Amazon Bedrock AgentCore sul OpenSearch sito Web.

SageMaker Modelli Amazon

I SageMaker CloudFormation modelli Amazon definiscono più AWS risorse per configurare il plug-in neurale e la ricerca semantica per te.

Inizia utilizzando il modello Integrazione con modelli di incorporamento del testo tramite Amazon per distribuire un SageMaker modello di incorporamento di testo in SageMaker Runtime come server. Se non fornisci un endpoint modello, CloudFormation crea un ruolo IAM che consente a SageMaker Runtime di scaricare gli artefatti del modello da Amazon S3 e distribuirli sul server. Se fornisci un endpoint, CloudFormation crea un ruolo IAM che consente alla funzione Lambda di accedere al dominio OpenSearch del servizio o, se il ruolo esiste già, aggiorna e riutilizza il ruolo. L'endpoint serve il modello remoto utilizzato per il connettore ML con il plug-in ML Commons.

Quindi, utilizza il SageMaker modello Integration with Sparse Encoders through Amazon per creare una funzione Lambda che consenta al tuo dominio di configurare connettori di inferenza remoti. Dopo aver creato il connettore in OpenSearch Service, l'inferenza remota può eseguire una ricerca semantica utilizzando il modello remoto in Runtime. SageMaker Il modello ti restituisce l'ID del modello nel tuo dominio in modo che tu possa iniziare la ricerca.

Per utilizzare i CloudFormation modelli Amazon SageMaker AI

1. Apri la console Amazon OpenSearch Service.
2. Nel riquadro di navigazione a sinistra, scegli Integrazioni.
3. In ciascuno dei modelli Amazon SageMaker AI, scegli Configura dominio, Configura dominio pubblico.
4. Segui le istruzioni nella CloudFormation console per effettuare il provisioning dello stack e configurare un modello.

> **Note**
>
> OpenSearch Il servizio fornisce anche un modello separato per configurare il dominio VPC. Se utilizzi questo modello, devi fornire l'ID VPC per la funzione Lambda.

Inferenza remota per modelli di evidenziazione semantica

L'evidenziazione semantica è una funzionalità di ricerca avanzata che migliora la pertinenza dei risultati analizzando il significato e il contesto delle query anziché basarsi esclusivamente sulle corrispondenze esatte delle parole chiave. Questa funzionalità utilizza modelli di apprendimento automatico per valutare la somiglianza semantica tra le query di ricerca e il contenuto dei documenti, identificando ed evidenziando le frasi o i passaggi più rilevanti dal punto di vista contestuale all'interno dei documenti. A differenza dei metodi di evidenziazione tradizionali che si concentrano sulle corrispondenze esatte dei termini, l'evidenziazione semantica sfrutta i modelli di intelligenza artificiale per valutare ogni frase utilizzando informazioni contestuali provenienti sia dalla query che dal testo circostante, consentendole di far emergere informazioni pertinenti anche quando non sono presenti termini di ricerca esatti nei passaggi evidenziati. Questo approccio è particolarmente utile per le implementazioni di ricerca basate sull'intelligenza artificiale in cui gli utenti danno la priorità al significato semantico rispetto alla corrispondenza letterale delle parole, consentendo agli amministratori della ricerca di offrire esperienze di ricerca più intelligenti e contestualmente consapevoli che evidenziano intervalli di contenuti significativi anziché semplici occorrenze di parole chiave. Per ulteriori informazioni, consulta Utilizzo dell'evidenziazione semantica.

Utilizza la procedura seguente: apri ed esegui un CloudFormation modello che configura automaticamente i SageMaker modelli Amazon per l'evidenziazione semantica.

Per utilizzare il modello di evidenziazione semantica CloudFormation

1. Apri la console Amazon OpenSearch Service a https://console.aws.amazon.com/aos/casa.
2. Nella barra di navigazione a sinistra, scegli Integrazioni.
3. In Abilita l'evidenziazione semantica tramite l' SageMakerintegrazione di Amazon, scegli Configura dominio, Configura dominio pubblico.
4. Segui le istruzioni per configurare il tuo modello.

> **Note**
>
> OpenSearch Il servizio fornisce anche un modello separato per configurare il dominio VPC. Se utilizzi questo modello, devi fornire l'ID VPC per la funzione Lambda.

Impostazioni ML Commons non supportate

Amazon OpenSearch Service non supporta l'uso delle seguenti impostazioni ML Commons:

- `plugins.ml_commons.allow_registering_model_via_url`
- `plugins.ml_commons.allow_registering_model_via_local_file`

> **Important**
>
> Nei cluster di produzione, non disabilitare l'impostazione del cluster `plugins.ml_commons.only_run_on_ml_node` (non impostarla `false` su). L'opzione per disabilitare questa protezione serve a facilitare lo sviluppo, ma i cluster di produzione dovrebbero utilizzare i connettori. Per ulteriori informazioni, consulta the section called “Connettori per Servizi AWS”.

Per ulteriori informazioni sulle impostazioni del cluster ML Commons, vedere Impostazioni del cluster ML Commons.

OpenSearch Modelli di framework di flussi di servizio

I modelli di framework di flusso di Amazon OpenSearch Service consentono di automatizzare attività complesse di configurazione e preelaborazione del OpenSearch servizio fornendo modelli per casi d'uso comuni. Ad esempio, puoi utilizzare i modelli di Flow Framework per automatizzare le attività di configurazione dell'apprendimento automatico. I modelli di framework di flusso di Amazon OpenSearch Service forniscono una descrizione compatta del processo di configurazione in un documento JSON o YAML. Questi modelli descrivono configurazioni automatizzate di flussi di lavoro per chat conversazionali o generazione di query, connettori AI, strumenti, agenti e altri componenti che preparano OpenSearch Service per l'uso in backend per modelli generativi.

I modelli di framework di flusso di Amazon OpenSearch Service possono essere personalizzati per soddisfare esigenze specifiche. Per vedere un esempio di modello di framework di flusso personalizzato, consulta flow-framework. Per i modelli forniti dal OpenSearch servizio, vedere workflow-templates. Per una documentazione completa, che include passaggi dettagliati, un riferimento all'API e un riferimento a tutte le impostazioni disponibili, vedi Automating Configuration nella documentazione open source. OpenSearch

> **Note**
>
> Flow-framework non supporta il filtraggio dei ruoli di backend per Service 2.17. OpenSearch

Creazione di connettori ML in Service OpenSearch

I modelli di framework di flusso di Amazon OpenSearch Service consentono di configurare e installare connettori ML utilizzando l'API create connector offerta in ml-commons. Puoi utilizzare i connettori ML per connettere OpenSearch Service ad altri AWS servizi o piattaforme di terze parti. Per ulteriori informazioni su questo argomento, consulta Creazione di connettori per piattaforme ML di terze parti. L'API del framework di flusso di Amazon OpenSearch Service consente di automatizzare le attività di configurazione e preelaborazione del OpenSearch servizio e può essere utilizzata per creare connettori ML.

Prima di poter creare un connettore in OpenSearch Service, devi fare quanto segue:

- Crea un dominio Amazon SageMaker AI.
- Crea un ruolo IAM.
- Configura l'autorizzazione al pass role.
- Mappa i ruoli flow-framework e ml-commons nelle dashboard. OpenSearch

Per ulteriori informazioni su come configurare i connettori ML per AWS i servizi, consulta i connettori ML OpenSearch di Amazon Service per AWS i servizi. Per ulteriori informazioni sull'utilizzo dei connettori OpenSearch Service ML con piattaforme di terze parti, consulta Connettori Amazon OpenSearch Service ML per piattaforme di terze parti.

Creazione di un connettore tramite un servizio flow-framework

Per creare un modello di flow-framework con connettore, è necessario inviare una `POST` richiesta all'endpoint del dominio di servizio. OpenSearch Puoi usare cURL, un client Python di esempio,

Postman o un altro metodo per inviare una richiesta firmata. La POST richiesta ha il seguente formato:

```
POST /_plugins/_flow_framework/workflow
{
  "name": "Deploy Claude Model",
  "description": "Deploy a model using a connector to Claude",
  "use_case": "PROVISION",
  "version": {
    "template": "1.0.0",
    "compatibility": [
      "2.12.0",
      "3.0.0"
    ]
  },
  "workflows": {
    "provision": {
      "nodes": [
        {
          "id": "create_claude_connector",
          "type": "create_connector",
          "user_inputs": {
            "name": "Claude Instant Runtime Connector",
            "version": "1",
            "protocol": "aws_sigv4",
            "description": "The connector to Bedrock service for Claude model",
            "actions": [
              {
                "headers": {
                  "x-amz-content-sha256": "required",
                  "content-type": "application/json"
                },
                "method": "POST",
                "request_body": "{ \"prompt\":\"${parameters.prompt}\",
 \"max_tokens_to_sample\":${parameters.max_tokens_to_sample},
 \"temperature\":${parameters.temperature},  \"anthropic_version\":
\"${parameters.anthropic_version}\" }",
                "action_type": "predict",
                "url": "https://bedrock-runtime.us-west-2.amazonaws.com/model/
anthropic.claude-instant-v1/invoke"
              }
            ],
            "credential": {
```

```
                "roleArn": "arn:aws:iam::account-id:role/opensearch-secretmanager-
role"
             },
            "parameters": {
              "endpoint": "bedrock-runtime.us-west-2.amazonaws.com",
              "content_type": "application/json",
              "auth": "Sig_V4",
              "max_tokens_to_sample": "8000",
              "service_name": "bedrock",
              "temperature": "0.0001",
              "response_filter": "$.completion",
              "region": "us-west-2",
              "anthropic_version": "bedrock-2023-05-31"
            }
          }
        }
      ]
    }
  }
}
```

Se il tuo dominio risiede in un cloud privato virtuale (Amazon VPC), devi essere connesso ad Amazon VPC affinché la richiesta crei correttamente il connettore AI. L'accesso a un Amazon VPC varia in base alla configurazione di rete, ma di solito comporta la connessione a una VPN o a una rete aziendale. Per verificare di poter accedere al dominio del OpenSearch servizio, accedi a `https://`*`your-vpc-domain`*`.`*`region`*`.es.amazonaws.com` In un browser Web e verifica di ricevere la risposta JSON predefinita. (*`placeholder text`*Sostituiscili con i tuoi valori.

Client Python di esempio

Il client Python è più semplice da automatizzare rispetto a una HTTP richiesta e ha una migliore riutilizzabilità. Per creare il connettore AI con il client Python, salva il seguente codice di esempio in un file Python. Il client richiede i pacchetti AWS SDK for Python (Boto3), requests:HTTP forHumans e requests-aws4auth 1.2.3.

```
import boto3
import requests
from requests_aws4auth import AWS4Auth

host = 'domain-endpoint/'
region = 'region'
service = 'es'
```

```
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

path = '_plugins/_flow_framework/workflow'
url = host + path

payload = {
  "name": "Deploy Claude Model",
  "description": "Deploy a model using a connector to Claude",
  "use_case": "PROVISION",
  "version": {
    "template": "1.0.0",
    "compatibility": [
      "2.12.0",
      "3.0.0"
    ]
  },
  "workflows": {
    "provision": {
      "nodes": [
        {
          "id": "create_claude_connector",
          "type": "create_connector",
          "user_inputs": {
            "name": "Claude Instant Runtime Connector",
            "version": "1",
            "protocol": "aws_sigv4",
            "description": "The connector to Bedrock service for Claude model",
            "actions": [
              {
                "headers": {
                  "x-amz-content-sha256": "required",
                  "content-type": "application/json"
                },
                "method": "POST",
                "request_body": "{ \"prompt\":\"${parameters.prompt}\",
 \"max_tokens_to_sample\":${parameters.max_tokens_to_sample},
 \"temperature\":${parameters.temperature},  \"anthropic_version\":
\"${parameters.anthropic_version}\" }",
                "action_type": "predict",
                "url": "https://bedrock-runtime.us-west-2.amazonaws.com/model/
anthropic.claude-instant-v1/invoke"
              }
```

```
            ],
            "credential": {
                "roleArn": "arn:aws:iam::account-id:role/opensearch-secretmanager-
role"
             },
            "parameters": {
              "endpoint": "bedrock-runtime.us-west-2.amazonaws.com",
              "content_type": "application/json",
              "auth": "Sig_V4",
              "max_tokens_to_sample": "8000",
              "service_name": "bedrock",
              "temperature": "0.0001",
              "response_filter": "$.completion",
              "region": "us-west-2",
              "anthropic_version": "bedrock-2023-05-31"
            }
          }
        }
      ]
    }
  }
}

headers = {"Content-Type": "application/json"}

r = requests.post(url, auth=awsauth, json=payload, headers=headers)
print(r.status_code)
print(r.text)
```

Modelli di workflow predefiniti

Amazon OpenSearch Service fornisce diversi modelli di flusso di lavoro per alcuni casi d'uso comuni di machine learning (ML). L'utilizzo di un modello semplifica le configurazioni complesse e fornisce molti valori predefiniti per casi d'uso come la ricerca semantica o conversazionale. Puoi specificare un modello di workflow quando chiami l'API Create Workflow.

- Per utilizzare un modello di workflow fornito dal OpenSearch servizio, specifica il caso d'uso del modello come parametro di `use_case` query.
- Per utilizzare un modello di workflow personalizzato, fornisci il modello completo nel corpo della richiesta. Per un esempio di modello personalizzato, vedi un esempio di modello JSON o un modello YAML di esempio.

Casi d'uso del modello

Questa tabella fornisce una panoramica dei diversi modelli disponibili, una descrizione dei modelli e i parametri richiesti.

Caso d'uso del modello	Description	Parametri obbligatori
`bedrock_titan_embedding_model_deploy`	Crea e distribuisce un modello di incorporamento Amazon Bedrock (per impostazione predefinita, `titan-embed-text-v1`	`create_connector.credential.roleArn`
`bedrock_titan_embedding_model_deploy`	Crea e distribuisce un modello di incorporamento multimodale Amazon Bedrock (per impostazione predefinita, `titan-embed-text-v1`	`create_connector.credential.roleArn`
`cohere_embedding_model_deploy`	Crea e distribuisce un modello di incorporamento Cohere (per impostazione predefinita, 3.0). embed-english-v	`create_connector.credential.roleArn` , `create_connector.credential.secretArn`
`cohere_chat_model_deploy`	Crea e distribuisce un modello di chat Cohere (per impostazione predefinita, Cohere Command).	`create_connector.credential.roleArn` , `create_connector.credential.secretArn`
`open_ai_embedding_model_deploy`	Crea e distribuisce un modello di incorporamento OpenAI (per impostazione predefinita text-embedding-ada, -002).	`create_connector.credential.roleArn` , `create_connector.credential.secretArn`
`openai_chat_model_deploy`	Crea e distribuisce un modello di chat OpenAI (per impostazione predefinita, gpt-3.5-turbo).	`create_connector.credential.roleArn` , `create_connector.credential.secretArn`

Caso d'uso del modello	Description	Parametri obbligatori
`semantic_search_with_cohere_embedding`	Configura la ricerca semantica e implementa un modello di incorporamento Cohere. È necessario fornire la chiave API per il modello Cohere.	`create_connector.credential.roleArn` , `create_connector.credential.secretArn`
`semantic_search_with_cohere_embedding_query_enricher`	Configura la ricerca semantica e implementa un modello di incorporamento Cohere. Aggiunge un processore di ricerca query_enricher che imposta un ID modello predefinito per le query neurali. È necessario fornire la chiave API per il modello Cohere.	`create_connector.credential.roleArn` , `create_connector.credential.secretArn`
`multimodal_search_with_bedrock_titan`	Implementa un modello multimodale Amazon Bedrock e configura una pipeline di ingestione con un processore text_image_embedding e un indice k-NN per la ricerca multimodale. Devi AWS fornire le tue credenziali.	`create_connector.credential.roleArn`

> **Note**
>
> Per tutti i modelli che richiedono un ARN segreto, l'impostazione predefinita prevede l'archiviazione del segreto con il nome chiave «key» in AWS Secrets Manager.

Modelli predefiniti con modelli preaddestrati

Amazon OpenSearch Service offre due modelli di flusso di lavoro predefiniti aggiuntivi non disponibili nel servizio opensource OpenSearch .

Caso d'uso del modello	Description
`semantic_search_with_local_model`	Configura la ricerca semantica e distribuisce un modello preaddestrato (). `msmarco-d istilbert-base-tas-b` Aggiunge un processore `neural_query_enricher` di ricerca che imposta un ID modello predefini to per le query neurali e crea un indice k-NN collegato chiamato ". my-nlp-index
`hybrid_search_with_local_model`	Configura la ricerca ibrida e implementa un modello preaddestrato (). `msmarco-d istilbert-base-tas-b` Aggiunge un processore `neural_query_enricher` di ricerca che imposta un ID modello predefini to per le query neurali e crea un indice k-NN collegato chiamato ". my-nlp-index

Configurazione delle autorizzazioni

Se crei un nuovo dominio con la versione 2.13 o successiva, le autorizzazioni sono già disponibili. Se abiliti il framework di flusso su un dominio di OpenSearch servizio preesistente con la versione 2.11 o precedente e poi esegui l'aggiornamento alla versione 2.13 o successiva, devi definire il ruolo. `flow_framework_manager` Gli utenti senza privilegi di amministratore devono essere mappati a questo ruolo in modo da gestire gli indici a caldo sui domini che utilizzano il controllo granulare degli accessi. Per creare manualmente il ruolo `flow_framework_manager`, procedere nel seguente modo:

1. In OpenSearch Dashboards, vai su Sicurezza e scegli Autorizzazioni.
2. Scegliere Crea gruppo di operazioni e configurare i seguenti gruppi:

Group name (Nome gruppo)	Permissions
`flow_framework_full_access`	• `cluster:admin/opensearch/flow_framework/*` • `cluster_monitor`
`flow_framework_read_accesss`	• `cluster:admin/opensearch/flow_framework/workflow/get` • `cluster:admin/opensearch/flow_framework/workflow/search` • `cluster:admin/opensearch/flow_framework/workflow_state/get` • `cluster:admin/opensearch/flow_framework/workflow_state/search`

3. Scegliere Ruoli, quindi selezionare Crea ruolo.
4. Assegna un nome al ruolo flow_framework_manager.
5. Per Autorizzazioni cluster, selezionare `flow_framework_full_access` e `flow_framework_read_access`.
6. Per Indice, digitare `*`.
7. Per Autorizzazioni indice, selezionare `indices:admin/aliases/get`, `indices:admin/mappings/get` e `indices_monitor`.
8. Scegli Create (Crea).
9. Dopo aver creato il ruolo, associalo a qualsiasi ruolo utente o di backend che gestirà gli indici del framework di flusso.

Osservabilità in Amazon Service OpenSearch

L'osservabilità è la pratica di comprendere lo stato interno e le prestazioni di sistemi complessi esaminandone i risultati. Il monitoraggio tradizionale indica che un sistema non funziona; l'osservabilità aiuta a capirne il motivo, consentendoti di porre nuove domande sui dati di telemetria.

Cosa offre Amazon OpenSearch Service

Amazon OpenSearch Service fornisce una soluzione di osservabilità unificata raccogliendo, correlando e visualizzando tre tipi di dati di telemetria:

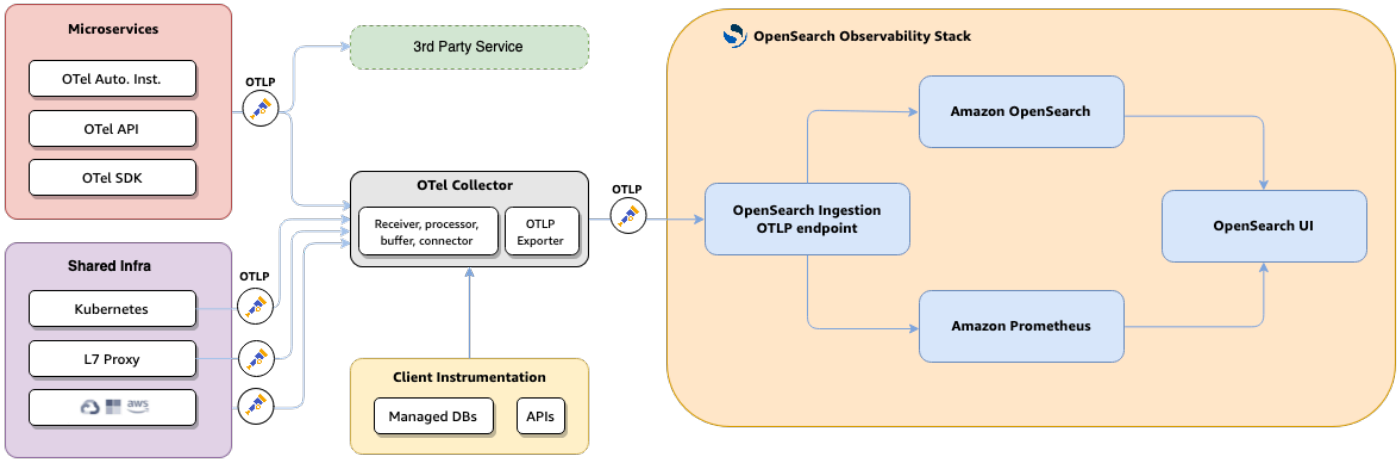
- Registri: registrazioni degli eventi con data e ora.
- Tracce: End-to-end percorso delle richieste attraverso servizi distribuiti.
- Metriche: serie temporali di dati che rappresentano lo stato del sistema, tramite l'integrazione diretta delle query di Amazon Managed Service for Prometheus.

Riunendoli in un'unica interfaccia, Amazon OpenSearch Service aiuta i team operativi e gli sviluppatori a rilevare, diagnosticare e risolvere i problemi più rapidamente. SREs

L'approccio OpenSearch di Amazon Service all'osservabilità

Amazon OpenSearch Service si differenzia in tre modi principali:

- OpenTelemetry-nativo con OpenSearch Ingestion come ultimo miglio: standardizza la strumentazione e la raccolta. OTel Amazon OpenSearch Ingestion funge da pipeline completamente gestita che filtra, arricchisce, trasforma e indirizza i dati di telemetria prima dell'indicizzazione.
- Registri, tracce e metriche unificati nell'interfaccia utente: analizza tutti e tre i tipi di segnale da un'unica area di lavoro di osservabilità OpenSearch . Correla una traccia lenta ai registri delle applicazioni o sovrapponi le metriche di Prometheus alle dashboard dei tuoi servizi.
- Linguaggi di interrogazione orientati allo scopo: utilizza il Piped Processing Language (PPL) per log e tracce e PromQL per le metriche. Ogni linguaggio è ottimizzato per il tipo di segnale corrispondente, offrendo interrogazioni espressive senza compromessi.



> **Note**
>
> Le funzionalità di osservabilità descritte in questa sezione sono disponibili solo nell'interfaccia utente. OpenSearch Non sono disponibili nei OpenSearch dashboard. Per i nuovi carichi di lavoro di osservabilità, consigliamo di configurare uno spazio di lavoro OpenSearch Interfaccia utente con osservabilità.

Nozioni di base

Attiva il tuo stack di osservabilità e inizia a inviare dati di telemetria in pochi minuti AWS .

Avvio rapido

Il modo più veloce per implementare uno stack di end-to-end osservabilità è il programma di installazione AWS CLI. Crea le seguenti risorse:

- Un dominio Amazon OpenSearch Service
- Un Amazon Managed Service per l'area di lavoro Prometheus
- Una pipeline di Amazon OpenSearch Ingestion
- Un'applicazione OpenSearch UI con un'area di lavoro di osservabilità

Facoltativamente, il programma di installazione avvia un'istanza EC2 con la Demo per la telemetria di esempio. OpenTelemetry

Esegui il seguente comando per avviare l'installazione:

```
bash -c "$(curl -fsSL https://raw.githubusercontent.com/opensearch-project/
observability-stack/main/install.sh)" -- --deployment-target=aws
```

L'installazione richiede circa 15 minuti.

Implementazione CDK

Per infrastructure-as-code, usa AWS CDK. La distribuzione CDK crea due stack:

Pila	Cosa crea	Orario di distribuzione
ObsInfra	OpenSearch dominio, spazio di lavoro Amazon Managed Service for Prometheus, origine dati di query diretta, ruolo IAM della pipeline	~17 min
ObservabilityStack	Mappatura dettagliata del controllo degli accessi, pipeline di inserimento, applicazione OpenSearch UI, OpenSearch inizializzazione del dashboard, carico di lavoro dimostrativo (opzionale)	~6 min

Esegui i seguenti comandi per la distribuzione:

```
cd aws/cdk
npm install
cdk deploy --all
```

Per ulteriori informazioni, consultate il file README sulla distribuzione CDK su. GitHub

Invio di telemetria

Entrambi i metodi di implementazione creano un endpoint di OpenSearch ingestione che accetta dati OTLP. Configura il tuo OTel Collector per l'esportazione utilizzando l'autenticazione SigV4:

```
extensions:
```

```
  sigv4auth:
    region: us-west-2
    service: osis

exporters:
  otlphttp/logs:
    logs_endpoint: ${OSIS_ENDPOINT}/v1/logs
    auth: { authenticator: sigv4auth }
    compression: none
  otlphttp/traces:
    traces_endpoint: ${OSIS_ENDPOINT}/v1/traces
    auth: { authenticator: sigv4auth }
    compression: none
  otlphttp/metrics:
    metrics_endpoint: ${OSIS_ENDPOINT}/v1/metrics
    auth: { authenticator: sigv4auth }
    compression: none
```

> **Note**
>
> Il principale IAM che invia i dati richiede l'`aps:RemoteWrite`autorizzazione `osis:Ingest` e l'autorizzazione sull'ARN della pipeline.

Ulteriori informazioni

Utilizza le seguenti risorse per saperne di più sull'invio di dati di telemetria:

- OpenTelemetry guide alla strumentazione (per lingua)
- Monitoraggio dell'infrastruttura (Docker AWS, Kubernetes, Prometheus)
- OTel Configurazione del collettore
- Pipeline e dosaggio dei dati
- OpenSearch Ingestione di Amazon in questa guida

Acquisizione della telemetria delle applicazioni

Per utilizzare le funzionalità di osservabilità in Amazon OpenSearch Service, devi importare tracce, log e metriche delle applicazioni. Questa pagina descrive la configurazione delle pipeline

OpenTelemetry Collector e OpenSearch Ingestion per elaborare e instradare i dati di telemetria verso Amazon Managed Service for Prometheus. OpenSearch

OpenTelemetry Configurazione del Collector

Il OpenTelemetry (OTel) Collector è il punto di ingresso per tutta la telemetria delle applicazioni. Riceve dati tramite OTLP e indirizza le tracce e i log verso OpenSearch Ingestion mentre invia le metriche a Prometheus.

È possibile configurare Collector utilizzando uno dei seguenti approcci: OTel

OTel Collector esporta tracce e log su un endpoint OpenSearch Ingestion utilizzando l'autenticazione SigV4 e i parametri su Amazon Managed Service for Prometheus utilizzando l'esportatore di scrittura remoto Prometheus. OpenSearch Ingestion gestisce l' OpenSearchelaborazione, l'arricchimento e il routing verso.

OTel Configurazione del collettore con Ingestion OpenSearch

La seguente configurazione di esempio utilizza l'autenticazione SigV4 per esportare tracce e log su un endpoint OpenSearch Ingestion e le metriche su Prometheus:

```
extensions:
  sigv4auth:
    region: us-west-2
    service: osis

exporters:
  otlphttp/osis-traces:
    traces_endpoint: ${OSIS_ENDPOINT}/v1/traces
    auth: { authenticator: sigv4auth }
    compression: none
  otlphttp/osis-logs:
    logs_endpoint: ${OSIS_ENDPOINT}/v1/logs
    auth: { authenticator: sigv4auth }
    compression: none
  # Amazon Managed Service for Prometheus via Prometheus Remote Write with SigV4 auth
  prometheusremotewrite/amp:
    endpoint: "https://aps-workspaces.region.amazonaws.com/workspaces/workspace-id/api/
v1/remote_write"
    auth:
      authenticator: sigv4auth
```

```
service:
  extensions: [sigv4auth]
  pipelines:
    traces:
      receivers: [otlp]
      processors: [batch]
      exporters: [otlphttp/osis-traces]
    logs:
      receivers: [otlp]
      processors: [batch]
      exporters: [otlphttp/osis-logs]
    metrics:
      receivers: [otlp]
      processors: [batch]
      exporters: [prometheusremotewrite/amp]
```

> **Note**
>
> Il principale IAM che invia i dati richiede l'`aps:RemoteWrite`autorizzazione `osis:Ingest` e l'autorizzazione sull'ARN della pipeline.

Configurazione delle pipeline OpenSearch di ingestione

OpenSearch Ingestion (o Data Prepper autogestito) riceve la telemetria dal Collector e la elabora per il OTel monitoraggio delle prestazioni delle applicazioni (APM).

Architettura della pipeline

La pipeline elabora i dati di telemetria nelle seguenti fasi:

1. La pipeline di ingresso riceve tutta la telemetria e indirizza i registri e le tracce verso sotto-pipeline separate.
2. La pipeline di log scrive i dati di registro utilizzando il tipo di indice. OpenSearch `log-analytics-plain`
3. La pipeline di traccia distribuisce gli intervalli alla pipeline di archiviazione raw e alla pipeline della mappa dei servizi.
4. La pipeline di traccia non elaborata elabora i dati relativi al `otel_traces` processore e li memorizza nel tipo di indice. `trace-analytics-plain-raw`

5. La pipeline delle mappe dei servizi utilizza il `otel_apm_service_map` processore per generare la topologia e le metriche RED (Rate, Errors, Duration). Scrive da OpenSearch e verso Prometheus tramite scrittura remota.

Configurazione della pipeline

L'esempio seguente mostra una configurazione completa della pipeline per OpenSearch Ingestion che copre tutti i tipi di segnali di osservabilità: log, tracce e metriche. Puoi includere tutte le pipeline o solo quelle pertinenti al tuo caso d'uso. Sostituisci i *`placeholder`* valori con le tue informazioni.

```
version: '2'

# Main OTLP pipeline - receives all telemetry and routes by signal type
otlp-pipeline:
  source:
    otlp:
      logs_path: '/pipeline-name/v1/logs'
      traces_path: '/pipeline-name/v1/traces'
      metrics_path: '/pipeline-name/v1/metrics'
  route:
    - logs: 'getEventType() == "LOG"'
    - traces: 'getEventType() == "TRACE"'
    - metrics: 'getEventType() == "METRIC"'
  processor: []
  sink:
    - pipeline:
        name: otel-logs-pipeline
        routes:
          - logs
    - pipeline:
        name: otel-traces-pipeline
        routes:
          - traces
    - pipeline:
        name: otel-metrics-pipeline
        routes:
          - metrics

# Log processing pipeline
otel-logs-pipeline:
  source:
    pipeline:
```

```
      name: otlp-pipeline
  processor:
    - copy_values:
        entries:
          - from_key: "time"
            to_key: "@timestamp"
  sink:
    - opensearch:
        hosts:
          - 'https://opensearch-endpoint'
        index_type: log-analytics-plain
        aws:
          serverless: false
          region: 'region'
          sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"

# Trace fan-out pipeline
otel-traces-pipeline:
  source:
    pipeline:
      name: otlp-pipeline
  processor: []
  sink:
    - pipeline:
        name: traces-raw-pipeline
        routes: []
    - pipeline:
        name: service-map-pipeline
        routes: []

# Raw trace storage pipeline
traces-raw-pipeline:
  source:
    pipeline:
      name: otel-traces-pipeline
  processor:
    - otel_traces:
  sink:
    - opensearch:
        hosts:
          - 'https://opensearch-endpoint'
        index_type: trace-analytics-plain-raw
        aws:
          serverless: false
```

```
          region: 'region'
          sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"

# Service map generation pipeline (APM)
service-map-pipeline:
  source:
    pipeline:
      name: otel-traces-pipeline
  processor:
    - otel_apm_service_map:
        group_by_attributes:
          - telemetry.sdk.language # Add any resource attribute to group by
        window_duration: 30s
  route:
    - otel_apm_service_map_route: 'getEventType() == "SERVICE_MAP"'
    - service_processed_metrics: 'getEventType() == "METRIC"'
  sink:
    - opensearch:
        hosts:
          - 'https://opensearch-endpoint'
        aws:
          serverless: false
          region: 'region'
          sts_role_arn: "arn:aws:iam::account-id:role/pipeline-role"
        routes:
          - otel_apm_service_map_route
        index_type: otel-v2-apm-service-map
    - prometheus:
        url: 'https://aps-workspaces.region.amazonaws.com/workspaces/workspace-id/api/
v1/remote_write'
        aws:
          region: 'region'
        routes:
          - service_processed_metrics

# Metrics processing pipeline
otel-metrics-pipeline:
  source:
    pipeline:
      name: otlp-pipeline
  processor:
    - otel_metrics:
  sink:
    - prometheus:
```

```
        url: 'https://aps-workspaces.region.amazonaws.com/workspaces/workspace-id/api/
v1/remote_write'
        aws:
          region: 'region'
```

Verifica dell'ingestione

Dopo aver configurato OTel Collector e le pipeline, verifica che i dati di telemetria scorrano correttamente.

- Verifica OpenSearch gli indici: verifica che nel tuo dominio siano presenti i seguenti indici:, e. `otel-v1-apm-span-*` `otel-v2-apm-service-map` `logs-otel-v1-*`
- Verifica degli obiettivi Prometheus: verifica che il target di scrittura remoto Prometheus riceva le metriche dalla pipeline della mappa dei servizi.
- Verifica nell' OpenSearch interfaccia utente: accedi a Observability, quindi Application Monitoring per confermare che i tuoi servizi vengano visualizzati.

Fasi successive

Dopo aver verificato che i dati di telemetria siano stati acquisiti, esplora i seguenti argomenti:

- the section called “Monitoraggio dell'applicazione”— Monitora lo stato delle applicazioni con mappe dei servizi e metriche RED.
- Scopri le tracce: scopri e analizza le tracce distribuite.
- Scopri i log: scopri e interroga i log delle applicazioni.
- Scopri le metriche: scopri e interroga le metriche di Prometheus utilizzando PromQL.

Set di dati

I set di dati sono raccolte di indici che rappresentano un raggruppamento logico dei dati di osservabilità. I set di dati vengono utilizzati per organizzare i log e le tracce dei dati in modo da poter interrogare e analizzare insieme gli indici correlati nell'esperienza Discover. Ogni set di dati è mappato su uno o più indici nel dominio del OpenSearch servizio e definisce il tipo di dati, il campo temporale e il linguaggio di query per la pagina Discover.

Tipi di set di dati

La tabella seguente descrive i tipi di set di dati che è possibile creare.

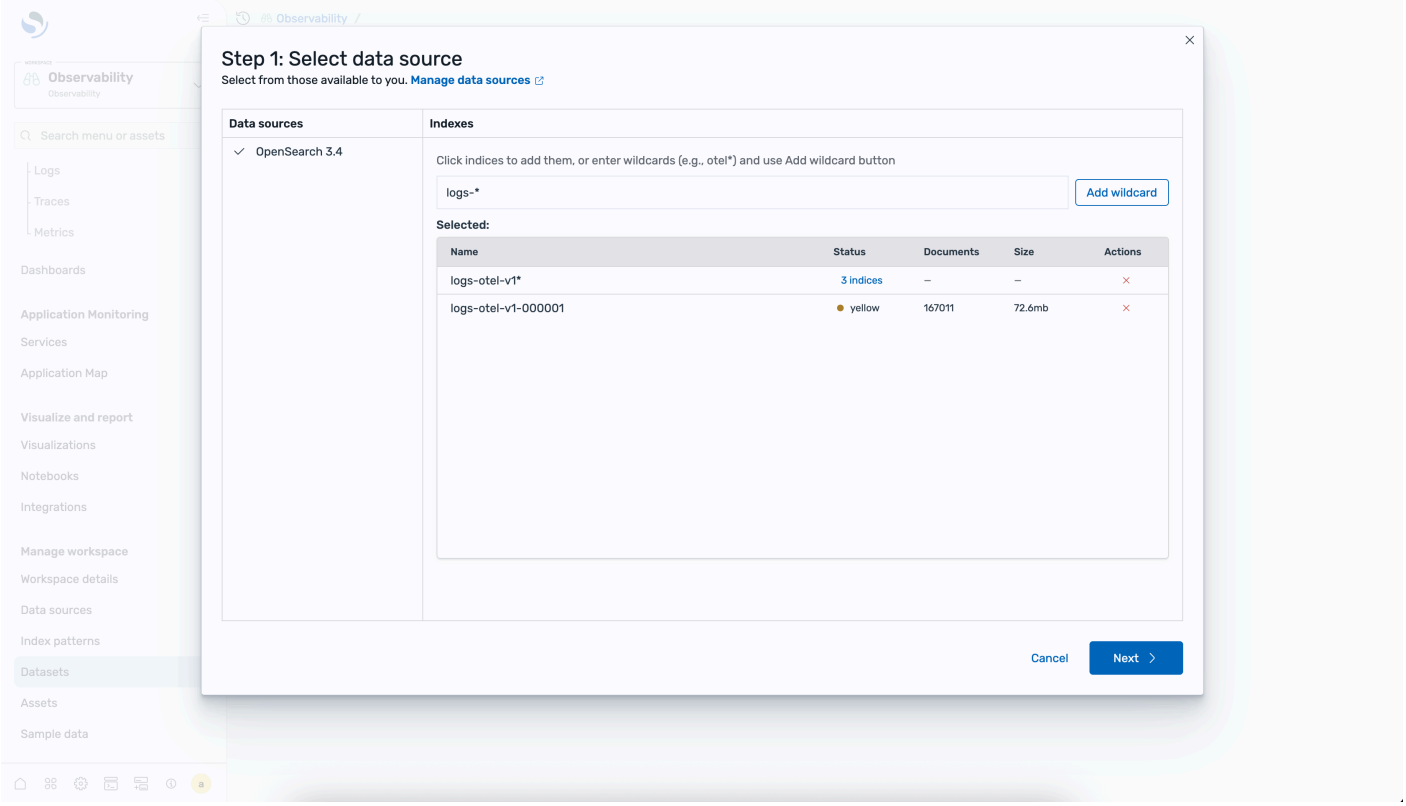
Tipo	Description	Linguaggio di query
Log	Raggruppa uno o più indici di log per l'interrogazione e la visualizzazione nella pagina Discover Logs.	PPL
Tracce	Raggruppa gli indici trace span per l'interrogazione e la visualizzazione nella pagina Discover Traces.	PPL

> **Note**
>
> Le metriche non richiedono un set di dati perché i dati metrici non sono archiviati in. OpenSearch Le metriche vengono richieste direttamente da Amazon Managed Service for Prometheus utilizzando PromQL.

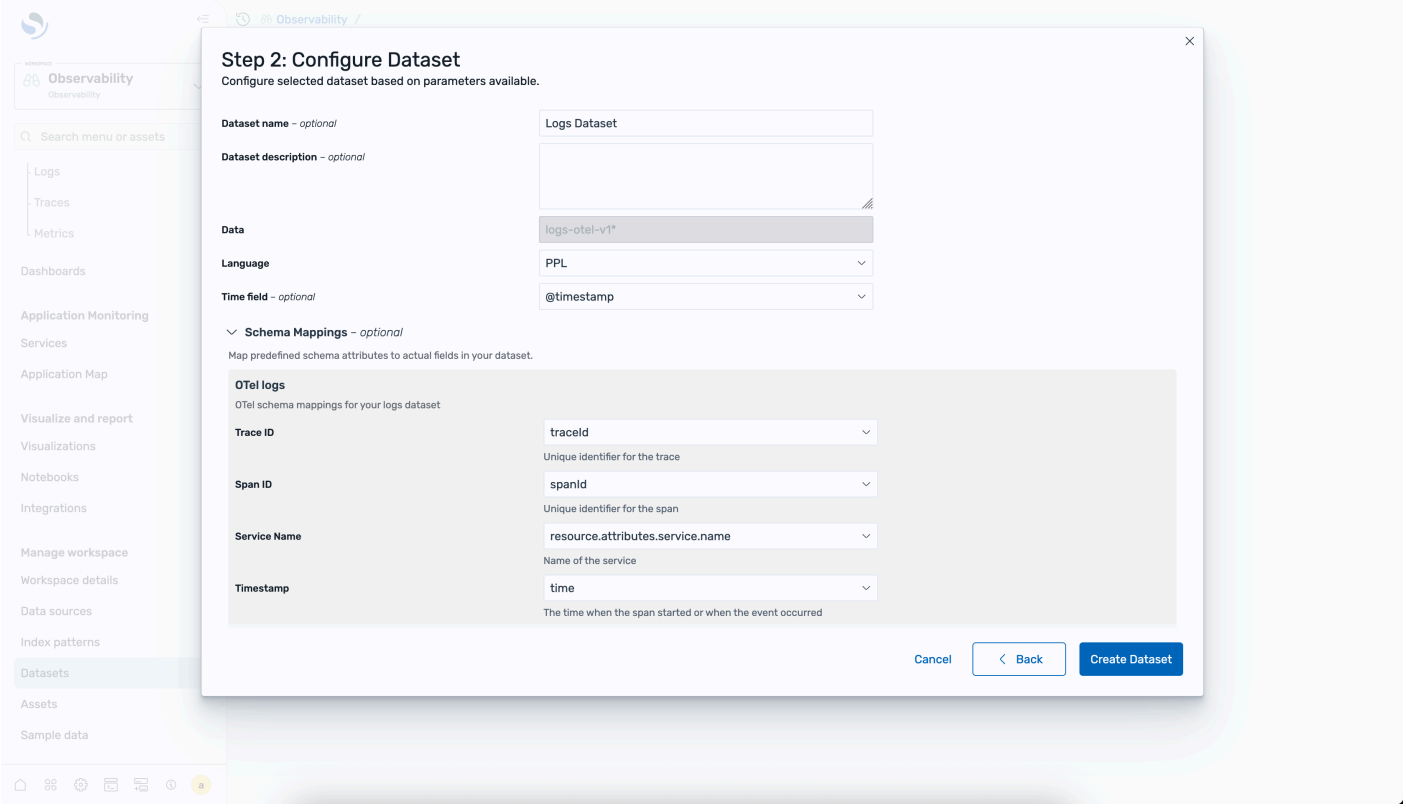
Per creare un set di dati di log

Completa i seguenti passaggi per creare un set di dati di log nell'interfaccia utente. OpenSearch

1. Nel tuo spazio di lavoro di osservabilità, espandi Discover nella barra di navigazione a sinistra e scegli Logs.
2. Scegli Crea set di dati.
3. Seleziona una fonte di dati dall'elenco delle connessioni di servizio disponibili OpenSearch .



4. Configura il set di dati inserendo un nome, selezionando l'indice e specificando il campo timestamp.

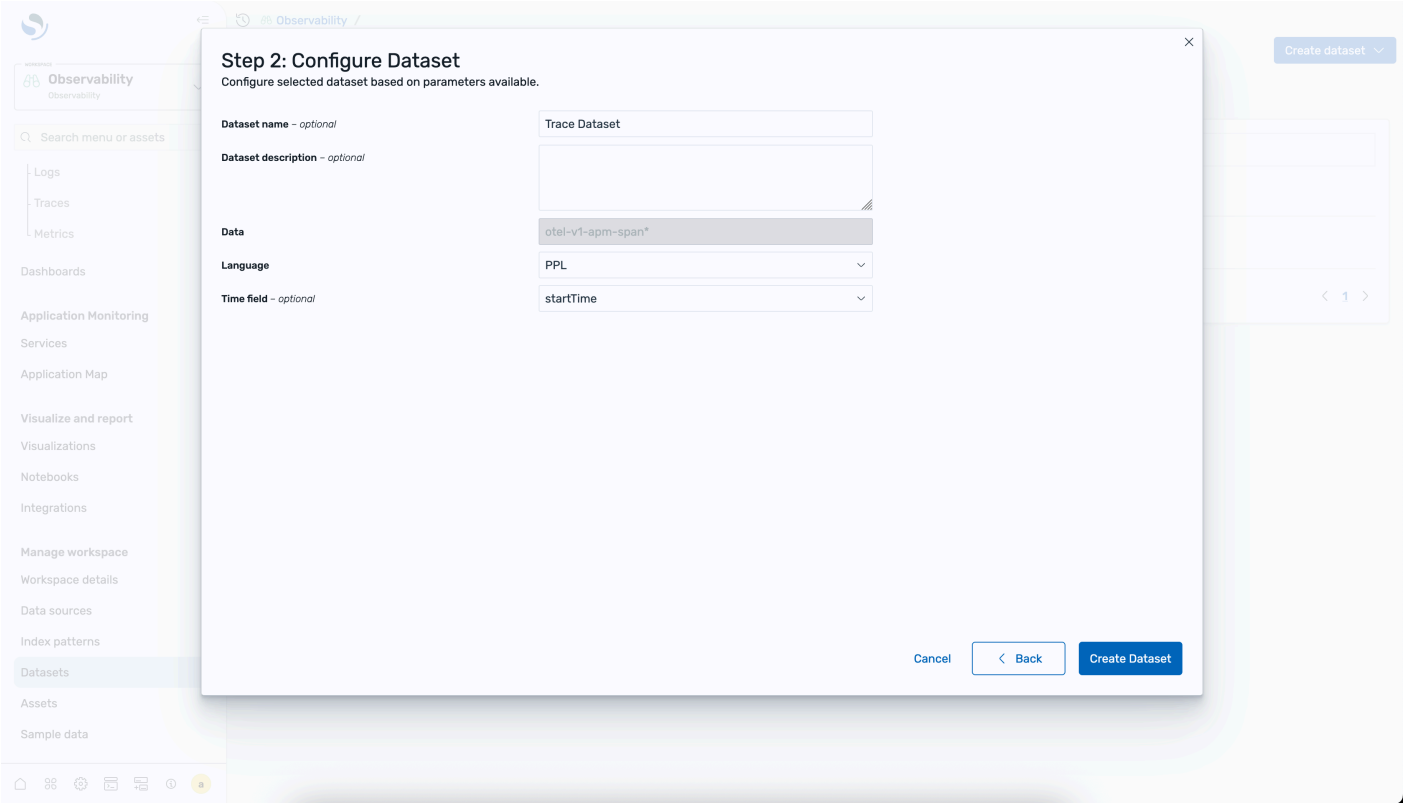


5. Scegli Crea set di dati per salvare la configurazione.

Per creare un set di dati Traces

Completa i seguenti passaggi per creare un set di dati di tracce nell' OpenSearch interfaccia utente.

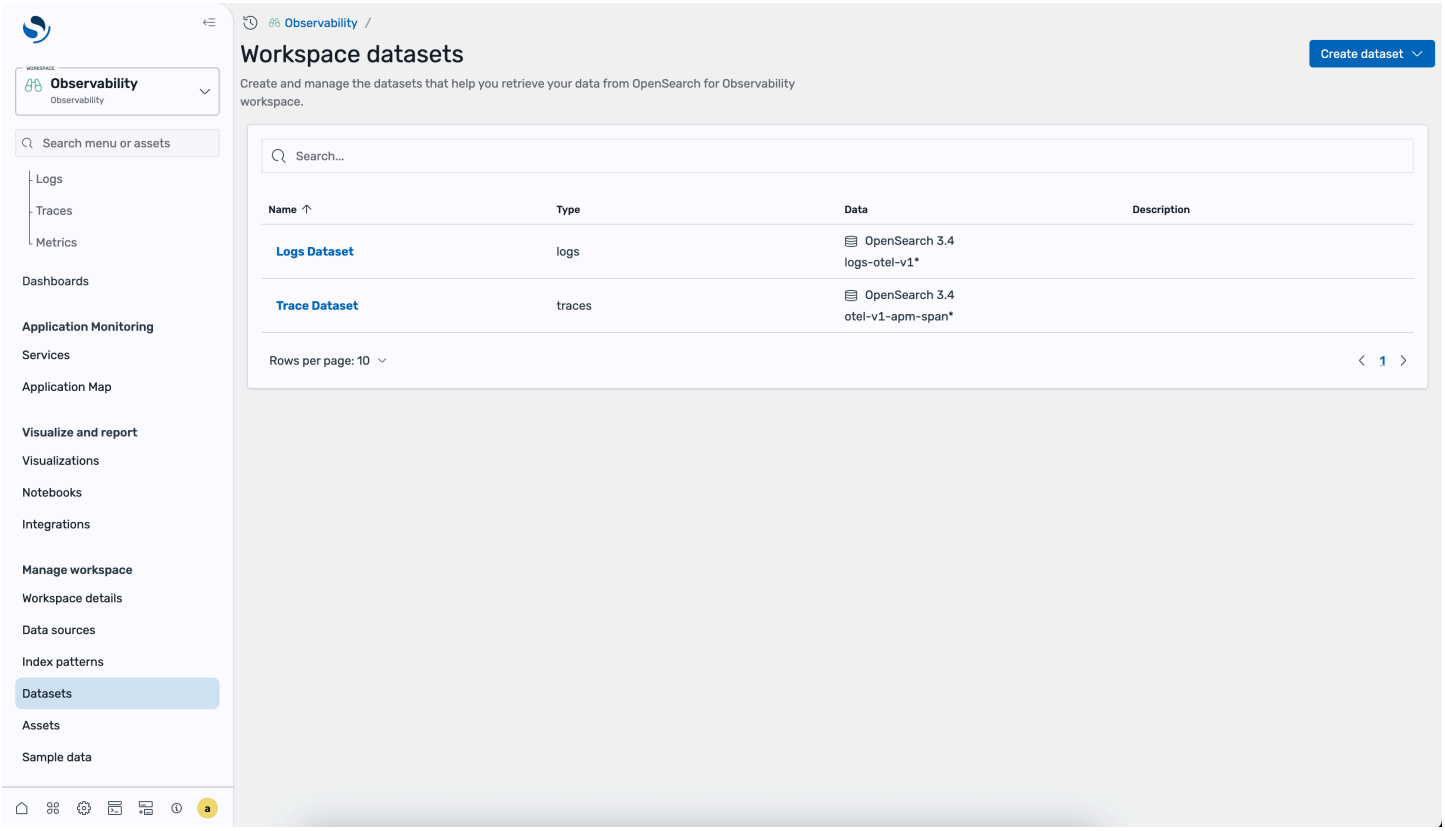
1. Nel tuo spazio di lavoro sull'osservabilità, espandi Discover nella barra di navigazione a sinistra e scegli Traces.
2. Scegli Crea set di dati.
3. Seleziona una fonte di dati dall'elenco delle connessioni di OpenSearch servizio disponibili.
4. Configura il set di dati inserendo un nome, selezionando l'indice span e specificando il campo timestamp.



5. Scegli Crea set di dati per salvare la configurazione.

Per visualizzare i set di dati

Puoi visualizzare tutti i set di dati configurati dal selettore di set di dati nella pagina Discover Logs o Discover Traces. L'elenco dei set di dati mostra il nome, il tipo, l'origine dati e il campo timestamp di ogni set di dati.



Analisi dei set di dati in Discover

Dopo aver creato un set di dati, puoi analizzarlo nella pagina Discover corrispondente.

Log

Seleziona un set di dati di log dal selettore di set di dati nella pagina Discover Logs per interrogare e visualizzare i dati di registro utilizzando PPL. Per ulteriori informazioni, consulta the section called “Scopri i log”.

Tracce

Seleziona un set di dati di tracce dal selettore del set di dati nella pagina Discover Traces per esplorare gli intervalli di traccia, visualizzare le metriche RED e approfondire le singole tracce. Per ulteriori informazioni, consulta the section called “Scopri le tracce”.

Scopri i log

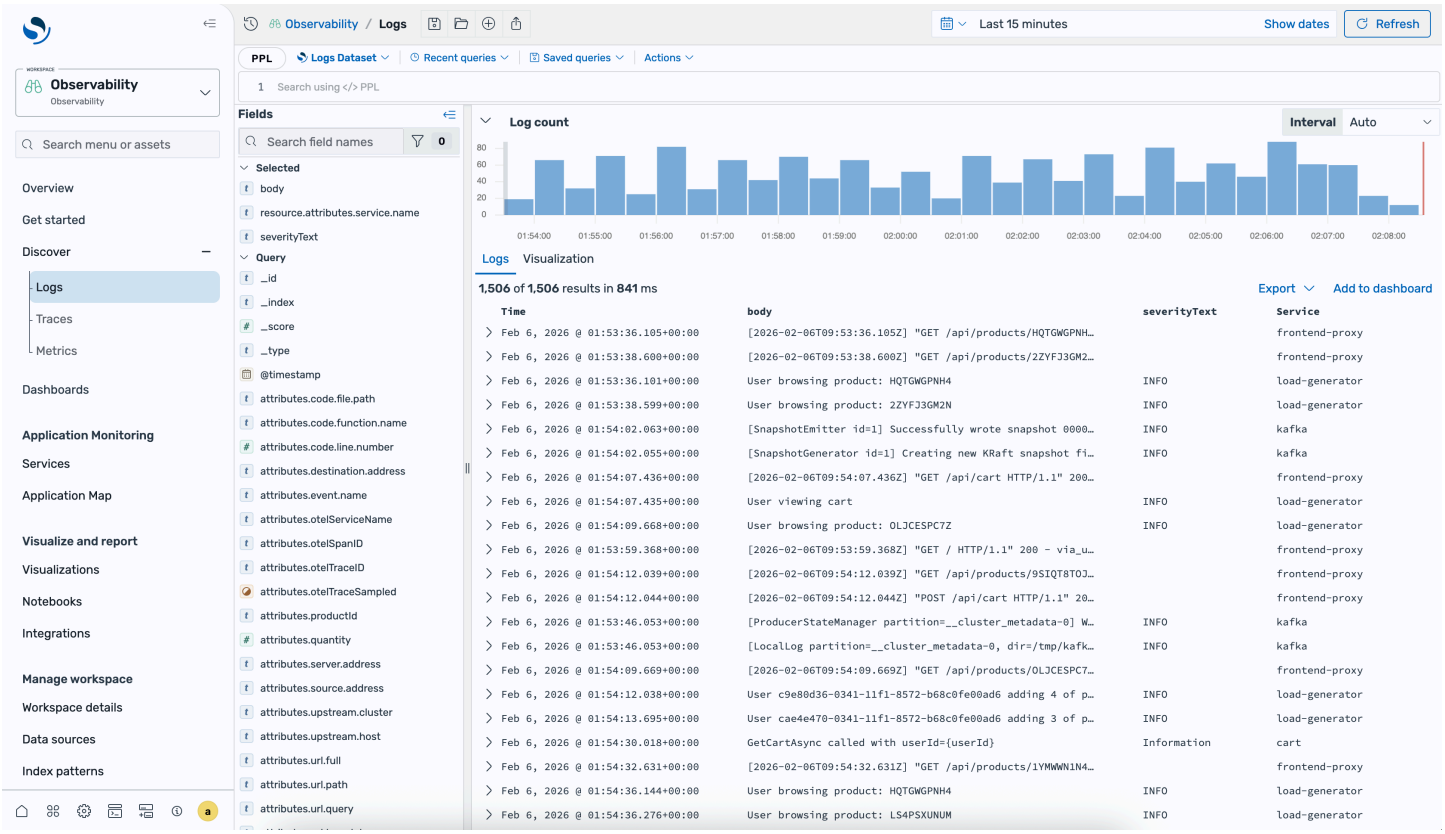
La pagina Discover Logs fornisce un'interfaccia dedicata per esplorare e analizzare i dati di registro nell'area di lavoro di osservabilità del OpenSearch servizio. Puoi scrivere query PPL per filtrare e aggregare i dati di registro, creare visualizzazioni direttamente dai risultati delle query e aggiungere tali visualizzazioni ai dashboard. La pagina fornisce anche assistenza per le interrogazioni in linguaggio naturale fornita dall'assistente AI. OpenSearch

Per accedere alla pagina dei registri

Nel tuo spazio di lavoro di osservabilità, espandi Discover nella barra di navigazione a sinistra e scegli Logs.

Esplorazione dei dati di registro

L'interfaccia Discover Logs fornisce i seguenti componenti per l'esplorazione dei dati di registro.



- Selettore del set di dati: scegli il set di dati dei log su cui desideri interrogare. Ogni set di dati è mappato su uno o più indici nel dominio del servizio. OpenSearch

- Editor di query: scrivi query PPL per filtrare, aggregare e trasformare i dati di registro. L'editor fornisce suggerimenti per il completamento automatico e l'evidenziazione della sintassi.
- Filtro temporale: specifica l'intervallo di tempo per i risultati della query. È possibile scegliere un intervallo relativo o specificare orari di inizio e fine assoluti.
- Pannello dei risultati: visualizza i risultati della query come tabella di eventi di registro. È possibile espandere i singoli eventi per visualizzare tutti i campi.
- Istogramma: visualizza la distribuzione degli eventi di registro nel tempo. L'istogramma si aggiorna automaticamente in base alla query e al filtro temporale.
- Pannello Campi: sfoglia i campi disponibili nel set di dati e aggiungili come colonne alla tabella dei risultati.

Interrogazione dei log tramite PPL

Il linguaggio PPL (Piped Processing Language) è un linguaggio di interrogazione che utilizza la sintassi pipe-based () per concatenare i comandi. | È possibile utilizzare PPL per filtrare, aggregare e trasformare i dati di registro.

Interrogazioni di base

Per recuperare tutti gli eventi di registro da un set di dati, usa il comando: `source`

```
source = my-logs-dataset
```

Per limitare il numero di risultati, usa il `head` comando:

```
source = my-logs-dataset | head 20
```

Filtraggio con WHERE

Utilizza la `where` clausola per filtrare gli eventi di registro in base ai valori dei campi:

```
source = my-logs-dataset | where severity_text = 'ERROR'
```

È possibile combinare più condizioni:

```
source = my-logs-dataset |
    where severity_text = 'ERROR' and service_name = 'payment-service'
```

Gestione delle interrogazioni

È possibile salvare le interrogazioni utilizzate di frequente per riutilizzarle. Per salvare un'interrogazione, scegli Salva nella barra degli strumenti dell'editor di query e inserisci un nome per l'interrogazione. Per caricare un'interrogazione salvata, scegliete Apri e selezionate l'interrogazione dall'elenco.

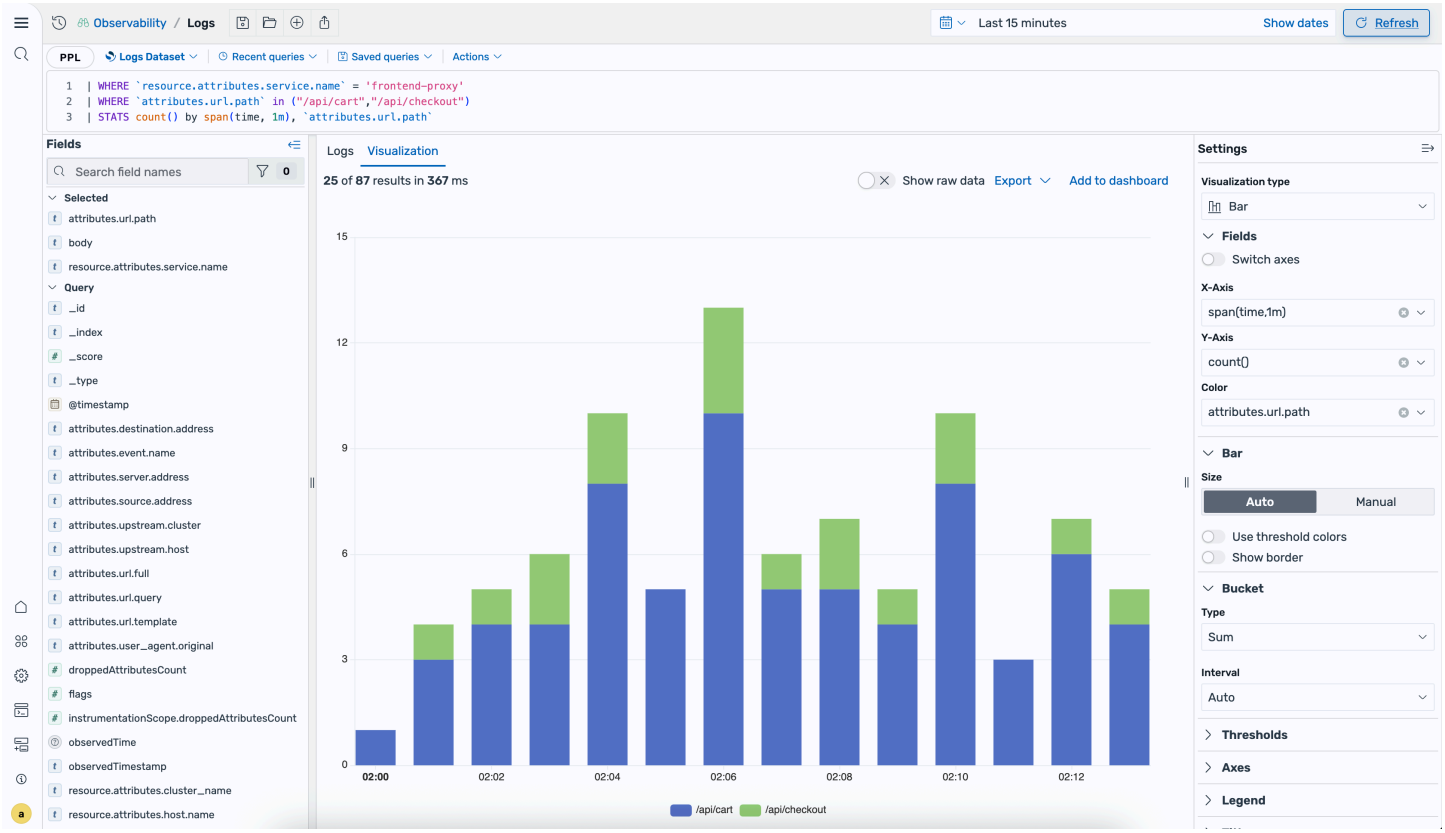
Per l'elenco completo dei comandi e delle funzioni PPL, vedete il riferimento al Piped Processing Language.

Creazione di visualizzazioni dai log

È possibile creare visualizzazioni direttamente dai risultati delle query PPL. Usa il `stats` comando per aggregare i dati per la visualizzazione:

```
source = my-logs-dataset |
    stats count() as error_count by service_name, span(timestamp, 1h)
```

Dopo aver eseguito una `stats` query, scegli la scheda Visualizzazione per visualizzare i risultati sotto forma di grafico.



Tipi di visualizzazione

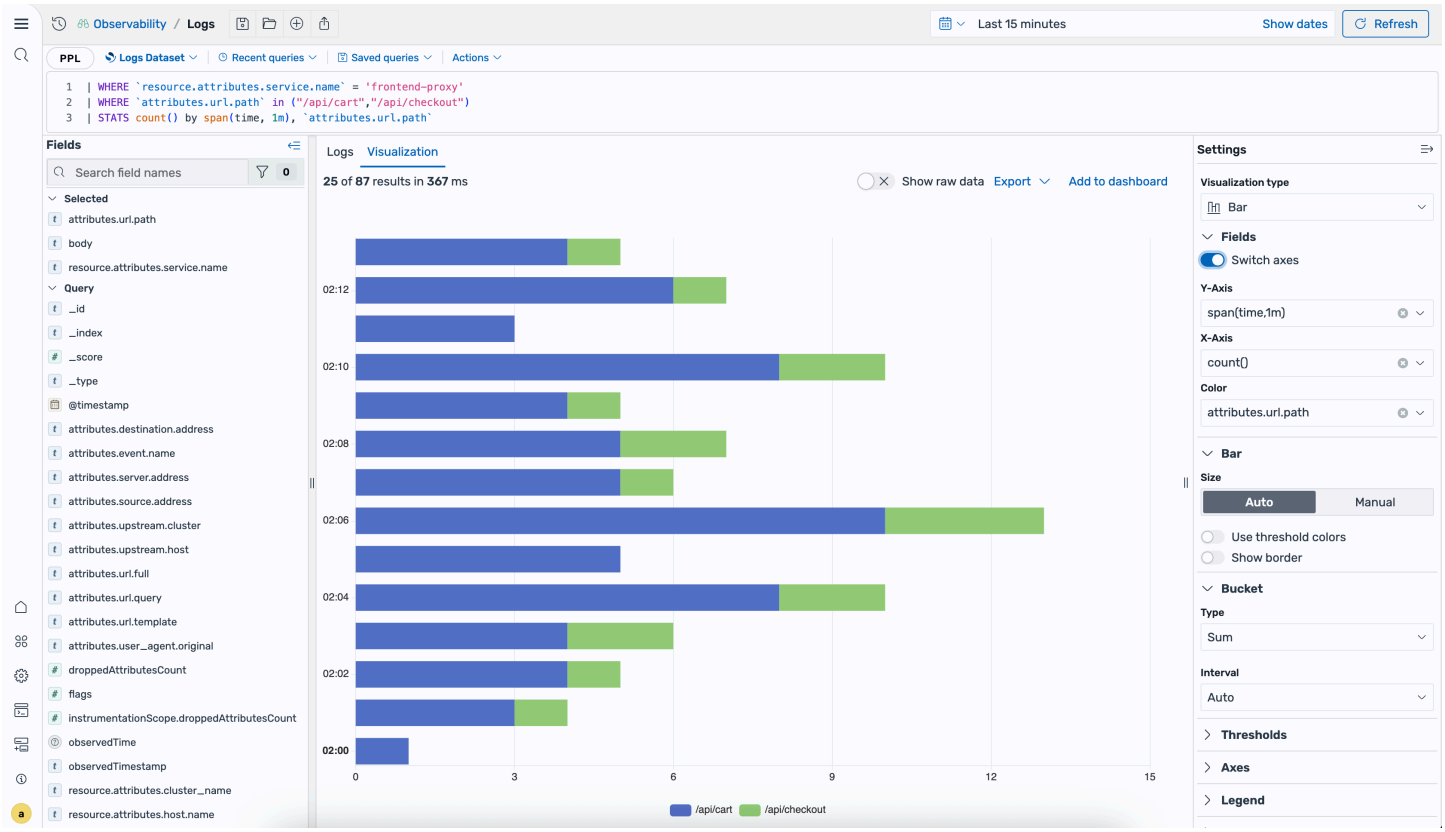
La tabella seguente descrive i tipi di visualizzazione che è possibile utilizzare.

Tipo	Description
Line (Linea)	Visualizza punti dati collegati da linee, utili per mostrare le tendenze nel tempo.
Area	Simile a un grafico a linee con l'area sotto la linea riempita, utile per mostrare il volume nel tempo.
Barra	Visualizza i dati come barre verticali o orizzonta li, utili per confrontare i valori tra le categorie.
Metrica	Visualizza un singolo valore numerico, utile per mostrare gli indicatori chiave delle prestazioni.
Cronologia dello stato	Visualizza i cambiamenti di stato nel tempo sotto forma di bande colorate, utili per monitorare le transizioni di stato.
Mappa termica	Visualizza i dati come una matrice di celle colorate, utile per mostrare densità e motivi.
Manometro a barra	Visualizza un singolo valore come barra piena all'interno di un intervallo, utile per mostrare l'avanzamento verso una soglia.
Torta	Visualizza i dati come sezioni proporzionali di un cerchio, utili per mostrare la composizione.

Impostazioni di visualizzazione

Quando la scheda Visualizzazione è attiva, viene visualizzato un pannello delle impostazioni sul lato destro dello schermo. Usa questo pannello per configurare il tipo di grafico, mappare i campi sugli assi e personalizzare gli stili visivi come colori e legende.

Per cambiare gli assi di una visualizzazione, usa la configurazione degli assi nel pannello delle impostazioni.



Aggiungere visualizzazioni alle dashboard

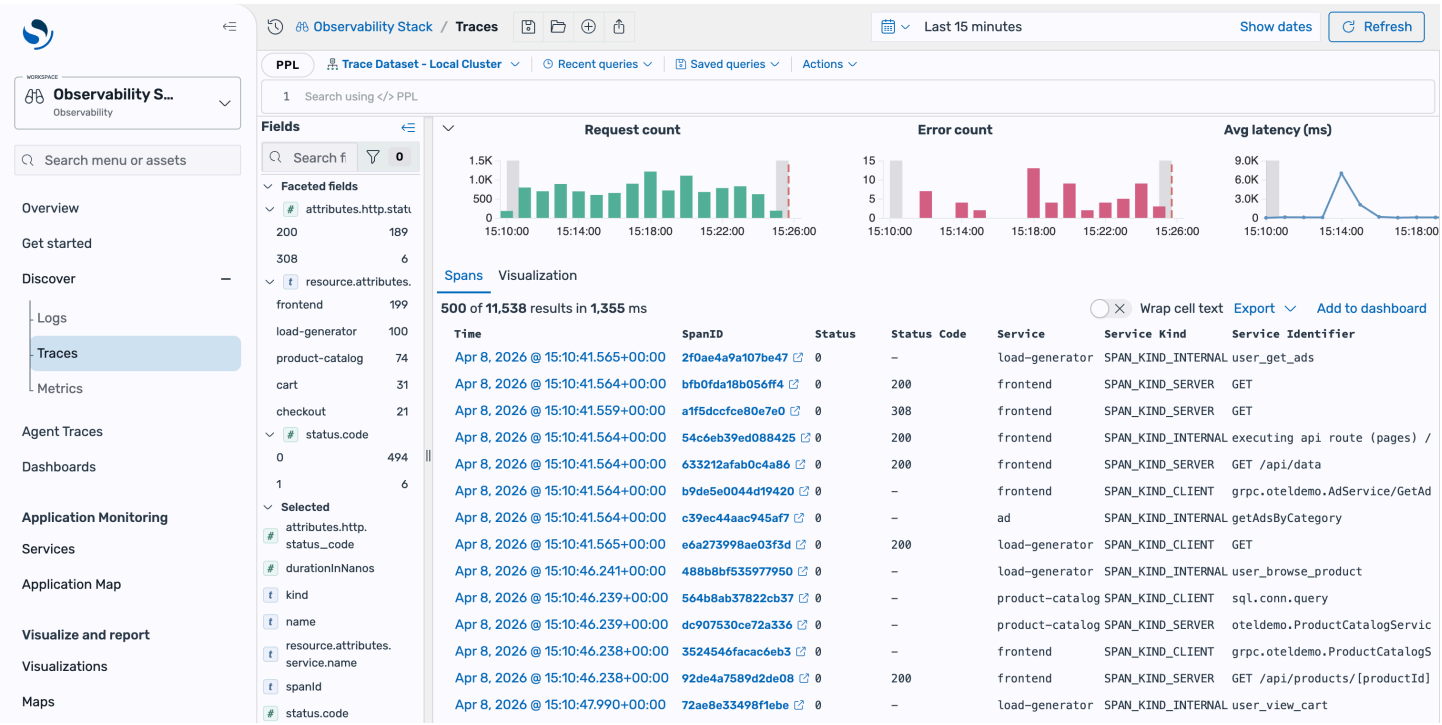
Dopo aver creato una visualizzazione, puoi aggiungerla a una dashboard per un monitoraggio continuo. Scegli Salva nella dashboard nella barra degli strumenti di visualizzazione, quindi seleziona una dashboard esistente o creane una nuova. La visualizzazione viene salvata con la relativa query PPL sottostante in modo che si aggiorni automaticamente all'apertura del pannello di controllo.

Scopri Traces

La pagina Discover Traces fornisce un'interfaccia dedicata per esplorare i dati di traccia distribuiti nell'area di lavoro di osservabilità OpenSearch del Servizio. Puoi visualizzare le metriche RED (frequenza, tasso di errore, durata) per i tuoi servizi, sfogliare gli intervalli di tracciamento con filtri sfaccettati e approfondire i singoli intervalli e le tracce per diagnosticare problemi di prestazioni. La pagina supporta anche la correlazione delle tracce con i dati di registro correlati.

Per accedere alla pagina Traces

Nel tuo spazio di lavoro sull'osservabilità, espandi Discover nella barra di navigazione a sinistra e scegli Traces.

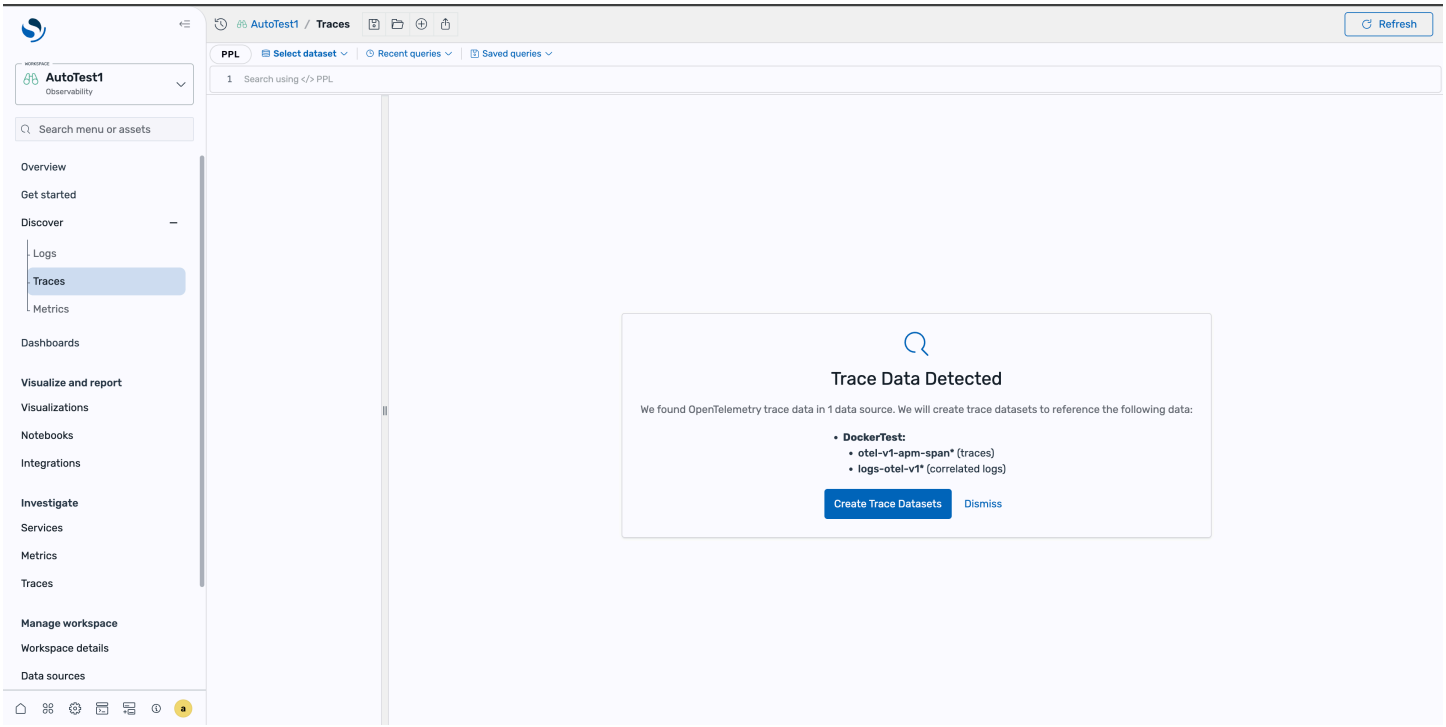


Configurazione dei set di dati di traccia

Prima di poter esplorare i dati di traccia, è necessario configurare un set di dati di trace. È possibile creare un set di dati automaticamente o manualmente.

Creazione automatica di set di dati

Quando accedi alla pagina Discover Traces per la prima volta e nel tuo dominio sono presenti dati di traccia, la pagina ti chiede di creare automaticamente un set di dati. Scegli Crea set di dati per accettare la configurazione predefinita.



Creazione manuale del set di dati

Per creare manualmente un set di dati Traces, segui la procedura riportata di seguito. the section called “Per creare un set di dati Traces” La creazione manuale consente di controllare il modello dell'indice, il campo del timestamp e il nome del set di dati.

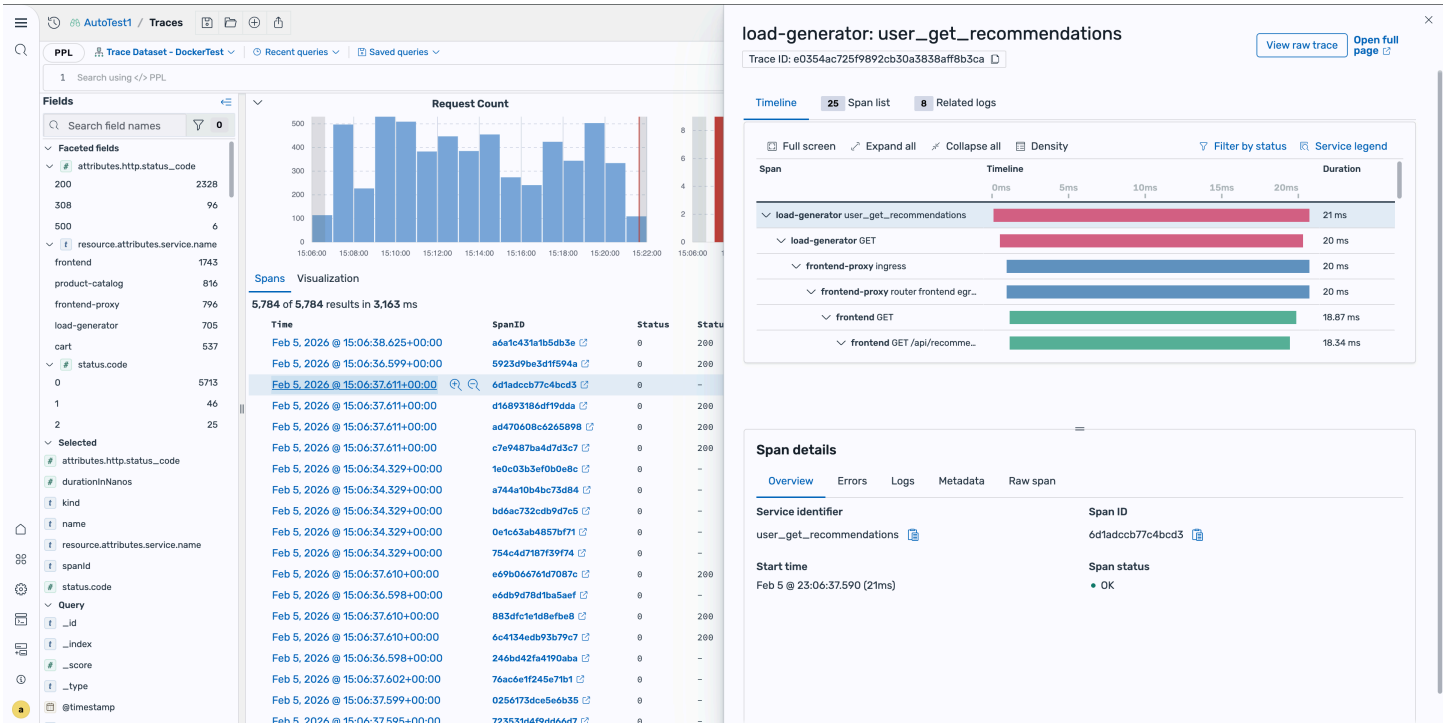
Esplorazione dei dati di traccia

La pagina Discover Traces fornisce i seguenti componenti per esplorare i dati di traccia.

- Metriche RED: frequenza di visualizzazione (richieste al secondo), tasso di errore (percentuale di richieste non riuscite) e durata (percentili di latenza) per il set di dati selezionato. Queste metriche si aggiornano in base al filtro temporale.
- Campi sfaccettati: filtra gli intervalli di traccia per nome del servizio, operazione, codice di stato e altri attributi span. Seleziona i valori nel pannello dei campi sfaccettati per restringere i risultati.
- Tabella Span: sfoglia i singoli intervalli con le colonne per Trace ID, span ID, nome del servizio, operazione, durata e stato. Puoi ordinare in base a qualsiasi colonna ed espandere le righe per visualizzare i dettagli degli intervalli.

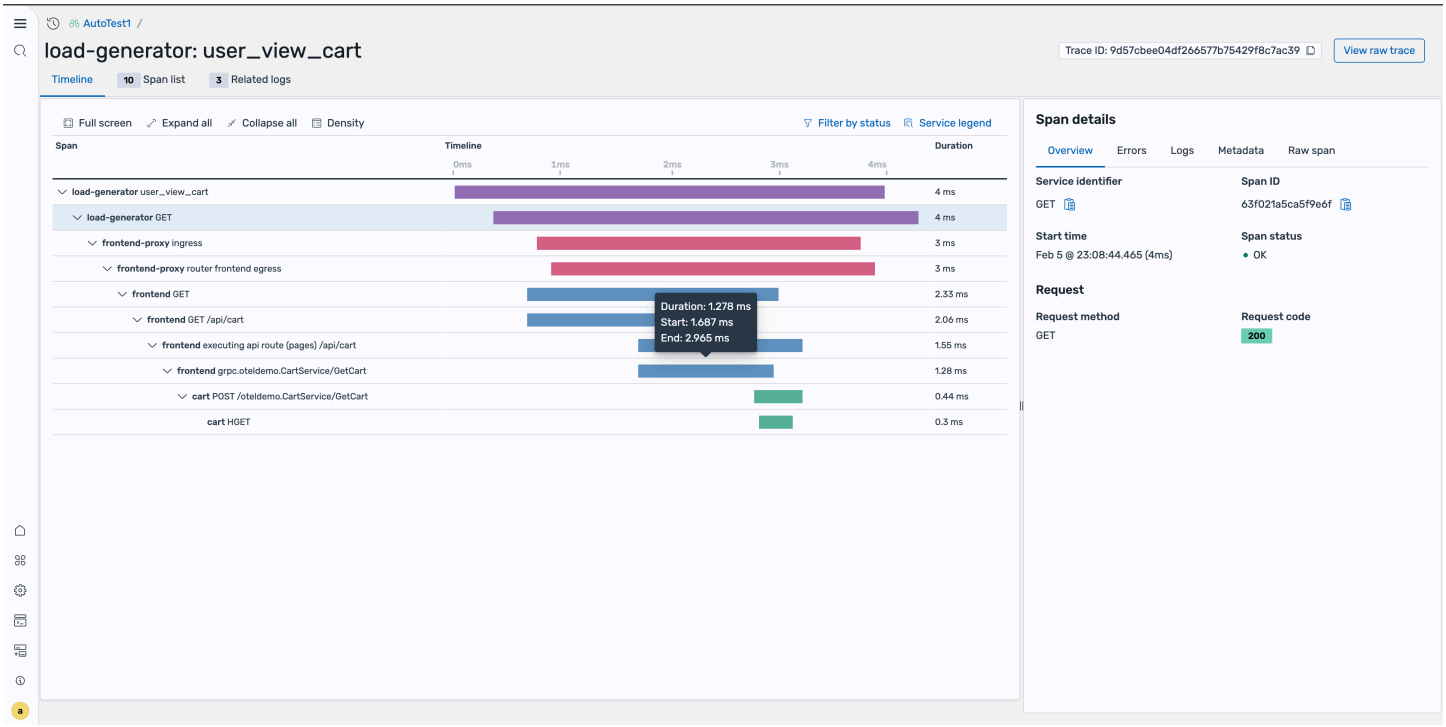
Visualizzazione di un intervallo specifico

Per visualizzare i dettagli di un intervallo specifico, scegliete la riga di estensione nella tabella degli intervalli. Si apre un pannello a comparsa con gli attributi span, gli attributi delle risorse e le informazioni sugli eventi.



Pagina dei dettagli di Trace

Per visualizzare la traccia completa, scegliete il link Trace ID nella tabella di span o nel pannello a comparsa. La pagina dei dettagli della traccia mostra un grafico a cascata che mostra tutti gli intervalli della traccia, le relative relazioni temporali e la durata complessiva della traccia. È possibile espandere i singoli intervalli per visualizzarne gli attributi e identificare i punti deboli.

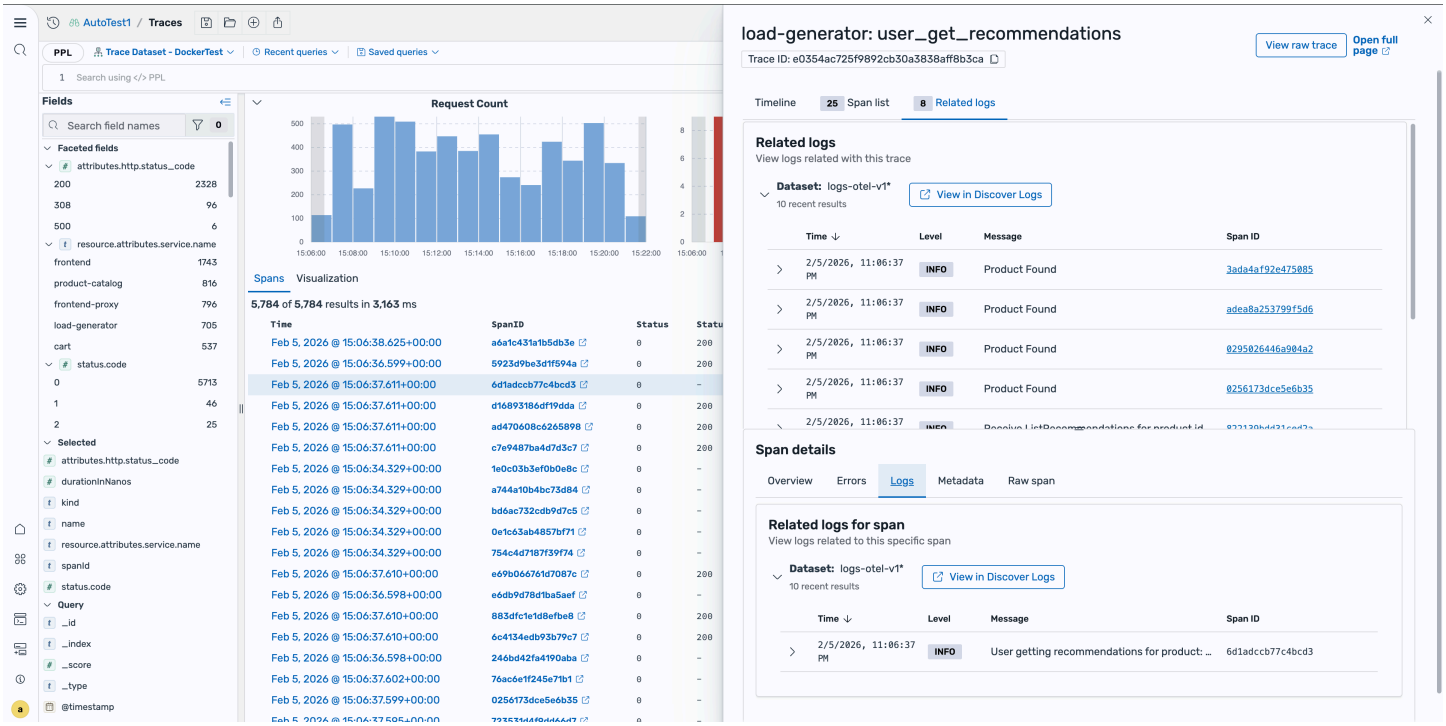


Correlazione delle tracce con i log

Quando configuri una correlazione tra un set di dati di trace e un set di dati di log, puoi visualizzare le voci di registro correlate direttamente dalla pagina Discover Traces. Per informazioni sulla creazione di correlazioni, vedere. the section called “Correlazioni”

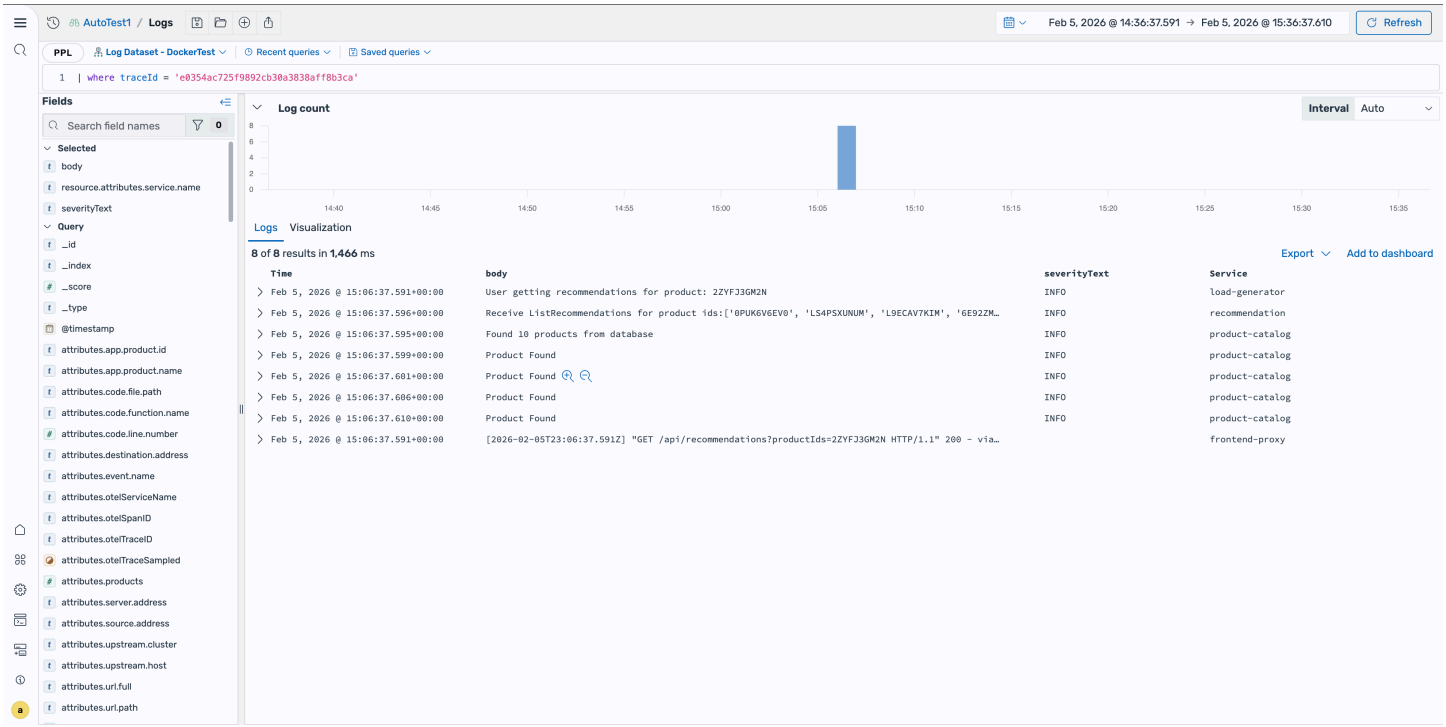
Visualizzazione dei log correlati

Nel riquadro a comparsa dei dettagli dell'intervallo o nella pagina dei dettagli di traccia, scegliete la scheda Log correlati per visualizzare le voci di registro che corrispondono all'ID di traccia, al nome del servizio e all'intervallo di tempo dello span. Questa correlazione ti aiuta a capire cosa è successo nella tua applicazione durante l'esecuzione dello span.



Reindirizzamento dei log con contesto

È possibile passare da un trace span direttamente alla pagina Discover Logs mantenendo il contesto pertinente. Scegli Visualizza nei log dal pannello dei log correlati per aprire la pagina Discover Logs con la query precompilata per filtrare in base all'ID di traccia e all'intervallo di tempo dell'intervallo di tempo dell'intervallo.



Interrogazione delle tracce tramite PPL

È possibile utilizzare PPL per interrogare direttamente i dati di traccia. PPL concatena i comandi che utilizzano il carattere pipe per filtrare, trasformare e aggregare i dati span.

L'esempio seguente trova le 10 tracce più lente:

```
source = otel-v1-apm-span-*
| where durationInNanos > 5000000000
| fields traceId, serviceName, name, durationInNanos
| sort - durationInNanos
| head 10
```

L'esempio seguente conta gli errori per servizio:

```
source = otel-v1-apm-span-*
| where status.code = 2
| stats count() as errorCount by serviceName
| sort - errorCount
```

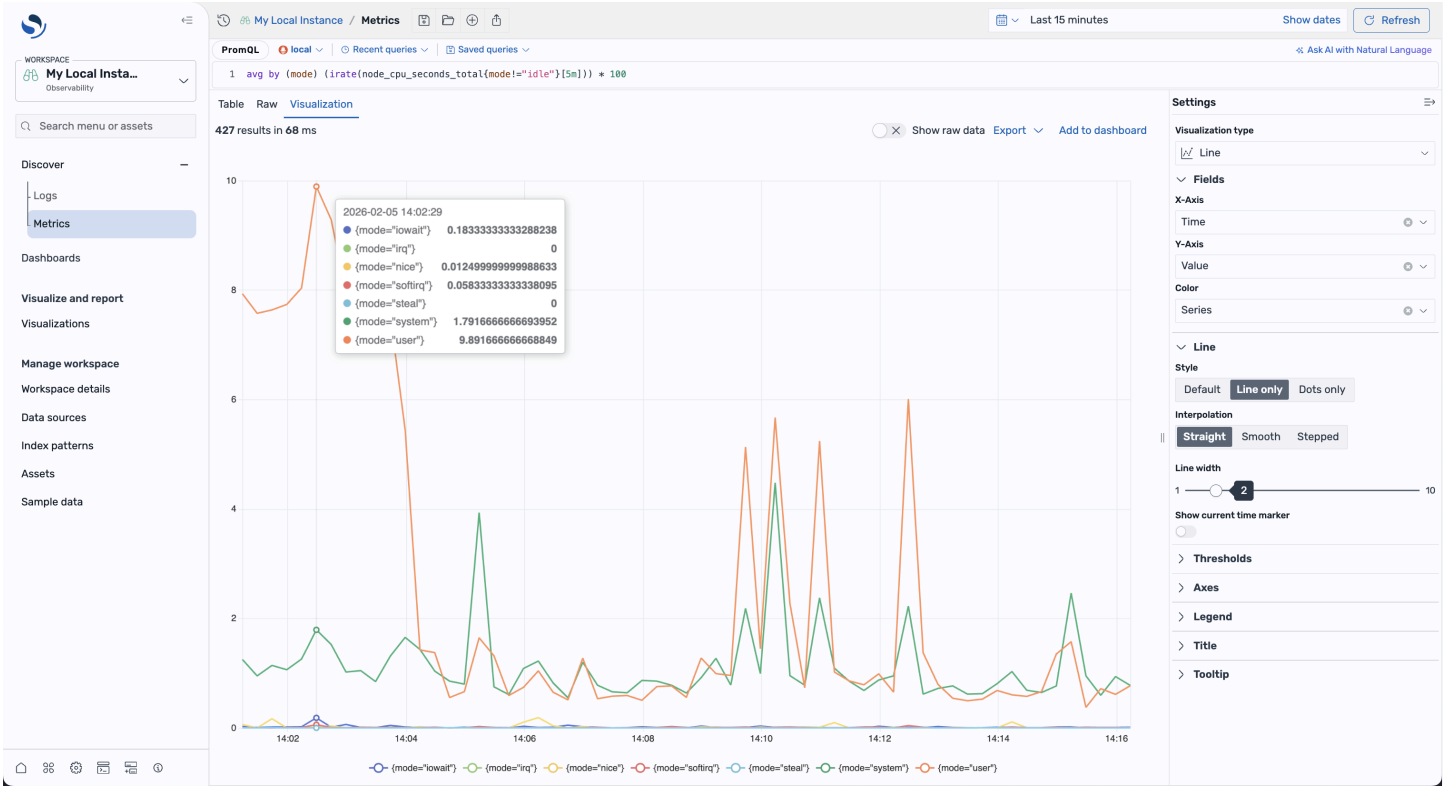
L'esempio seguente trova le tracce per un servizio specifico:

```
source = otel-v1-apm-span-*
```

```
| where serviceName = 'checkout-service'
| where parentSpanId = ''
| sort - startTime
| head 20
```

Metriche

La pagina Discover Metrics nell'interfaccia OpenSearch utente fornisce un'interfaccia dedicata per l'individuazione, l'interrogazione e la visualizzazione dei dati metrici di serie temporali. Questa pagina è ottimizzata per lavorare con le metriche di Prometheus utilizzando PromQL.



La pagina Discover Metrics è disponibile negli spazi di lavoro di Observability. Per accedervi, accedi a un'area di lavoro di Observability, espandi Discover nella barra di navigazione a sinistra e seleziona Metrics.

Configurazione di un'origine dati Prometheus

Prima di iniziare, configura un'origine dati Prometheus utilizzando uno dei seguenti metodi:

- the section called “Creazione di un'origine dati Prometheus gestita”nella console di gestione AWS
- AWS CLI

pannello Query

Puoi scrivere ed eseguire query metriche nel pannello delle query nella parte superiore della pagina Discover Metrics. L'editor di query fornisce suggerimenti per il completamento automatico e l'evidenziazione della sintassi per PromQL.

Scrittura di interrogazioni

Scrivi query utilizzando la sintassi PromQL. Esempio:

```
up{job="prometheus"}
```

Esecuzione di query

Per eseguire una query, inserisci la tua query nell'editor di query e seleziona Aggiorna.

Puoi eseguire più query PromQL insieme separandole con un punto e virgola (): `;`

```
up{job="prometheus"};
node_cpu_seconds_total{mode="idle"};
```

Ogni query viene eseguita in modo indipendente e i risultati vengono combinati nell'output.

Filtro temporale

Usa il filtro temporale per specificare l'intervallo di tempo per i tuoi dati metrici:

- Selezione rapida: scegli un intervallo di tempo relativo (ad esempio, gli ultimi 15 minuti o l'ultima 1 ora).
- Usato comunemente: seleziona tra intervalli di tempo predefiniti.
- Personalizzato: specifica l'ora di inizio e di fine assoluta.
- Aggiornamento automatico: imposta un intervallo di aggiornamento automatico.

Visualizzazione dei risultati

Dopo aver eseguito una query, i risultati vengono visualizzati in un'interfaccia a schede:

- Metriche: visualizza il punto dati più recente per ogni serie in un formato tabellare.

- Raw: mostra il punto dati più recente per ogni serie come JSON non elaborato restituito dall'origine dati.
- Visualizzazione: fornisce grafici interattivi per i dati metrici.

Configurazione delle visualizzazioni

Quando è selezionata la scheda Visualizzazione, viene visualizzato un pannello delle impostazioni sul lato destro dello schermo. Utilizzate questo pannello per:

- Seleziona un tipo di grafico: scegli tra visualizzazioni a linee, a barre, a torta, a indicatore o tabellari.
- Assi della mappa: assegna campi agli assi X e Y.
- Personalizza gli stili: regola colori, legende, griglie e altre opzioni visive.

Quando modifichi le impostazioni, la visualizzazione viene aggiornata automaticamente.

Correlazioni

Le correlazioni collegano un set di dati di tracce a un set di dati di log in modo da poter visualizzare le voci di registro correlate quando si esaminano gli intervalli di traccia. Definendo una correlazione, abiliti la pagina Discover Traces a visualizzare i log che si sono verificati durante l'esecuzione di un intervallo, aiutandoti a diagnosticare i problemi più velocemente senza passare da una pagina all'altra.

Requisiti di correlazione

Per creare una correlazione, i dati di log e trace devono contenere campi corrispondenti. La tabella seguente descrive i campi utilizzati dalla correlazione per unire i dati di traccia e di registro.

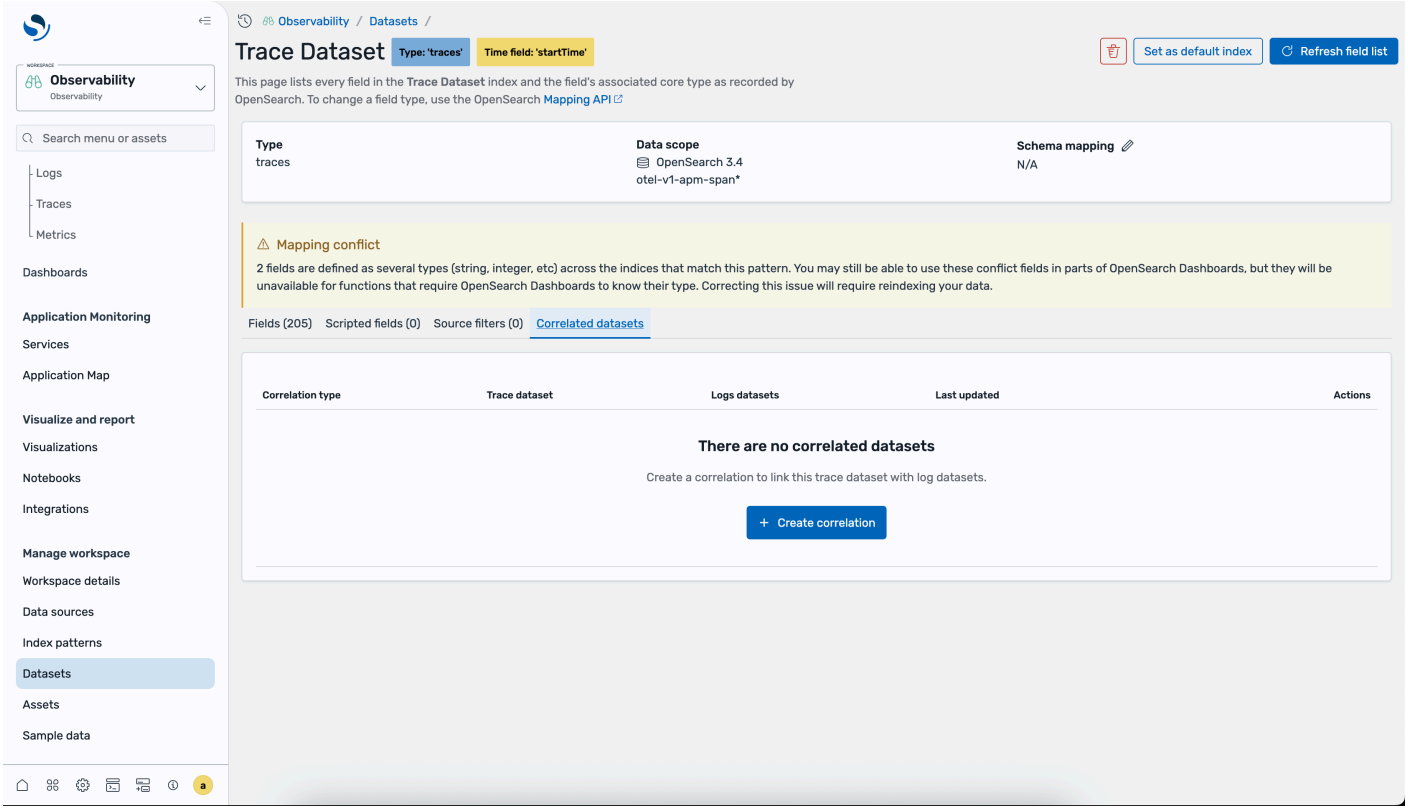
Campo	Description	Richiesto
ID di traccia	L'identificatore univoco per la traccia. Deve esistere sia nell'indice trace span che nell'indice di registro.	Sì

Campo	Description	Richiesto
ID Span	L'identificatore univoco dello span. Utilizzato per abbinare i log a un intervallo specifico all'interno di una traccia.	No
Nome del servizio	Il nome del servizio che ha generato la telemetria. Utilizzato per filtrare i log correlati in base al servizio.	No
Time stamp	Il campo orario utilizzato per includere i log correlati all'intervallo di tempo dell'intervallo di tempo dell'intervallo.	Sì

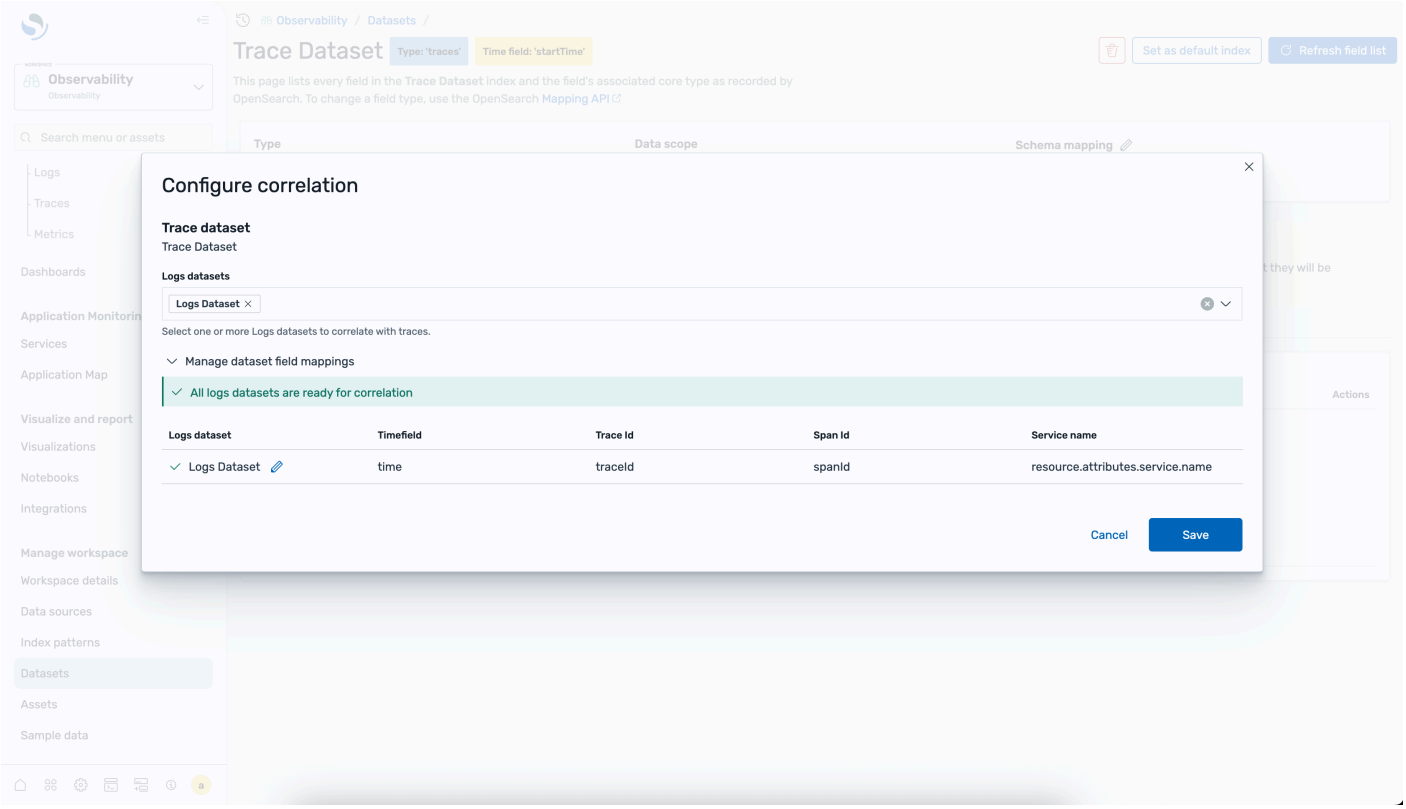
Per creare una correlazione trace-to-logs

Completa i passaggi seguenti per creare una correlazione tra un set di dati di trace e un set di dati di log.

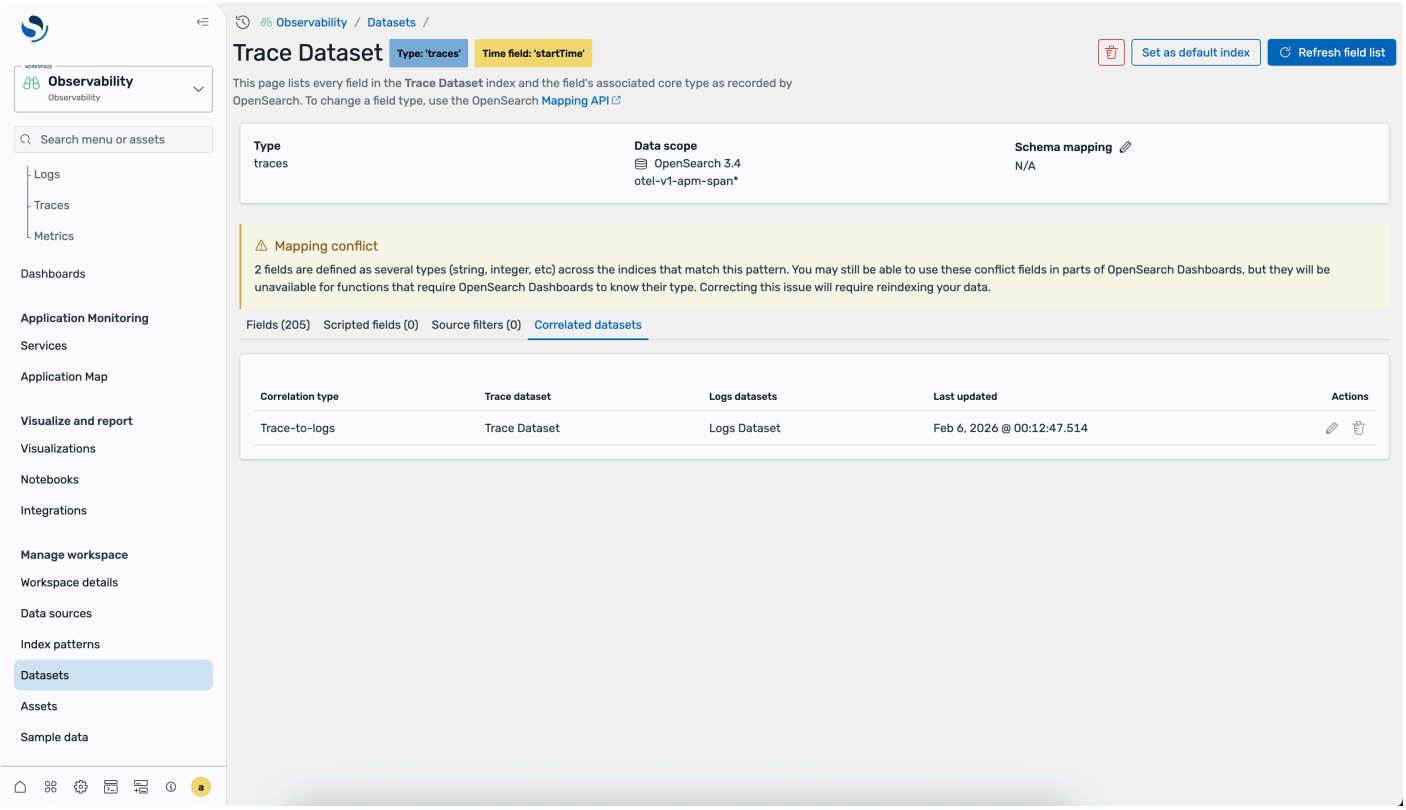
1. Nel tuo spazio di lavoro sull'osservabilità, espandi Discover nella barra di navigazione a sinistra e scegli Traces.
2. Seleziona il set di dati delle tracce che desideri correlare.
3. Scegli la scheda Correlazioni nel pannello di configurazione del set di dati.



4. Scegli Crea correlazione.
5. Nella finestra di dialogo di configurazione, seleziona il set di dati dei log di destinazione e mappa i campi di correlazione richiesti (ID di traccia e timestamp). Facoltativamente, mappa l'ID span e il nome del servizio per una corrispondenza più precisa.

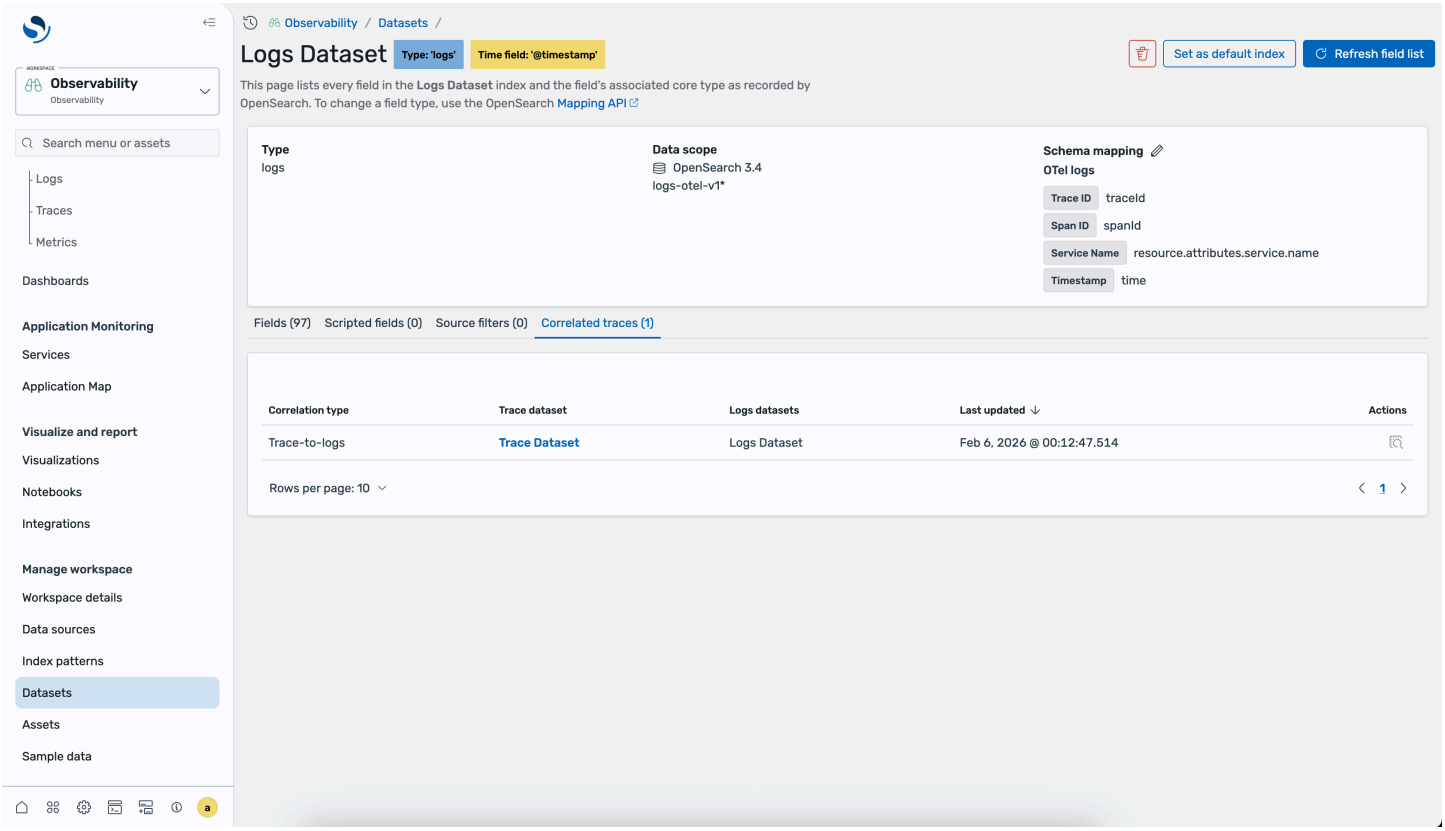


6. Scegli Crea per salvare la correlazione.
7. Verifica che la correlazione compaia nella tabella delle correlazioni.



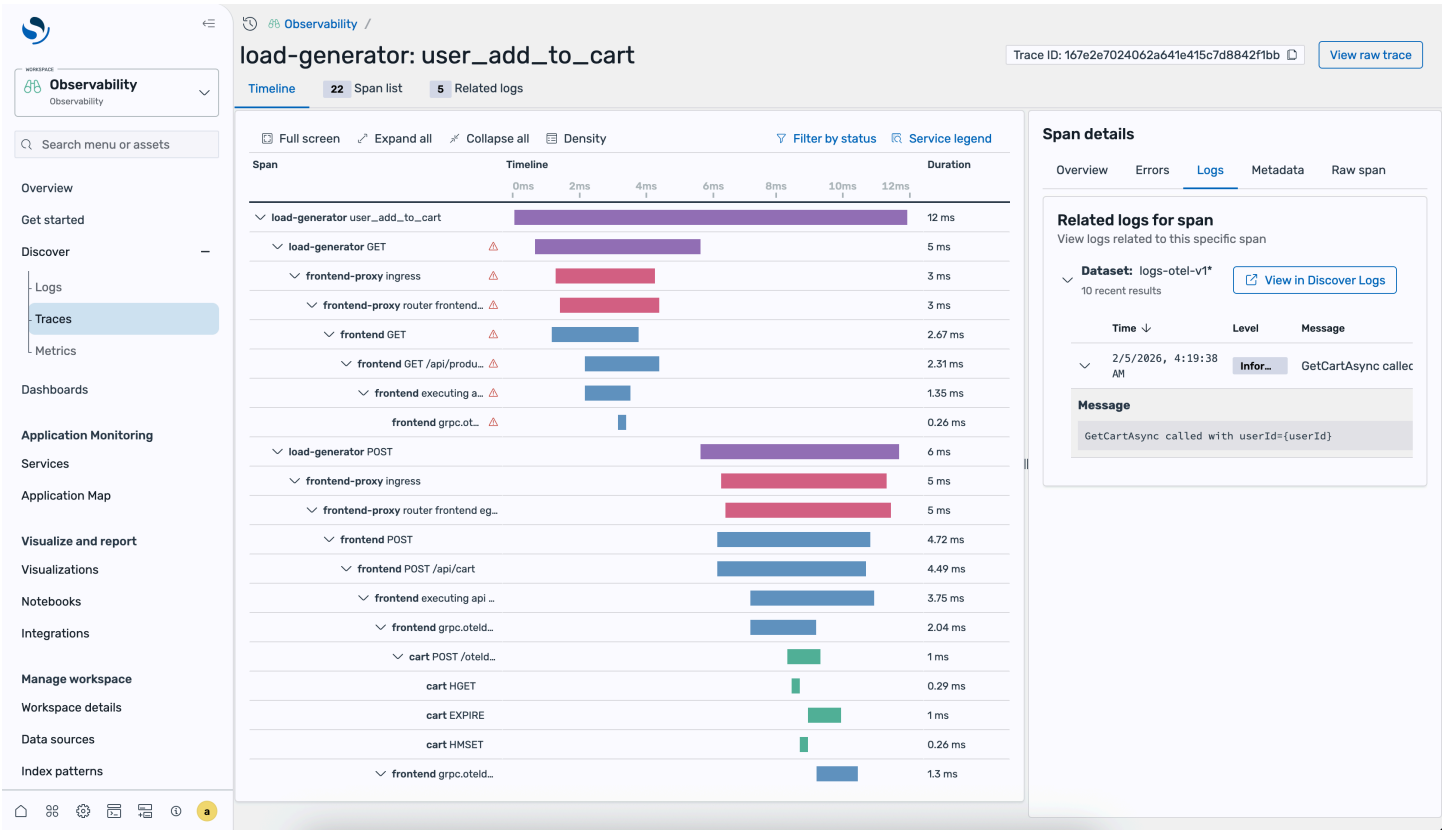
Visualizzazione delle correlazioni nei log, nei set di dati

Dopo aver creato una correlazione, puoi anche visualizzarla dal lato del set di dati dei log. Vai alla pagina Discover Logs, seleziona il set di dati dei log correlati e scegli la scheda Correlazioni per visualizzare il set di dati delle tracce collegato.



Utilizzo delle correlazioni nella pagina Traces

Quando esiste una correlazione, la pagina Discover Traces mostra i log correlati nella visualizzazione dei dettagli di span. Scegliete un intervallo nella tabella degli span per aprire il riquadro a discesa dei dettagli, quindi scegliete la scheda Log correlati per visualizzare le voci di registro correlate.



Gestione delle correlazioni

È possibile modificare o rimuovere le correlazioni dalla scheda Correlazioni del set di dati Traces o Logs.

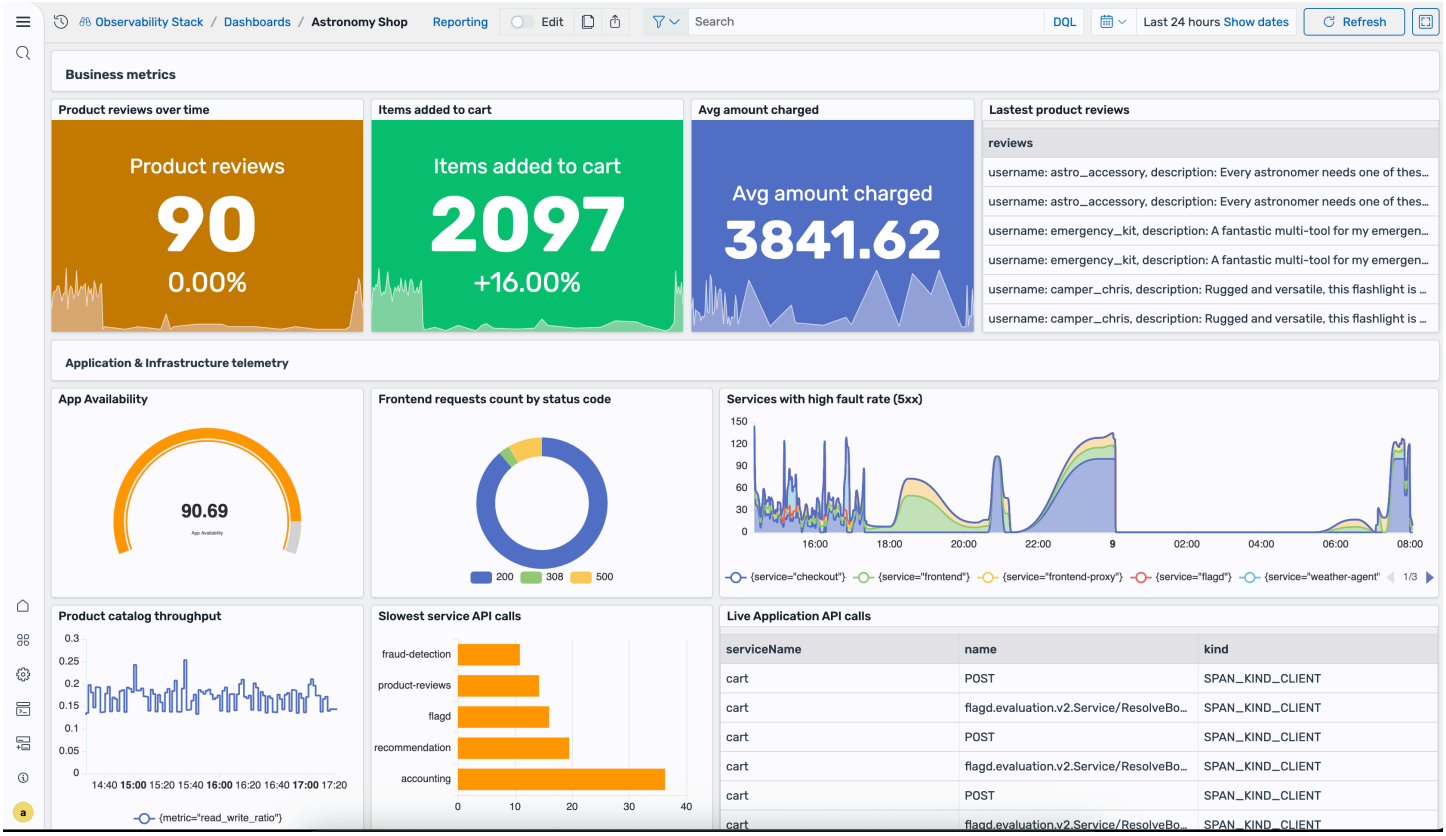
- Modifica: scegli la correlazione nella tabella e scegli Modifica per aggiornare le mappature dei campi o il set di dati di destinazione.
- Rimozione: scegli la correlazione nella tabella e scegli Elimina per rimuovere la correlazione. La rimozione di una correlazione non comporta l'eliminazione di alcun dato.

Pannelli di controllo

Le dashboard combinano le visualizzazioni di log, tracce e metriche in un'unica visualizzazione. È possibile utilizzare i dashboard per monitorare lo stato operativo, rispondere agli incidenti e tenere traccia dell'utilizzo delle risorse nel sistema distribuito.

La tabella seguente descrive i casi d'uso più comuni per i dashboard.

Caso d’uso	Esempio
Monitoraggio operativo	Monitora lo stato del servizio, la velocità effettiva e i tassi di errore in tempo reale.
Risposta agli incidenti	Correla log, tracce e metriche durante un incidente attivo.
Pianificazione della capacità	Monitora le tendenze di utilizzo delle risorse per pianificare la scalabilità.
Monitoraggio della disponibilità	Misura l'operatività e la disponibilità rispetto agli obiettivi a livello di servizio.
Revisione post-incidente	Analizza i dati storici per comprendere la causa principale degli incidenti passati.



Struttura del pannello di controllo

Una dashboard è una raccolta di pannelli disposti su una griglia. Ogni pannello è composto dai seguenti componenti.

- Origine dati: l' OpenSearch indice o l'origine dati di Amazon Managed Service for Prometheus interrogata dal pannello.
- Query: una query PPL o PromQL che recupera i dati per il pannello.
- Tipo di visualizzazione: il tipo di grafico utilizzato per visualizzare i risultati della query, ad esempio linee, barre o valori metrici.
- Configurazione opzionale: assi, legende, soglie e opzioni di formattazione.

Il selettore dell'intervallo di tempo nella parte superiore della dashboard si applica a tutti i pannelli. Puoi sovrascrivere l'intervallo di tempo per i singoli pannelli quando necessario.

Creazione di dashboard da Discover

Il flusso di lavoro consigliato per la creazione di dashboard inizia in Discover. Questo flusso di lavoro è coerente tra log, tracce e metriche.

1. Interroga i tuoi dati in Discover: accedi a Discover Logs, Discover Traces o Discover Metrics e scrivi una query utilizzando PPL (per log e tracce) o PromQL (per metriche).
2. Crea una visualizzazione: quando la query restituisce risultati, utilizza la scheda Visualizzazione per scegliere un tipo di grafico e configurare la visualizzazione. Per le query di log and trace, i comandi di aggregazione, ad esempio, passano `stats` automaticamente alla visualizzazione di visualizzazione.
3. Salva in una dashboard: scegli Aggiungi alla dashboard per salvare la visualizzazione in una dashboard nuova o esistente. Il pannello rimane attivo e si aggiorna man mano che arrivano nuovi dati.
4. Iterazione: ripeti l'operazione per ogni domanda a cui desideri che la dashboard risponda. Quando qualcosa sembra sbagliato su una dashboard, scegli un pannello qualsiasi per aprire la query sottostante in Discover per ulteriori indagini.

> **Important**
>
> Le visualizzazioni create tramite la pagina Visualizzazioni nell' OpenSearch interfaccia utente utilizzano DQL (Dashboards Query Language) e DSL (Domain Specific Language), che al momento non supportano il Piped Processing Language (PPL). Per creare visualizzazioni basate su PPL, utilizza il flusso di lavoro Discover descritto sopra.

Filtri del dashboard

I filtri consentono di restringere i dati visualizzati su tutti i pannelli di una dashboard senza modificare le singole query.

Per aggiungere un filtro

1. Apri la dashboard che desideri filtrare.
2. Scegli Aggiungi filtro nella barra dei filtri.
3. Seleziona il nome di un campo dall'elenco a discesa.
4. Seleziona un operatore e inserisci un valore.
5. Scegli Save (Salva).

La tabella seguente descrive i casi d'uso più comuni dei filtri.

Scenario	Campo	Operatore	Valore
Visualizza un singolo ambiente	`environment`	`is`	`production`
Errori isolati	`status_code`	`is greater than or equal to`	`400`
Concentrati su un servizio specifico	`service.name`	`is`	`order-service`
Escludi i controlli sanitari	`http.url`	`is not`	`/health`

Filtri bloccati rispetto a filtri non bloccati: un filtro bloccato persiste quando navighi tra le dashboard. Un filtro non aggiunto si applica solo alla dashboard corrente. Per aggiungere un filtro, scegli l'icona a forma di puntina accanto al badge del filtro.

Creazione di dashboard

Tipi di visualizzazione

La tabella seguente descrive i tipi di visualizzazione disponibili per i pannelli del dashboard.

Tipo	Caso d’uso
Line (Linea)	Tendenze nel tempo, come la frequenza delle richieste o la latenza
Area	Volume nel tempo con interruzioni cumulative
Barra	Confronto dei valori tra le categorie
Barra orizzontale	Confronti classificati, ad esempio i migliori servizi in base al numero di errori
Tabella dei dati	Dati tabulari con ordinamento e impaginazione
Valore del parametro	Indicatori chiave di prestazione singoli, come le richieste totali
Gauge	Progresso verso una soglia, ad esempio l'utilizz o della CPU
Torta	Composizione e proporzioni, come il traffico per regione
Mappa di calore	Schemi di densità e distribuzione su due dimensioni
Nuvola di tag	Frequenza relativa dei termini, come i messaggi di errore più comuni

Configurazione dei pannelli

Ogni pannello dispone di un editor di query in cui scrivere query PPL o PromQL. Gli esempi seguenti mostrano le query di pannello più comuni.

Conteggio degli errori per servizio (PPL):

```
source = logs-dataset |
    where severity_text = 'ERROR' |
    stats count() as error_count by service_name, span(timestamp, 5m)
```

Tasso di utilizzo della CPU (ProMQL):

```
rate(container_cpu_usage_seconds_total{namespace="production"}[5m])
```

È inoltre possibile configurare le seguenti opzioni del pannello.

- Assi: imposta le etichette degli assi, le scale (lineari o logaritmiche) e gli intervalli di valori.
- Legende: controlla la posizione della legenda e le serie da visualizzare.
- Soglie: aggiungi linee di soglia orizzontali per evidenziare i livelli di avviso o critici.

Suggerimenti per il layout

Utilizza i seguenti suggerimenti per organizzare i pannelli della dashboard in modo efficace.

- Posiziona pannelli di riepilogo di alto livello (valori metrici, indicatori) nella parte superiore della dashboard.
- Raggruppa i pannelli correlati, ad esempio tutti i pannelli per un unico servizio.
- Utilizzate larghezze coerenti per i pannelli della stessa riga.
- Trascina i bordi del pannello per ridimensionarlo e trascina le intestazioni dei pannelli per riposizionarlo.

Layout consigliati

Le tabelle seguenti descrivono i layout di pannello consigliati per i tipi di dashboard più comuni.

Dashboard sullo stato del servizio

Pannello	Tipo di visualizzazione
Tasso di richiesta	Line (Linea)
Tasso di errore	Line (Linea)
Latenza P99	Line (Linea)
Avvisi attivi	Valore del parametro
Principali errori per servizio	Barra orizzontale

Dashboard di risposta agli incidenti

Pannello	Tipo di visualizzazione
Log di errore	Tabella dei dati
Conteggio degli errori nel tempo	Area
Servizi interessati	Torta
Picchi di latenza	Line (Linea)

Dashboard sull'utilizzo delle risorse

Pannello	Tipo di visualizzazione
Utilizzo CPU	Gauge
Utilizzo della memoria nel tempo	Area
I/O del disco	Line (Linea)
Throughput di rete	Line (Linea)

Controlli dell'intervallo di tempo

Il selettore dell'intervallo di tempo nella parte superiore della dashboard controlla la finestra temporale di tutti i pannelli. È possibile selezionare un intervallo preimpostato (ad esempio Ultimi 15 minuti o Ultime 24 ore) o specificare un intervallo assoluto personalizzato.

Per abilitare l'aggiornamento automatico, scegli il menu a discesa dell'intervallo di aggiornamento accanto al selettore dell'intervallo di tempo e seleziona un intervallo. L'aggiornamento automatico riesegue tutte le interrogazioni del pannello all'intervallo specificato in modo che la dashboard mostri i dati più recenti.

Condivisione dei pannelli di controllo

Puoi condividere le dashboard con altri utenti dell'organizzazione tramite URLs istantanee ed esportazioni.

Condividi tramite URL

Copia l'URL della dashboard dalla barra degli indirizzi del browser e condividilo direttamente. L'URL conserva l'intervallo di tempo e i filtri correnti. Puoi includere collegamenti alla dashboard nei segnalibri, nei runbook o nella documentazione sulla risposta agli incidenti.

Snapshot

Un'istantanea cattura lo stato corrente di un pannello di controllo, inclusi tutti i dati del pannello, in un momento specifico. Le istantanee sono di sola lettura e non si aggiornano quando i dati sottostanti cambiano. Utilizza le istantanee per conservare una registrazione dello stato della dashboard durante gli incidenti o le revisioni.

Importa ed esporta definizioni

Puoi esportare una definizione di dashboard come JSON e importarla in un altro spazio di lavoro o ambiente. Questo approccio è utile per promuovere le dashboard dallo sviluppo alla produzione o per condividere layout standard tra i team.

Le migliori pratiche per la condivisione

- Destinatari: progetta dashboard per un pubblico specifico, ad esempio tecnici o dirigenti disponibili su richiesta.
- Focalizzazione: limita ogni dashboard a un unico scopo o flusso di lavoro.

- Convenzioni: utilizza convenzioni di denominazione coerenti per dashboard e pannelli in tutta l'organizzazione.
- Controllo della versione: esporta le definizioni JSON della dashboard e le archivia nel controllo delle versioni per tenere traccia delle modifiche nel tempo.

Dashboard di risoluzione dei problemi

Questa sezione descrive i problemi più comuni delle dashboard e come risolverli.

Nessun dato in un pannello

Se un pannello non visualizza dati, verificate le seguenti cause comuni.

Causa	Verifica	Correggere
Intervallo di tempo troppo ristretto	Verifica che l'intervallo di tempo del dashboard copra il periodo in cui i dati sono stati inseriti.	Espandi l'intervallo di tempo o seleziona Ultime 24 ore.
Filtro attivo, esclusi i dati	Controlla la barra dei filtri per individuare i filtri che potrebber o escludere tutti i documenti corrispondenti.	Rimuovi o modifica il filtro, quindi verifica che vengano visualizzati i dati.
Schema di indice errato	Verificate che l'origine dati del pannello punti a un indice che contiene dati.	Aggiorna la fonte di dati secondo lo schema di indice corretto nell'editor del pannello.
Errore di sintassi della query	Cerca i messaggi di errore nell'intestazione del pannello o nell'editor di query.	Correggi la sintassi PPL o PromQL ed esegui nuovament e la query.

Dati errati in un pannello

Se un pannello mostra risultati imprevisti, provate i seguenti passaggi.

- Verifica che la query restituisca i campi previsti eseguendola prima in Discover.
- Verifica che il tipo di visualizzazione corrisponda alla forma dei dati (ad esempio, utilizza un grafico a linee per i dati delle serie temporali).
- Verifica che sia selezionata la fonte di dati corretta nell'editor del pannello.

Dati obsoleti

Se i pannelli del dashboard visualizzano informazioni non aggiornate, prova i seguenti passaggi.

- Scegli l'icona di aggiornamento nella barra degli strumenti per aggiornare manualmente tutti i pannelli.
- Verifica che l'aggiornamento automatico sia abilitato e impostato su un intervallo appropriato.
- Verifica che la pipeline di ingestione stia inviando attivamente i dati agli indici configurati.

Problemi di prestazioni

I seguenti suggerimenti possono aiutarti a risolvere i problemi di prestazioni più comuni.

- Dashboard lento: riduci il numero di pannelli o restringi l'intervallo di tempo. Le dashboard con molti pannelli eseguono più query contemporaneamente, il che può aumentare i tempi di caricamento.
- Pannello lento: semplifica l'interrogazione del pannello. Evita di utilizzare schemi jolly nelle `where` clausole PPL e limita il numero di bucket di aggregazione.
- Browser lag: riduce la densità dei dati nelle visualizzazioni. Ad esempio, aumenta l'intervallo di tempo nei `stats` comandi per produrre meno punti dati.

Problemi di filtro

Se i filtri non si comportano come previsto, prova i seguenti passaggi.

- Verifica che il nome del campo nel filtro corrisponda al nome del campo nella mappatura dell'indice.
- Controlla se un filtro bloccato da un'altra dashboard influisce sui risultati.
- Rimuovi tutti i filtri e aggiungili nuovamente uno alla volta per isolare il problema.

Ispeziona un pannello

L'ispettore del pannello consente di eseguire il debug dei dati e dei problemi relativi alle interrogazioni. Per aprire l'inspector, scegliete il menu del pannello (tre puntini) e selezionate Ispeziona. L'ispettore fornisce le seguenti schede.

- Dati: visualizza i dati grezzi restituiti dalla query in formato tabulare.
- Richiesta: mostra la query inviata all'origine dati, inclusa l'istruzione PPL o PromQL completa.
- Risposta: mostra la risposta non elaborata dall'origine dati, comprese le informazioni sulla tempistica e sullo stato.

Strumenti per sviluppatori di browser

Per una risoluzione avanzata dei problemi, utilizzate gli strumenti di sviluppo del browser per esaminare le richieste di rete. Apri la scheda Rete, filtra le chiamate API e cerca le richieste non riuscite o le risposte lente. Controlla il corpo della risposta per i messaggi di errore che possono aiutarti a identificare la causa principale.

Monitoraggio dell'applicazione

Il monitoraggio delle applicazioni fornisce una visione in tempo reale delle prestazioni dei servizi. Combina i dati topologici archiviati OpenSearch con le metriche RED di serie temporali (Rate, Errors, Duration) di Amazon Managed Service for Prometheus per far emergere informazioni su salute, latenza, velocità effettiva ed errore nel sistema distribuito.

Per accedere al monitoraggio delle applicazioni, nell'interfaccia utente accedi a Observability > Application Monitoring. OpenSearch La barra laterale mostra due visualizzazioni:

- Mappa dell'applicazione: grafico topologico interattivo delle dipendenze dei servizi
- Servizi: catalogo di tutti i servizi strumentati con filtri, visualizzazioni dettagliate e collegamenti di correlazione

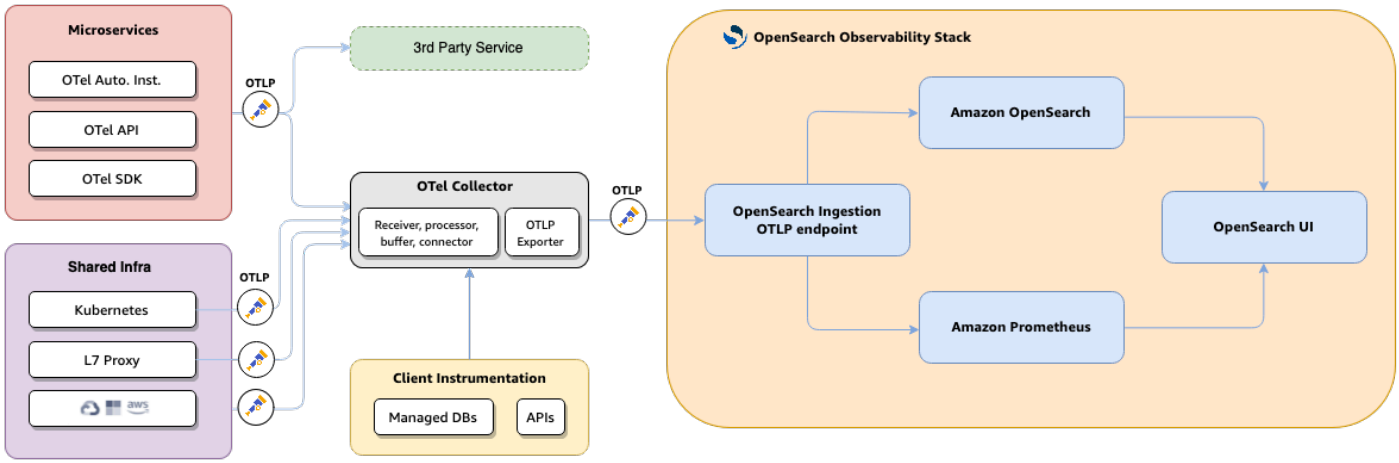
Prerequisiti

Prima di poter utilizzare il monitoraggio delle applicazioni, è necessario aver configurato le seguenti risorse.

- Dati di traccia OTLP che fluiscono dai OTel Collector a OpenSearch Ingestion (le metriche e i log sono opzionali)
- Amazon Managed Service per Prometheus configurato per ricevere la scrittura remota da Ingestion OpenSearch
- Uno spazio di lavoro dell'interfaccia utente con Observability abilitato OpenSearch

Come funziona

Il diagramma seguente mostra l' end-to-endarchitettura per il monitoraggio delle applicazioni.



1. Le applicazioni e l'infrastruttura inviano telemetria al Collector tramite OpenTelemetry SDKs la strumentazione automatica o l'API. OTel OTel
2. The OTel Collector inoltra i dati di traccia a Ingestion tramite OTLP. OpenSearch
3. Il processore OpenSearch Ingestion estrae le relazioni e calcola le metriche RED. `otel_apm_service_map` service-to-service
4. La topologia e i dati di traccia non elaborati vengono indicizzati in. OpenSearch Le metriche RED vengono esportate in Amazon Managed Service for Prometheus tramite scrittura remota.
5. OpenSearch L'interfaccia utente interroga entrambi gli archivi per visualizzare la mappa dell'applicazione, il catalogo dei servizi e le visualizzazioni dei dettagli del servizio.

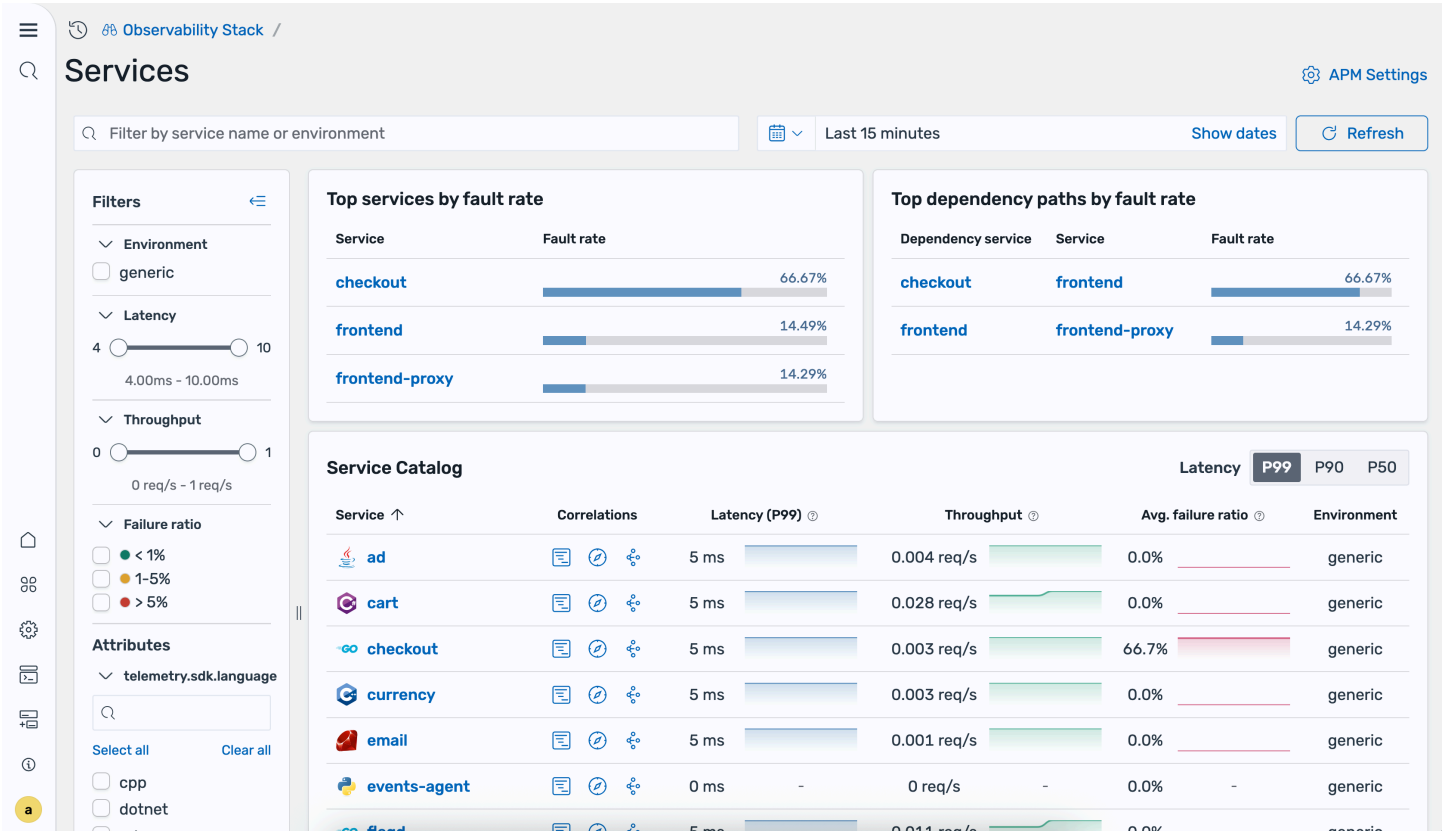
Servizi

La visualizzazione Servizi fornisce un catalogo centralizzato di tutti i servizi strumentati, che mostra le metriche RED (Rate, Errors, Duration) a colpo d'occhio. È possibile utilizzare questa visualizzazione

per identificare rapidamente i servizi non funzionanti e approfondire le visualizzazioni dettagliate per un'analisi più approfondita.

Per accedere alla vista Servizi, vai all'area di lavoro Observability nell' OpenSearch interfaccia utente e scegli APM > Servizi.

La home page dei Servizi mostra una tabella di tutti i servizi strumentati insieme a pannelli di riepilogo. L'immagine seguente mostra la home page dei Servizi.



La tabella seguente descrive le colonne della tabella dei servizi.

Colonna	Description
Nome del servizio	Il nome del servizio strumentato.
Latenza P99	La latenza del 99° percentile per il servizio.
Latenza P90	Latenza del 90° percentile per il servizio.
Latenza P50	La latenza del 50° percentile (mediana) del servizio.

Colonna	Description
Total Requests (Richieste totali)	Il numero totale di richieste elaborate nell'inte rvallo di tempo selezionato.
Rapporto di fallimento	Il rapporto tra richieste non riuscite e richieste totali.
Ambiente	L'ambiente di distribuzione del servizio, ad esempio `production ostaging`.

La home page include anche i seguenti pannelli di riepilogo:

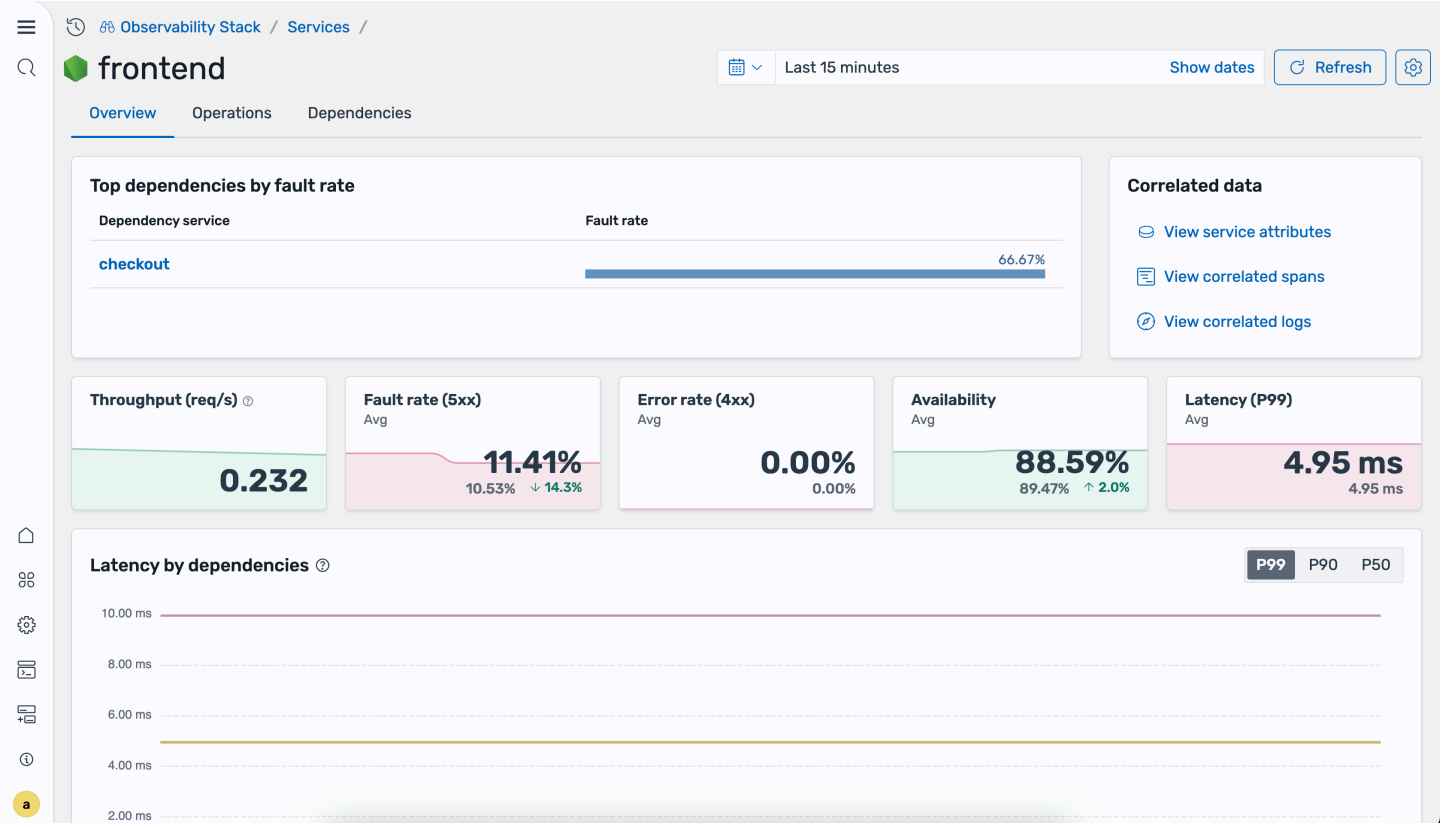
- I migliori servizi per percentuale di errore: servizi con la più alta percentuale di risposte 5xx.
- Principali percorsi di dipendenza in base alla frequenza di errore: percorsi di Service-to-service dipendenza con i tassi di errore più elevati.

È possibile filtrare la tabella dei servizi utilizzando i seguenti filtri:

- Ambiente: filtra per ambiente di distribuzione.
- Latenza: filtra per intervallo di latenza.
- Throughput: filtra in base all'intervallo di throughput della richiesta.
- Rapporto di guasto: filtra per intervallo del rapporto di guasto.

Panoramica del servizio

Per aprire la visualizzazione dettagliata del servizio, selezionare un nome di servizio nella tabella dei servizi. La scheda Panoramica mostra riquadri metrici e grafici delle serie temporali per il servizio selezionato.

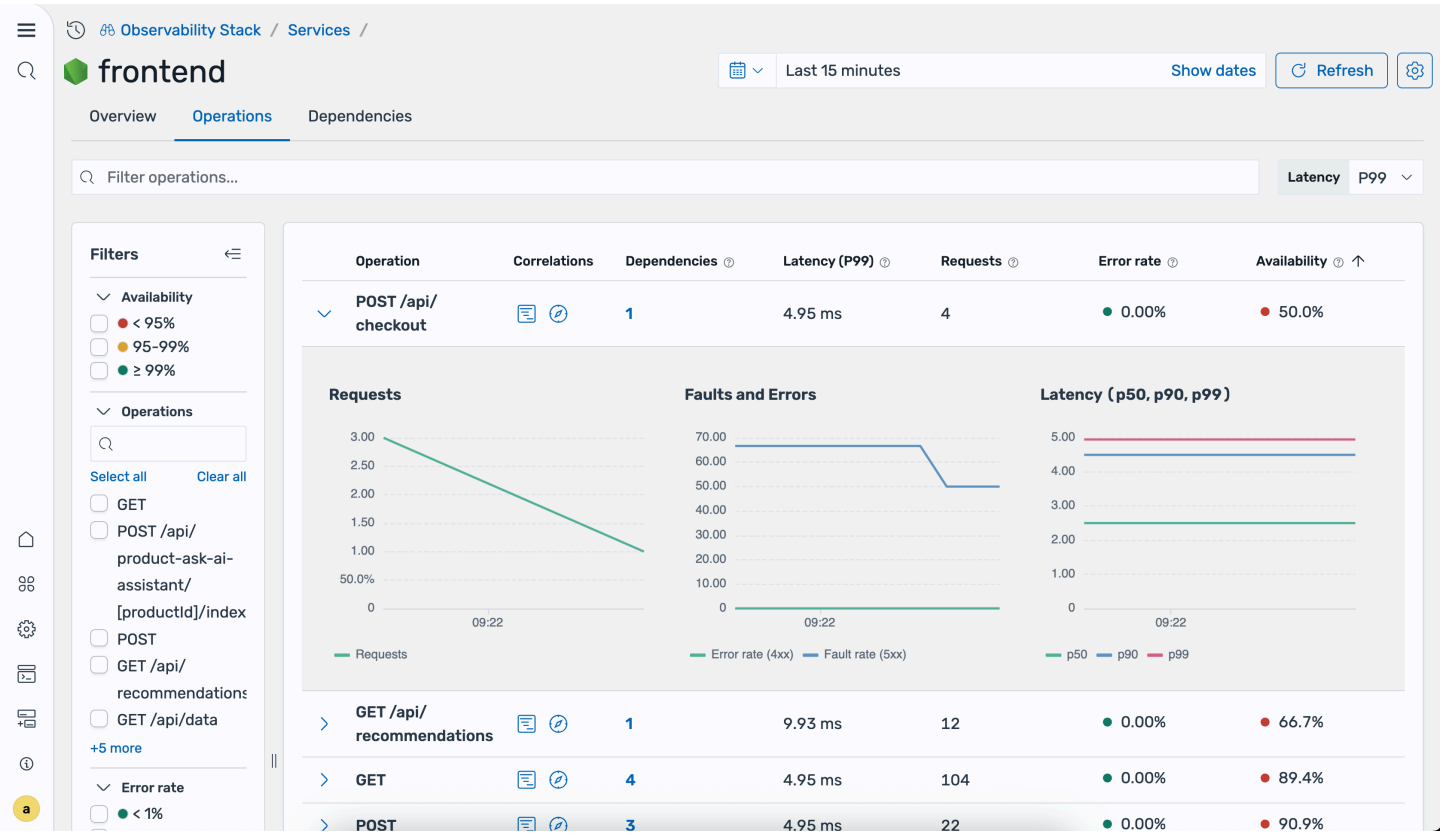


La scheda Panoramica include i seguenti grafici delle serie temporali:

- Latenza per dipendenze del servizio: latenza P50, P90 e P99 suddivisa per dipendenze a valle.
- Richieste per operazioni: volume delle richieste per ogni operazione del servizio.
- Disponibilità per operazioni: percentuale di risposte riuscite per ciascuna operazione.
- Percentuale di guasti e tasso di errore per operazione: percentuale di risposte 5xx e 4xx per ogni operazione.

Operazioni

La scheda Operazioni fornisce una suddivisione per operazione per il servizio selezionato. È possibile ordinare la tabella in base a qualsiasi colonna per identificare le operazioni problematiche.

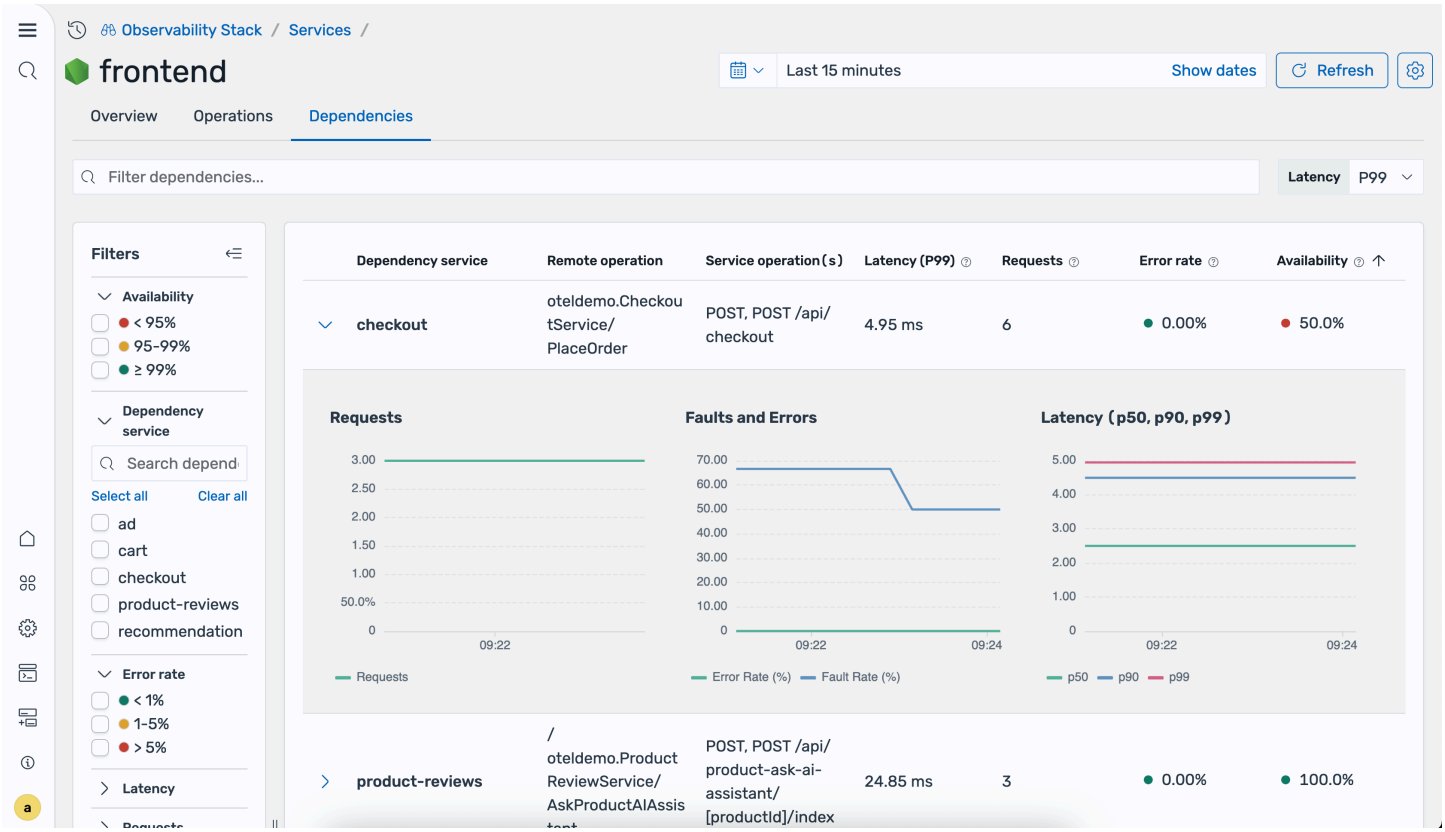


La tabella seguente descrive le colonne della tabella delle operazioni.

Colonna	Description
Nome operazione	Il nome dell'operazione.
Latenza P50/P90/P99	La latenza del 50°, 90° e 99° percentile per l'operazione.
Total Requests (Richieste totali)	Il numero totale di richieste per l'operazione nell'intervallo di tempo selezionato.
Tasso di errore	La percentuale di richieste che hanno restituito errori.
Disponibilità	La percentuale di risposte riuscite per l'operazione.

Dipendenze

La scheda Dipendenze mostra i servizi a valle richiamati dal servizio selezionato.



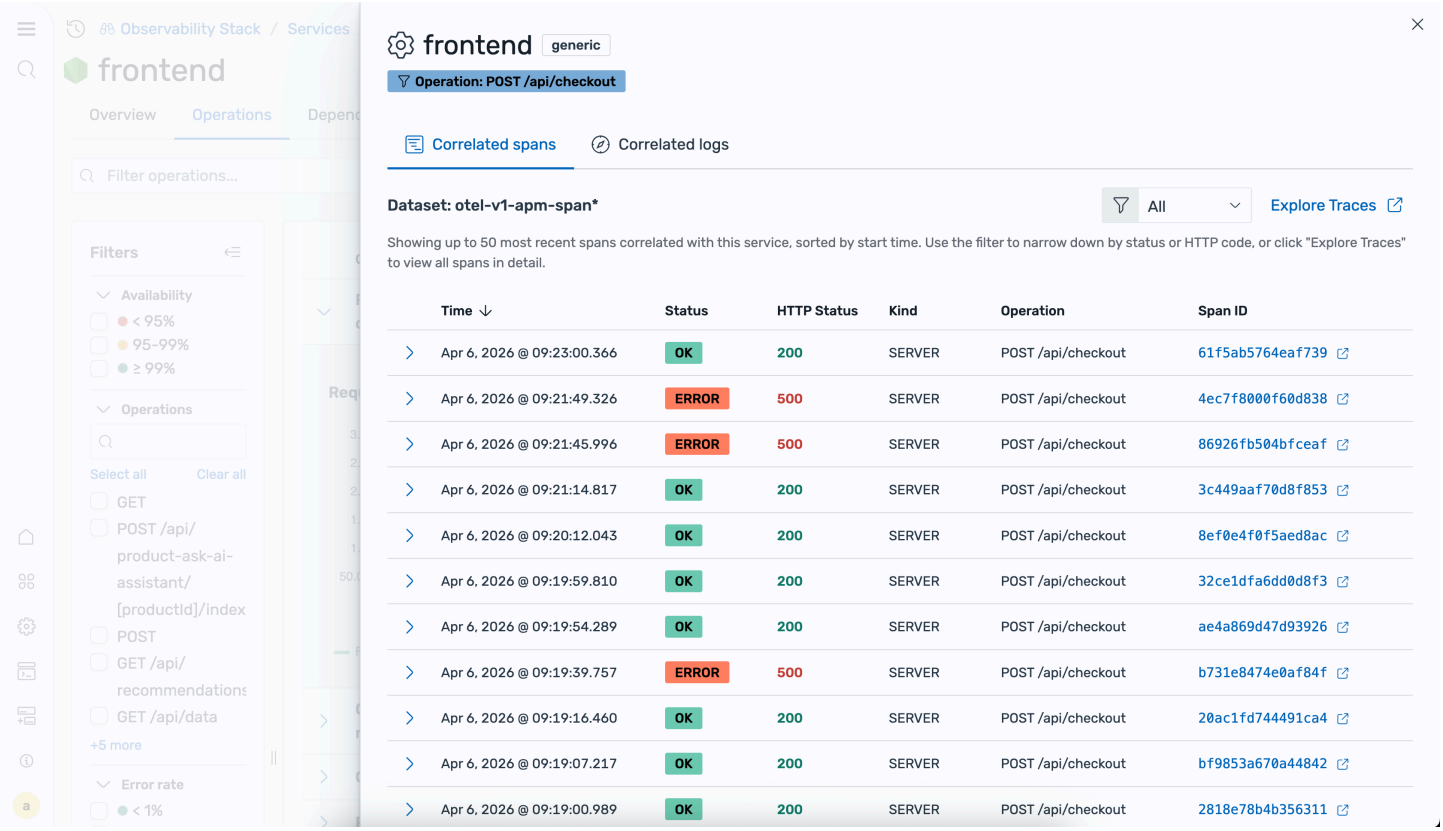
La tabella seguente descrive le colonne della tabella delle dipendenze.

Colonna	Description
Servizio di dipendenza	Il nome del servizio downstream.
Funzionamento remoto	L'operazione richiamata dal servizio downstrea m.
Operazioni di servizio	Le operazioni sul servizio corrente che richiamano questa dipendenza.
Latenza P99/P90/P50	La latenza del 99°, 90° e 50° percentile per il percorso di dipendenza.
Total Requests (Richieste totali)	Il numero totale di richieste alla dipendenza durante l'intervallo di tempo selezionato.

Colonna	Description
Tasso di errore	La percentuale di richieste alla dipendenza che hanno restituito errori.
Disponibilità	La percentuale di risposte riuscite derivanti dalla dipendenza.

Correlazioni

La visualizzazione dettagliata del servizio fornisce correlazioni contestuali che consentono di passare dalle metriche del servizio direttamente alle tracce e ai log correlati. È possibile utilizzare le correlazioni per analizzare la causa principale dei picchi di latenza o degli aumenti del tasso di errore.



Sono disponibili le seguenti opzioni di correlazione:

- Visualizza tracce correlate: apre una visualizzazione di traccia filtrata per il servizio o l'operazione selezionati.

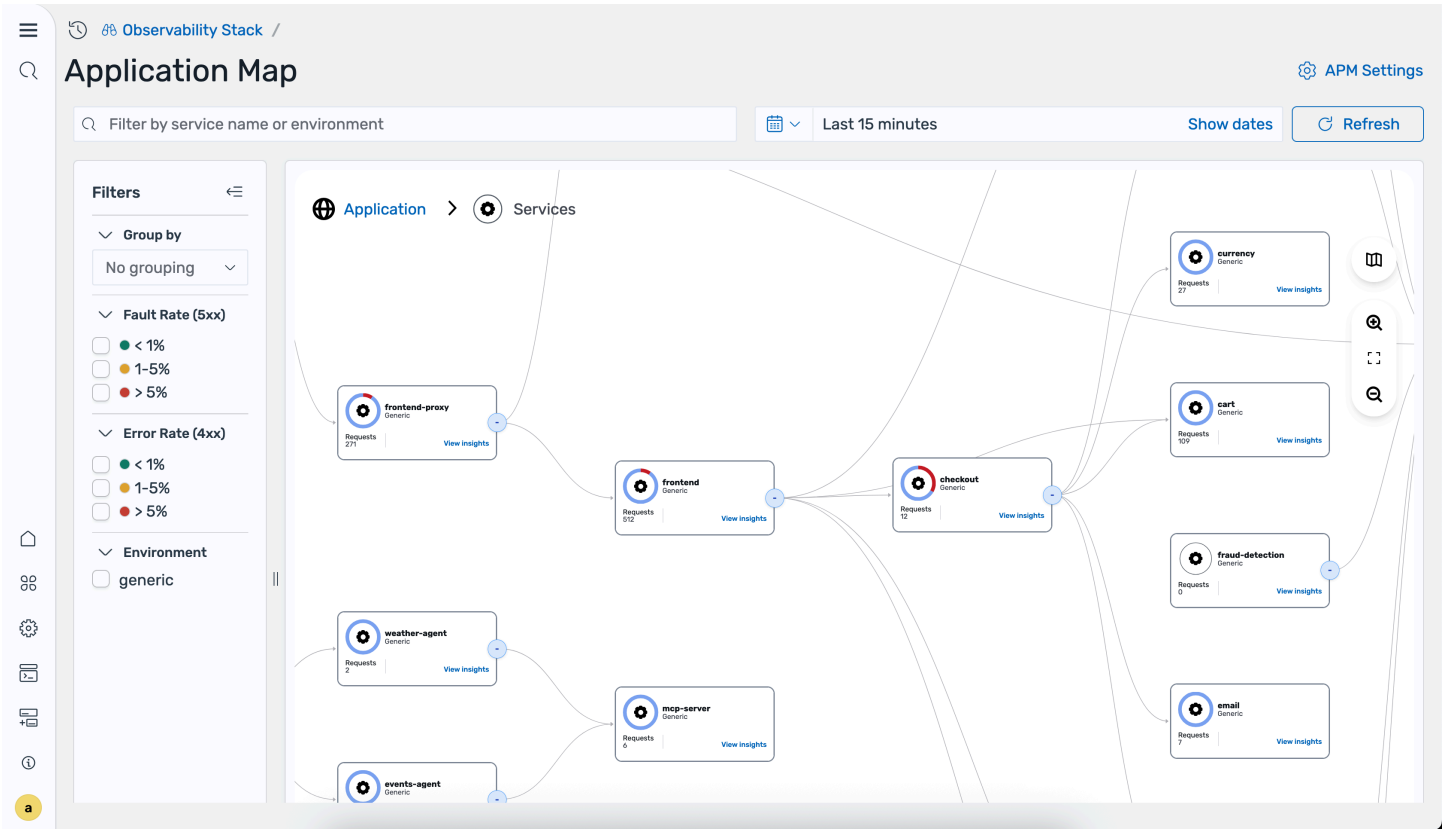
- Visualizza registri correlati: apre una visualizzazione di registro filtrata per il servizio o l'operazione selezionati.
- Filtra per attributi: limita i risultati di correlazione in base a specifici attributi span.

Mappa dell'applicazione

L'Application Map è una visualizzazione topologica interattiva che OpenSearch Ingestion genera automaticamente dai dati di traccia utilizzando il processore. `otel_apm_service_map` La mappa mostra i servizi come nodi con bordi direzionali che mostrano schemi di comunicazione, sovrapposti a metriche RED (Rate, Errors, Duration).

Per accedere alla mappa dell'applicazione, vai all'area di lavoro Observability nell' OpenSearch interfaccia utente e scegli APM > Mappa dell'applicazione.

L'immagine seguente mostra la mappa dell'applicazione.



La mappa mostra le seguenti metriche RED per ogni servizio:

- Frequenza: richieste al secondo elaborate dal servizio.
- Errori: percentuale di risposte 4xx e 5xx.

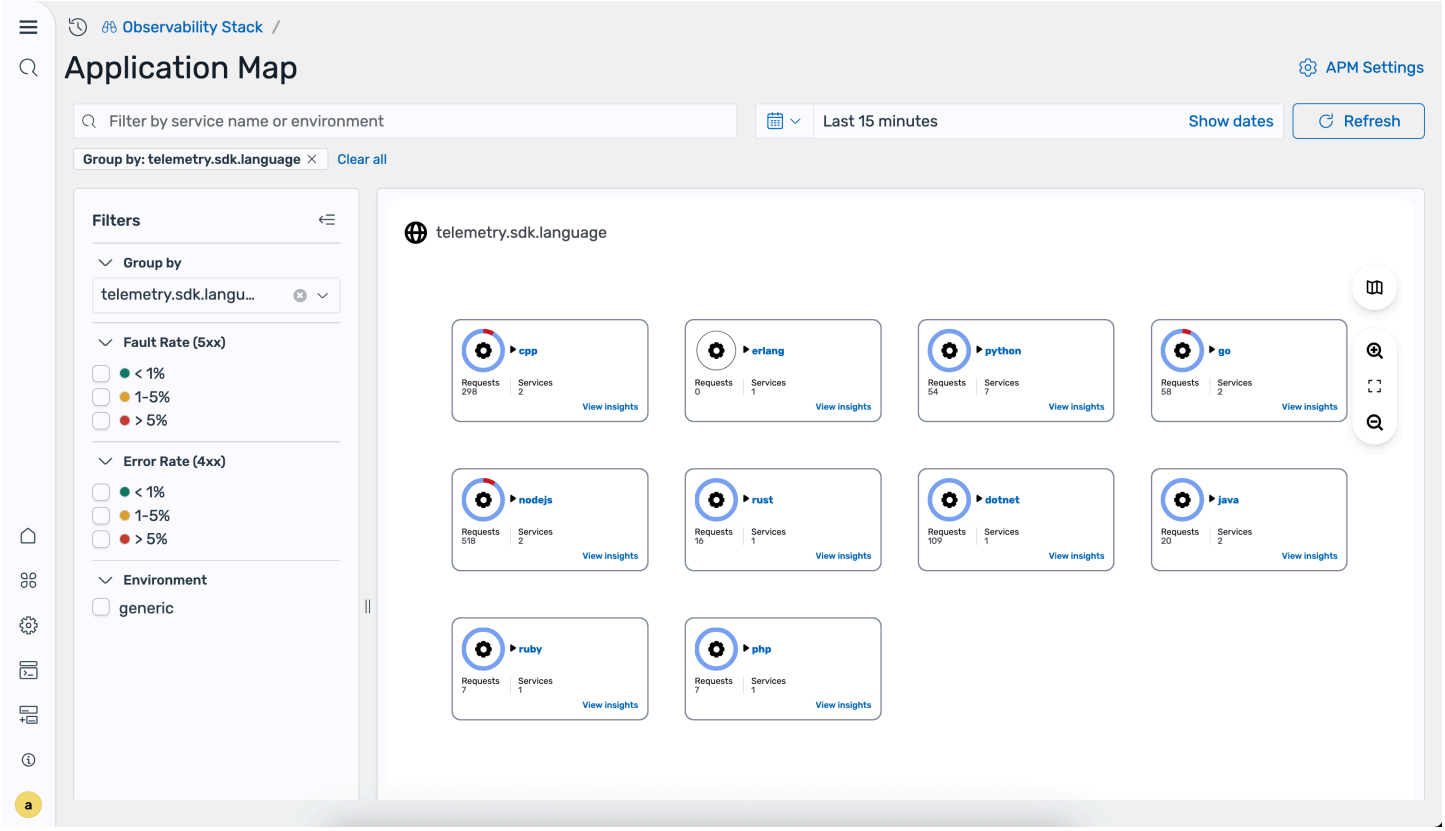
- Durata: latenza P50 e P99 per il servizio.

Il `otel_apm_service_map` processore genera queste metriche e le archivia in Amazon Managed Service for Prometheus tramite scrittura remota.

La visualizzazione della topologia rappresenta i servizi come nodi e la direzione della comunicazione come bordi. La codifica a colori indica lo stato di salute di ogni servizio. La mappa si aggiorna automaticamente non appena OpenSearch Ingestion acquisisce nuovi dati di traccia.

Servizi di raggruppamento

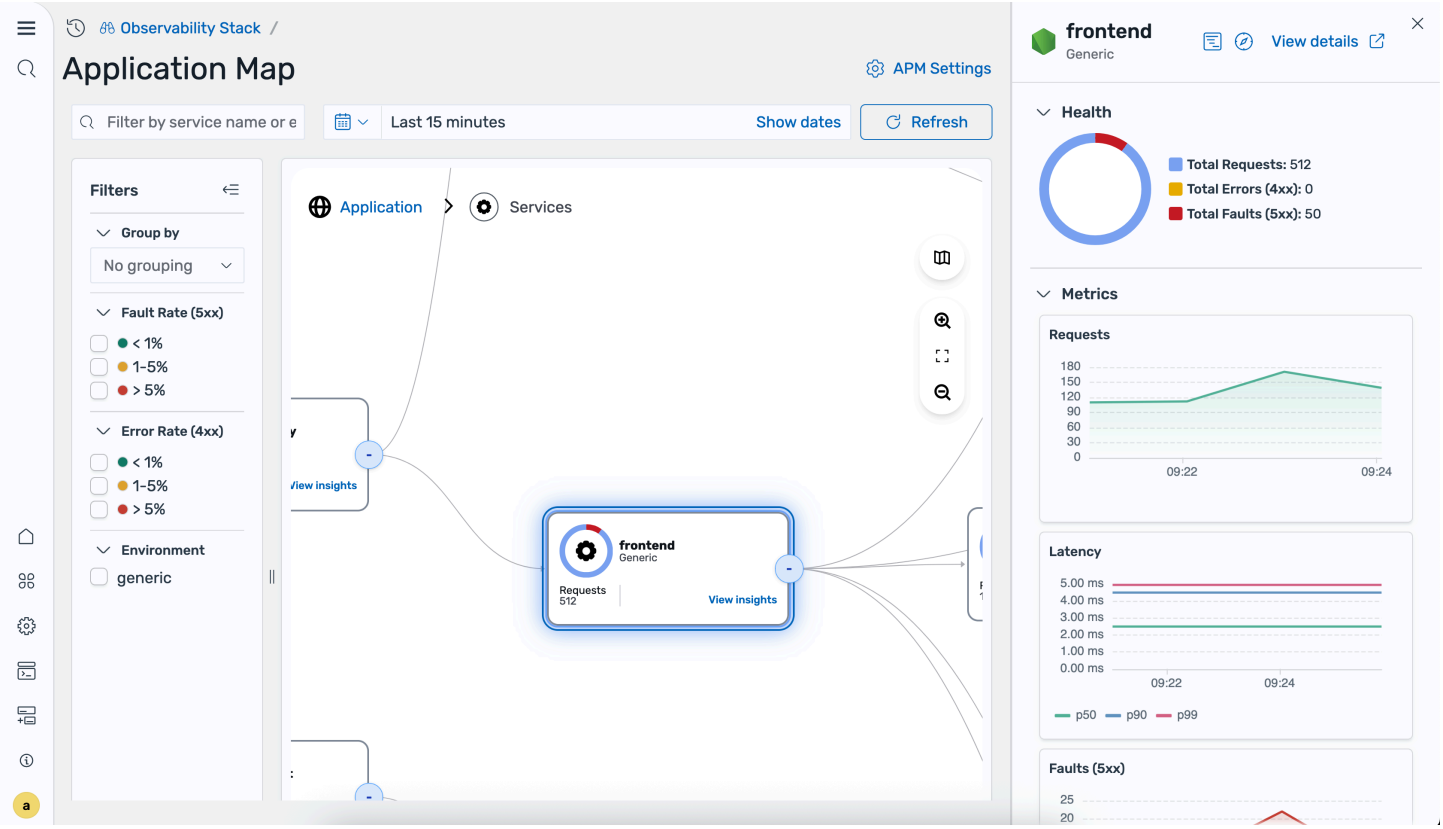
È possibile raggruppare i servizi in base ad attributi quali linguaggio di programmazione, team o ambiente. Quando si seleziona un attributo raggruppato per, la mappa passa da un grafico della topologia a una visualizzazione a griglia a schede. Ogni scheda rappresenta un gruppo di servizi che condividono lo stesso valore di attributo.



Gli attributi raggruppati per disponibili sono determinati dall'`group_by_attributes`impostazione nella configurazione del `otel_apm_service_map` processore in Ingestion. OpenSearch

Visualizzazione dei dettagli dei nodi

Per visualizzare i dettagli di un servizio, selezionare un nodo sulla mappa. Si apre un pannello di dettaglio con le seguenti sezioni.



La sezione Health mostra le seguenti metriche di riepilogo:

- Total Requests (Richieste totali)
- Errori totali 4xx
- Difetti totali 5xx

La sezione Metriche mostra i seguenti grafici delle serie temporali:

- Richieste
- Latenza P50/P90/P99
- Guasti 5xx
- Errori 4xx

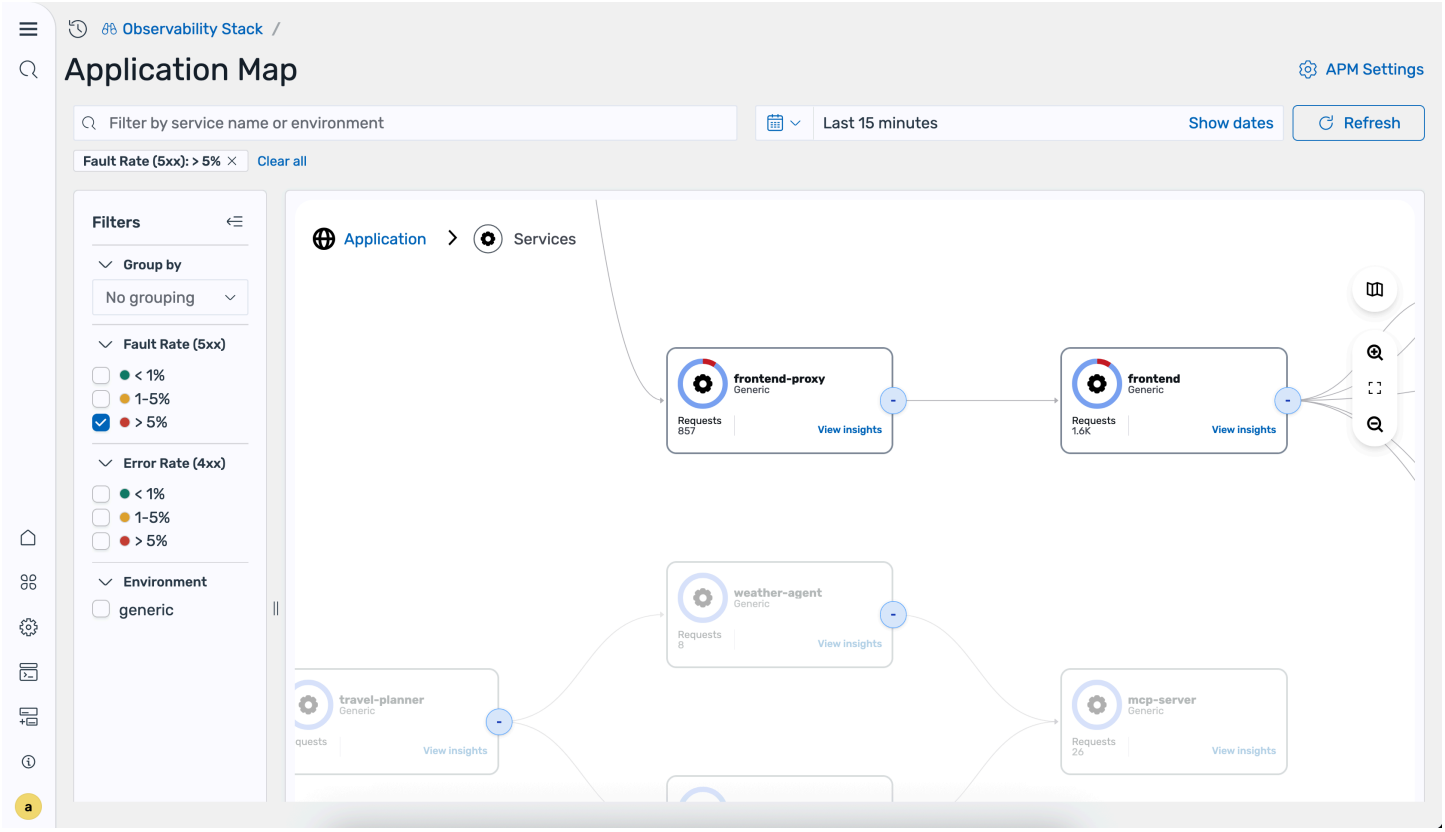
Scegli Visualizza dettagli per accedere alla visualizzazione dei dettagli dei servizi per il servizio selezionato.

Filtrare la mappa

È possibile filtrare la mappa dell'applicazione utilizzando i seguenti filtri:

- Frequenza di errore: filtra i servizi in base alla frequenza di errore lato server (5xx).
- Frequenza di errore: filtra i servizi in base alla frequenza di errore lato client (4xx).
- Ambiente: filtra i servizi in base all'ambiente di distribuzione.

L'immagine seguente mostra la mappa filtrata per tasso di errore.



Correlazioni contestuali

È possibile passare dalla vista topologica direttamente alle tracce e ai log correlati. Da qualsiasi nodo di servizio, sono disponibili le seguenti opzioni di correlazione:

- Visualizza tracce correlate: apre una visualizzazione di traccia filtrata per il servizio selezionato.
- Visualizza log correlati: apre una visualizzazione di log filtrata per il servizio selezionato.

Osservabilità dell'IA

L'osservabilità dell'IA OpenSearch fornisce end-to-end strumenti per il monitoraggio, il debug e l'ottimizzazione dei flussi di lavoro degli agenti AI e dei modelli di linguaggio di grandi dimensioni (LLM). Basato sulle convenzioni semantiche GenAI e integrato nativamente con OpenTelemetry (OTel), offre una visibilità completa sul comportamento delle applicazioni di intelligenza artificiale in produzione.

L'osservabilità dell'IA include le seguenti funzionalità:

- Tracciamento degli agenti: acquisisce le tracce di esecuzione gerarchiche tra le fasi di orchestrazione degli agenti, le chiamate LLM, le chiamate agli strumenti e le operazioni di recupero.
- Convenzioni semantiche GenAI: utilizza attributi standardizzati come e per descrivere la telemetria specifica dell'intelligenza artificiale. OTel `gen_ai.system` `gen_ai.request.model` `gen_ai.usage.input_tokens`
- Strumentazione automatica: acquisisci automaticamente le tracce dai framework e dai provider di intelligenza artificiale più diffusi, tra cui OpenAI, Anthropic, Amazon Bedrock e più di 20 librerie aggiuntive. LangChain
- Interrogazione PPL: interroga e aggrega i dati di traccia utilizzando Piped Processing Language (PPL) direttamente dall'interfaccia utente. OpenSearch

Nozioni di base

Questa sezione illustra come utilizzare un agente AI, inviare tracce ad Amazon OpenSearch Service e visualizzarle nell' OpenSearch interfaccia utente.

Per installare l'SDK

Installa il pacchetto di strumentazione OpenTelemetry GenAI:

```
pip install opentelemetry-instrumentation-openai-v2 opentelemetry-sdk opentelemetry-exporter-otlp
```

Per strumentare il codice del tuo agente

L'esempio seguente mostra come registrare l' OTel SDK, annotare le funzioni dell'agente con il `@observe` decoratore e arricchire gli intervalli con gli attributi GenAI.

```
from opentelemetry import trace
from opentelemetry.sdk.trace import TracerProvider
from opentelemetry.sdk.trace.export import BatchSpanProcessor
from opentelemetry.exporter.otlp.proto.grpc.trace_exporter import OTLPSpanExporter

# Register the tracer provider
provider = TracerProvider()
processor = BatchSpanProcessor(
    OTLPSpanExporter(endpoint="<your-osis-endpoint>")  # Use your OpenSearch Ingestion
 endpoint with SigV4
)
provider.add_span_processor(processor)
trace.set_tracer_provider(provider)

tracer = trace.get_tracer(__name__)

# Decorate your agent function
@observe
def run_agent(prompt: str):
    with tracer.start_as_current_span("invoke_agent") as span:
        span.set_attribute("gen_ai.operation.name", "invoke_agent")
        span.set_attribute("gen_ai.system", "openai")
        span.set_attribute("gen_ai.request.model", "gpt-4")

        # Enrich with token usage after the LLM call
        response = call_llm(prompt)
        span.set_attribute("gen_ai.usage.input_tokens", response.input_tokens)
        span.set_attribute("gen_ai.usage.output_tokens", response.output_tokens)
        return response
```

Note

Quando invii tracce ad Amazon OpenSearch Service, utilizza l'endpoint della pipeline OpenSearch Ingestion con l'autenticazione SigV4 anziché un endpoint localhost. Per ulteriori informazioni sulla configurazione delle pipeline di Ingestion, consulta. OpenSearch [OpenSearch Ingestione di Amazon]

Per visualizzare le tracce nell'interfaccia utente OpenSearch

Dopo che l'applicazione strumentata ha inviato i dati di traccia, puoi esplorarli nell' OpenSearch interfaccia utente. Nel tuo spazio di lavoro sull'osservabilità, espandi Discover nella barra di navigazione a sinistra e scegli Agent Traces.

Interfaccia utente di Agent Tracing

La pagina Agent Traces nell' OpenSearch interfaccia utente fornisce un'interfaccia appositamente creata per l'esplorazione, il debug e il monitoraggio delle tracce di esecuzione degli agenti LLM. Offre agli sviluppatori e agli operatori di piattaforma la piena osservabilità nelle applicazioni di intelligenza artificiale agentiche, tra cui viste di traccia gerarchiche, riepiloghi di dettaglio, visualizzazioni dei flussi e metriche aggregate.

Architecture

Il diagramma seguente mostra il flusso di dati dalle applicazioni strumentate all'interfaccia utente di Agent Traces:

```
LLM Application (with OTel SDK + GenAI instrumentation)
    |
    |  OTLP (gRPC/HTTP)
    v
OTel Collector (batch, transform)
    |
    +---- OTLP ----> OpenSearch Ingestion --> OpenSearch (otel-v1-apm-span-*)
    |
    +---- Prometheus Remote Write --> Prometheus (metrics)
                                          |
                                          v
                               OpenSearch UI
                               +-- Agent Traces Plugin
```

Prerequisiti

Prima di utilizzare Agent Traces, assicurati di disporre di quanto segue:

- Un OpenSearch cluster con dati di traccia indicizzati in `otel-v1-apm-span-*` indici.
- OpenTelemetry strumentazione con convenzioni semantiche GenAI abilitate nell'applicazione LLM.
- OpenSearch Inserimento configurato con il processore in cui inserire intervalli. `otel_trace_raw` OpenSearch

- Supporto per query PPL abilitato nell'interfaccia utente. OpenSearch

Attributi span richiesti

Agent Traces richiede attributi span specifici per visualizzare correttamente i dati di traccia. Le tabelle seguenti descrivono i campi principali e gli attributi specifici di GenAI.

Campi Core Span

Ogni intervallo deve includere i seguenti campi principali:

Campo	Tipo	Description
`traceId`	Stringa	Identificatore univoco per l'intera traccia.
`spanId`	Stringa	Identificatore univoco per questo intervallo.
`parentSpanId`	Stringa	Identificatore dell'intervallo principale. Vuoto per le estensioni principali.
`startTime`	Time stamp	Ora in cui è iniziato l'intervallo.
`endTime`	Time stamp	Ora in cui l'intervallo è terminato.
`durationInNanos`	Long	Durata dell'intervallo in nanosecondi.
`status.code`	Numero intero	Codice di stato Span (0 = non impostato, 1 = OK, 2 = errore).

Attributi GenAI

I seguenti `gen_ai.*` attributi abilitano funzionalità specifiche dell'intelligenza artificiale nell'interfaccia utente di Agent Traces:

Attributo	Valore di esempio	Description
`gen_ai.operation.name`	`chat`	Il tipo di operazione GenAI. Determina la categoria di estensione.

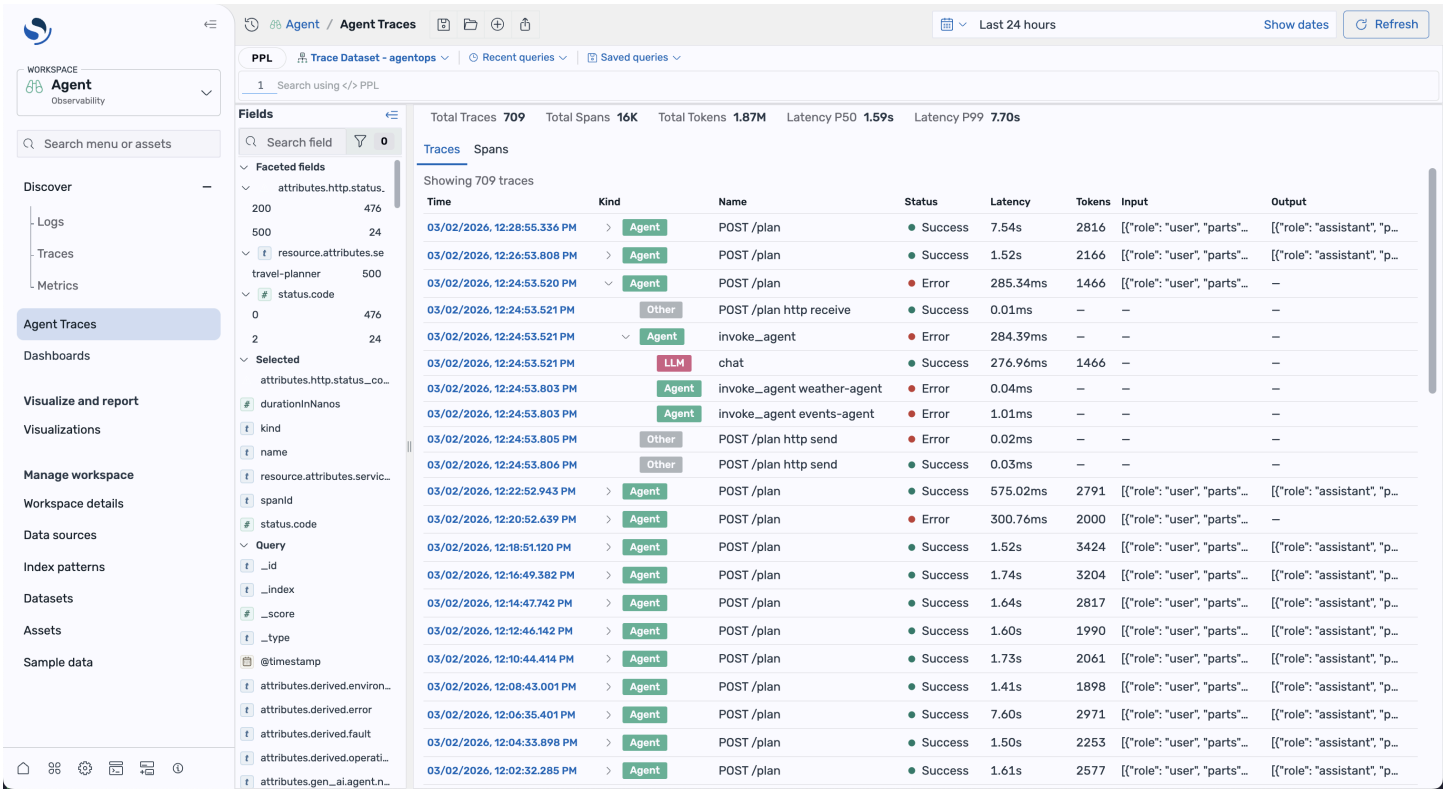
Attributo	Valore di esempio	Description
`gen_ai.system`	`openai`	Il sistema o il provider di intelligenza artificiale.
`gen_ai.request.model`	`gpt-4`	Il modello utilizzato per la richiesta.
`gen_ai.usage.input_tokens`	`150`	Numero di token di input consumati.
`gen_ai.usage.output_tokens`	`85`	Numero di token di output generati.
`gen_ai.response.finish_reasons`	`["stop"]`	Motivi per cui il modello ha smesso di generare.

Layout della pagina e barra delle metriche

La pagina Agent Traces mostra una barra delle metriche nella parte superiore che riassume le statistiche chiave su tutte le tracce visibili. Le metriche includono il numero totale di tracce, la durata media, il tasso di errore e l'utilizzo dei token. Questi valori vengono aggiornati dinamicamente in base al filtro temporale e alla query.

Scheda Tracce

La scheda Traces elenca tutte le tracce dell'agente principale che corrispondono alla query e all'intervallo di tempo correnti. Ogni riga rappresenta una singola chiamata all'agente.



La tabella seguente descrive le colonne della tabella trace:

Colonna	Description
ID di traccia	Identificatore univoco per la traccia. Scegli il link per aprire il riquadro a comparsa dei dettagli della traccia.
Nome dell'agente	Nome dell'agente che ha avviato la traccia.
Stato	Stato generale della traccia (OK o Errore).
Durata	Tempo totale dal primo all'ultimo intervallo della traccia.
Intervalli	Numero totale di intervalli nella traccia.
Token di input	Token di input totali consumati in tutte le chiamate LLM nella traccia.
Token di output	Token di output totali generati in tutte le chiamate LLM nella traccia.

Colonna	Description
Ora di inizio	Timestamp di inizio della traccia.

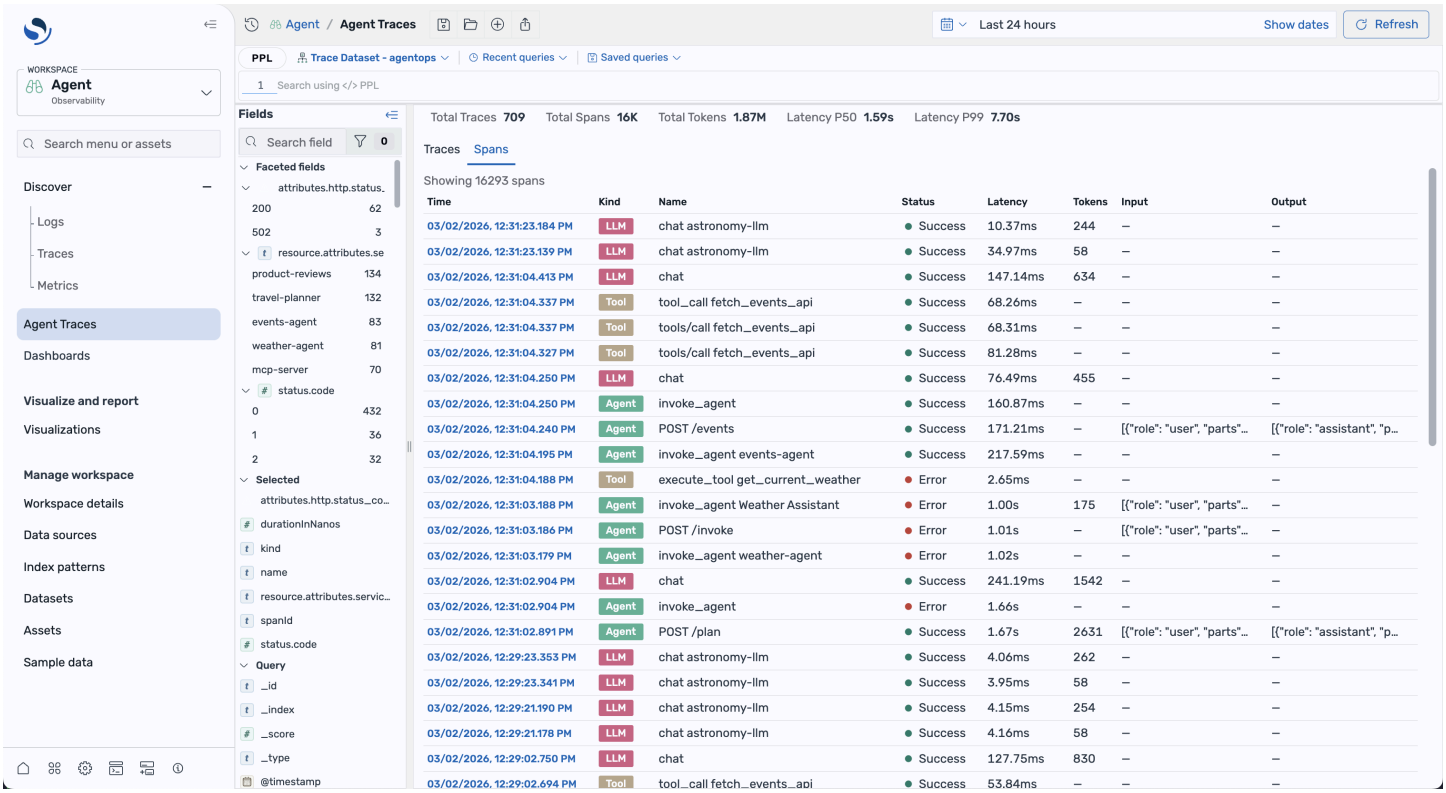
Amplia le categorie

Gli intervalli sono classificati in base all'attributo. `gen_ai.operation.name` Ogni categoria viene visualizzata con un colore e un'icona unici nell'interfaccia utente.

Nome operazione	Categoria	Description
`invoke_agent` , `create_agent`	Agente	Fase di orchestrazione degli agenti.
`chat`	LLM	Chiamata di completamento della chat LLM.
`text_completion` , `generate_content`	Contenuti	Operazione di generazione del testo.
`execute_tool`	Strumento	Invocazione dello strumento.
`embeddings`	Embedding	Generazione di incorporamento.
`retrieval`	Recupero	Operazione di recupero dei dati.

Scheda Spans

La scheda Estensioni mostra le singole estensioni su tutte le tracce. Puoi filtrare e ordinare gli intervalli per trovare operazioni specifiche.



Visualizza i dettagli di Trace

Quando scegli un ID di traccia, si apre un pannello a tendina con due aree principali:

- Pannello sinistro: visualizza l'albero di traccia, che mostra le relazioni gerarchiche padre-figlio tra gli intervalli. Include anche un DAG (grafico aciclico diretto) di flusso che visualizza il percorso di esecuzione dell'agente.
- Pannello destro: contiene due schede. La scheda Dettagli mostra gli attributi span, gli attributi delle risorse e i metadati specifici di GenAI. La scheda Cronologia mostra un grafico a cascata delle durate degli intervalli e delle relative relazioni temporali.

Interrogazione delle tracce

Agent Traces utilizza PPL (Piped Processing Language) per il recupero di tutti i dati. È possibile scrivere interrogazioni nel pannello delle interrogazioni nella parte superiore della pagina.

Per elencare le tracce principali

La seguente query restituisce le 100 tracce di root agent più recenti:

```
source = otel-v1-apm-span-*
```

```
| where parentSpanId = "" AND isnotnull(`attributes.gen_ai.operation.name`)
| sort - startTime
| head 100
```

Per recuperare tutti gli intervalli di una traccia

La seguente query restituisce lo span tree completo per una traccia specifica:

```
source = otel-v1-apm-span-*
| where traceId = "trace-id"
| head 1000
```

Per calcolare metriche aggregate

La seguente query calcola la durata media e l'utilizzo totale dei token raggruppati per modello:

```
source = otel-v1-apm-span-*
| where isnotnull(`attributes.gen_ai.request.model`)
| stats avg(durationInNanos) as avg_duration,
        sum(`attributes.gen_ai.usage.input_tokens`) as total_input_tokens,
        sum(`attributes.gen_ai.usage.output_tokens`) as total_output_tokens
    by `attributes.gen_ai.request.model`
```

Analisi di sicurezza per Amazon OpenSearch Service

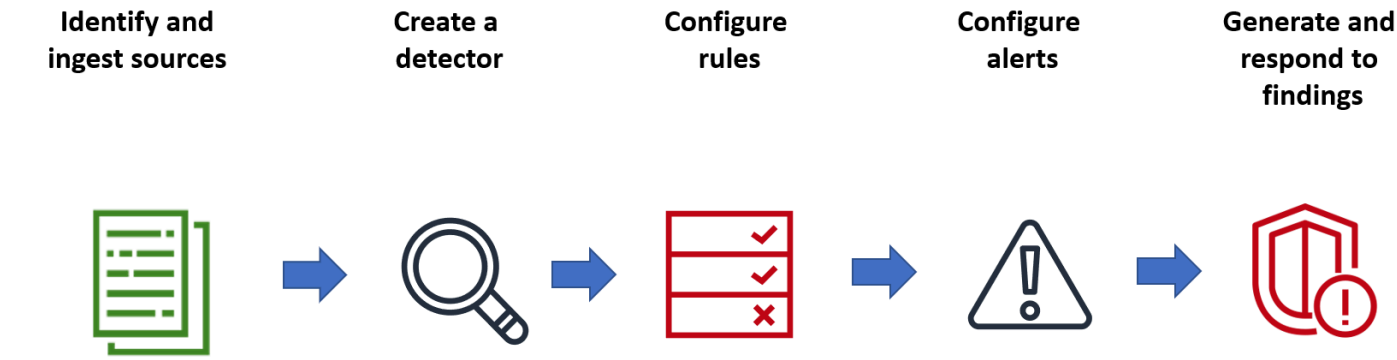
Security Analytics è una OpenSearch soluzione che offre visibilità sull'infrastruttura dell'organizzazione, monitora le attività anomale, rileva potenziali minacce alla sicurezza in tempo reale e attiva avvisi verso destinazioni preconfigurate. Puoi monitorare le attività dannose dai registri degli eventi di sicurezza valutando continuamente le regole di sicurezza e rivedendo i risultati di sicurezza generati automaticamente. Inoltre, Security Analytics può generare avvisi automatici e inviarli a un canale di notifica specifico, come Slack o e-mail.

Puoi utilizzare il plug-in Security Analytics per rilevare le minacce più comuni out-of-the-box e generare informazioni critiche sulla sicurezza dai registri degli eventi di sicurezza esistenti, come i registri del firewall, i registri di Windows e i registri di controllo dell'autenticazione. Per utilizzare Security Analytics, sul tuo dominio deve essere in esecuzione la OpenSearch versione 2.5 o successiva.

> **Note**
>
> Questa documentazione fornisce una breve panoramica di Security Analytics for Amazon OpenSearch Service. Definisce i concetti chiave e fornisce i passaggi per configurare le autorizzazioni. Per una documentazione completa, tra cui una guida alla configurazione, un riferimento all'API e un riferimento a tutte le impostazioni disponibili, consulta Security Analytics nella OpenSearch documentazione.

Componenti e concetti di analisi della sicurezza

Numerosi strumenti e funzionalità forniscono le basi per il funzionamento di Security Analytics. I componenti principali che compongono il plug-in includono rilevatori, tipi di registro, regole, risultati e avvisi.



Tipi di log

OpenSearch supporta diversi tipi di log e fornisce out-of-the-box mappature per ogni tipo. Si specifica il tipo di registro e si configura un intervallo di tempo quando si crea un rilevatore, e da lì Security Analytics attiva automaticamente un set di regole pertinenti che vengono eseguite in quell'intervallo.

Rilevatori

I rilevatori identificano una serie di minacce alla sicurezza informatica per un tipo di registro nei tuoi indici di dati. È possibile configurare il rilevatore in modo da utilizzare sia regole personalizzate che regole Sigma preconfezionate che valutano gli eventi che si verificano nel sistema. Il rilevatore genera quindi risultati di sicurezza a partire da questi eventi. Per ulteriori informazioni sui rilevatori, vedere Creazione di rilevatori nella documentazione. OpenSearch

Regole

Le regole di rilevamento delle minacce definiscono le condizioni che i rilevatori applicano ai dati di registro acquisiti per identificare un evento di sicurezza. Security Analytics supporta l'importazione, la creazione e la personalizzazione di regole per soddisfare le vostre esigenze e fornisce anche regole Sigma preconfezionate e open source per rilevare le minacce più comuni dai log. Security Analytics associa molte regole a una base di conoscenze in continua crescita di tattiche e tecniche degli avversari gestita dall'organizzazione MITRE ATT&CK. È possibile utilizzare entrambi i OpenSearch dashboard oppure creare e utilizzare regole. APIs Per ulteriori informazioni sulle regole, consulta Lavorare con le regole nella OpenSearch documentazione.

Esiti

Quando un rilevatore abbina una regola a un evento di registro, genera un risultato. Ogni risultato include una combinazione unica di regole selezionate, un tipo di registro e una gravità della regola. I risultati non indicano necessariamente minacce imminenti all'interno del sistema, ma isolano sempre un evento di interesse. Per ulteriori informazioni sui risultati, consulta Lavorare con i risultati nella OpenSearch documentazione.

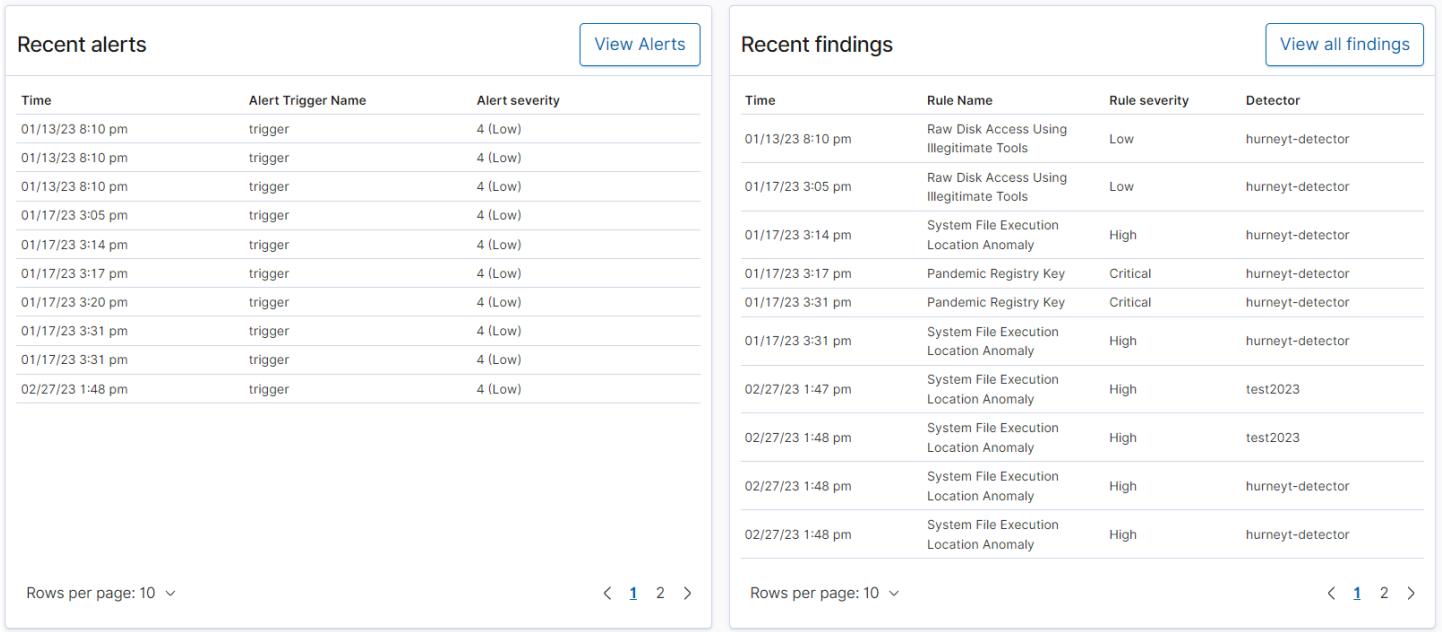
Avvisi

Quando si crea un rilevatore, è possibile specificare una o più condizioni che attivano un avviso. Un avviso è una notifica inviata a un canale preferito, come Slack o e-mail. Puoi impostare l'avviso in modo che venga attivato quando il rilevatore soddisfa una o più regole e puoi personalizzare il messaggio di notifica. Per ulteriori informazioni sugli avvisi, consulta Lavorare con gli avvisi nella documentazione. OpenSearch

Esplorazione dell'analisi della sicurezza

Puoi utilizzare OpenSearch le dashboard per visualizzare e ottenere informazioni dettagliate sul tuo plug-in Security Analytics. La visualizzazione Panoramica fornisce informazioni quali risultati e conteggio degli avvisi, scoperte e avvisi recenti, regole di rilevamento frequenti e un elenco dei rilevatori. È possibile visualizzare una visualizzazione riepilogativa composta da più visualizzazioni. Il grafico seguente, ad esempio, mostra l'andamento dei risultati e degli avvisi per vari tipi di log in un determinato periodo di tempo.

Più in basso nella pagina, puoi esaminare i risultati e gli avvisi più recenti.



Recent alerts

View Alerts

Time	Alert Trigger Name	Alert severity
01/13/23 8:10 pm	trigger	4 (Low)
01/13/23 8:10 pm	trigger	4 (Low)
01/13/23 8:10 pm	trigger	4 (Low)
01/17/23 3:05 pm	trigger	4 (Low)
01/17/23 3:14 pm	trigger	4 (Low)
01/17/23 3:17 pm	trigger	4 (Low)
01/17/23 3:20 pm	trigger	4 (Low)
01/17/23 3:31 pm	trigger	4 (Low)
01/17/23 3:31 pm	trigger	4 (Low)
02/27/23 1:48 pm	trigger	4 (Low)

Rows per page: 10 1 2

Recent findings

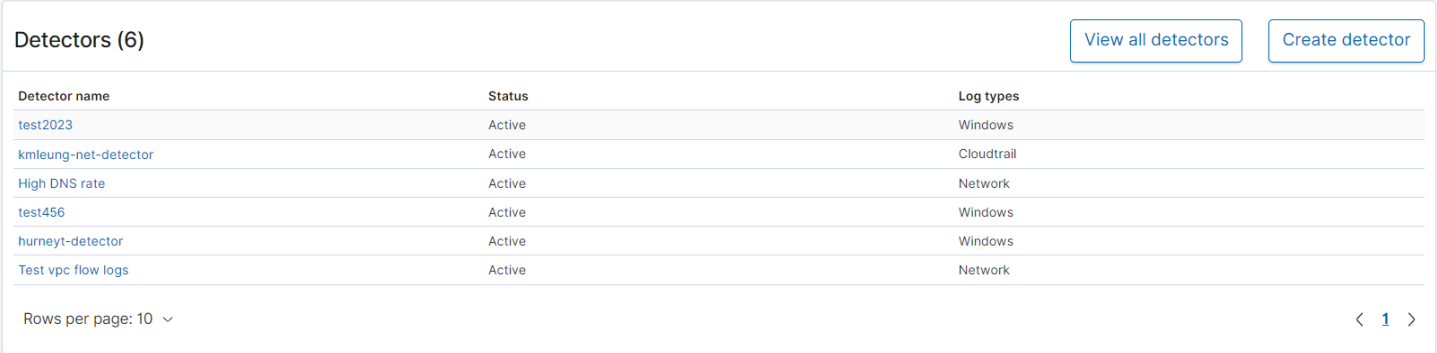
View all findings

Time	Rule Name	Rule severity	Detector
01/13/23 8:10 pm	Raw Disk Access Using Illegitimate Tools	Low	hurneyt-detector
01/17/23 3:05 pm	Raw Disk Access Using Illegitimate Tools	Low	hurneyt-detector
01/17/23 3:14 pm	System File Execution Location Anomaly	High	hurneyt-detector
01/17/23 3:17 pm	Pandemic Registry Key	Critical	hurneyt-detector
01/17/23 3:31 pm	Pandemic Registry Key	Critical	hurneyt-detector
01/17/23 3:31 pm	System File Execution Location Anomaly	High	hurneyt-detector
02/27/23 1:47 pm	System File Execution Location Anomaly	High	test2023
02/27/23 1:48 pm	System File Execution Location Anomaly	High	test2023
02/27/23 1:48 pm	System File Execution Location Anomaly	High	hurneyt-detector
02/27/23 1:48 pm	System File Execution Location Anomaly	High	hurneyt-detector

Rows per page: 10 1 2

Inoltre, puoi vedere una distribuzione delle regole attivate più frequentemente tra tutti i rilevatori attivi. Questo può aiutarti a rilevare e indagare su diversi tipi di attività dannose tra i tipi di registro.

Infine, è possibile visualizzare lo stato dei rilevatori configurati. Da questo pannello, puoi anche accedere al flusso di lavoro per la creazione del rilevatore.



Per configurare la configurazione di Security Analytics, crea delle regole con la pagina Regole e usale per scrivere i rilevatori nella pagina Rilevatori. Per una visualizzazione più mirata dei risultati di Security Analytics, puoi utilizzare le pagine Risultati e Avvisi.

Configurazione delle autorizzazioni

Se abiliti Security Analytics su un dominio di OpenSearch servizio preesistente, il `security_analytics_manager` ruolo potrebbe non essere definito nel dominio. Gli utenti senza privilegi di amministratore devono essere mappati a questo ruolo in modo da gestire gli indici a caldo sui domini che utilizzano il controllo granulare degli accessi. Per creare manualmente il ruolo `security_analytics_manager`, procedere nel seguente modo:

1. In OpenSearch Dashboard, vai su Sicurezza e scegli Autorizzazioni.
2. Scegliere Crea gruppo di operazioni e configurare i seguenti gruppi:

Group name (Nome gruppo)	Permissions
security_analytics_full_access	• cluster:admin/opensearch/securityanalytics/alerts/* • cluster:admin/opensearch/securityanalytics/detector/* • cluster:admin/opensearch/securityanalytics/findings/* • cluster:admin/opensearch/securityanalytics/mapping/* • cluster:admin/opensearch/securityanalytics/rule/*
security_analytics_read_access	• cluster:admin/opensearch/securityanalytics/alerts/get • cluster:admin/opensearch/securityanalytics/detector/get • cluster:admin/opensearch/securityanalytics/detector/search • cluster:admin/opensearch/securityanalytics/findings/get • cluster:admin/opensearch/securityanalytics/mapping/get • cluster:admin/opensearch/securityanalytics/mapping/view/get • cluster:admin/opensearch/securityanalytics/rule/get • cluster:admin/opensearch/securityanalytics/rule/search

3. Scegliere Ruoli, quindi selezionare Crea ruolo.
4. Assegna un nome al ruolo security_analytics_manager.

5. Per Autorizzazioni cluster, selezionare `security_analytics_full_access` e `security_analytics_read_access`.
6. Per Indice, digitare `*`.
7. Per le autorizzazioni di indicizzazione, seleziona e. `indices:admin/mapping/put` `indices:admin/mappings/get`
8. Scegli Create (Crea).
9. Dopo aver creato il ruolo, associalo a qualsiasi ruolo utente o di backend che gestirà gli indici di Security Analytics.

Risoluzione dei problemi

Nessun errore di indice di questo tipo

Se non hai rilevatori e apri la dashboard di Security Analytics, potresti vedere una notifica in basso a destra che dice`[index_not_found_exception] no such index [.opensearch-sap-detectors-config]`. Puoi ignorare questa notifica, che scompare nel giro di pochi secondi e non verrà più visualizzata una volta creato un rilevatore.

Best practice operative per Amazon OpenSearch Service

Questo capitolo fornisce le best practice per la gestione dei domini Amazon OpenSearch Service e include linee guida generali che si applicano a molti casi d'uso. Ogni carico di lavoro è unico, con caratteristiche uniche, quindi nessun suggerimento generico è adatto per ogni caso d'uso. La best practice più importante consiste nell'implementare, testare e ottimizzare i domini in un ciclo continuo per trovare la configurazione, la stabilità e il costo ottimali per il tuo carico di lavoro.

Argomenti

- Monitoraggio e avvisi
- Strategia di partizione
- Stabilità
- Performance
- Sicurezza
- Ottimizzazione dei costi
- CloudWatch Allarmi consigliati per Amazon Service OpenSearch
- Dimensionamento dei domini Amazon OpenSearch Service
- Scalabilità in petabyte in Amazon Service OpenSearch
- Nodi coordinatori dedicati in Amazon Service OpenSearch
- Nodi master dedicati in Amazon OpenSearch Service

Monitoraggio e avvisi

Le seguenti best practice si applicano al monitoraggio dei domini OpenSearch di servizio.

Configura CloudWatch gli allarmi

OpenSearch Il servizio invia metriche sulle prestazioni ad Amazon. CloudWatch Esamina regolarmente i parametri del cluster e dell'istanza e configura gli CloudWatch allarmi consigliati in base alle prestazioni del carico di lavoro.

Abilitazione della pubblicazione dei log

OpenSearch Il servizio espone i log degli OpenSearch errori, gli slow log di ricerca, gli slow log di indicizzazione e i log di controllo in Amazon Logs. CloudWatch I log di ricerca lenti, i log di

indicizzazione lenti e i log di errore sono utili per la risoluzione dei problemi relativi alle prestazioni e alla stabilità. I log di verifica, disponibili solo se abiliti il controllo granulare degli accessi, monitorano l'attività dell'utente. Per ulteriori informazioni, consulta Logs nella documentazione. OpenSearch

I log di ricerca lenti e i log di indicizzazione lenti sono uno strumento importante per comprendere e risolvere i problemi delle prestazioni delle operazioni di ricerca e indicizzazione. Abilita la distribuzione di log lenti di ricerca e di indice per tutti i domini di produzione. È inoltre necessario configurare le soglie di registrazione, altrimenti i registri CloudWatch non verranno acquisiti.

Strategia di partizione

Gli shard distribuiscono il carico di lavoro tra i nodi di dati del dominio di servizio. OpenSearch Gli indici configurati correttamente possono contribuire a migliorare le prestazioni complessive del dominio.

Quando invii dati a OpenSearch Service, invii tali dati a un indice. Un indice è analogo a una tabella di database, con documenti come le righe e campi come le colonne. Quando crei l'indice, dici OpenSearch quanti shard primari vuoi creare. Gli shard primari sono partizioni indipendenti dell'intero set di dati. OpenSearch Il servizio distribuisce automaticamente i dati tra gli shard primari di un indice. Puoi inoltre configurare le repliche dell'indice. Ogni partizione di replica comprende un set completo di copie delle partizioni primarie per quell'indice.

OpenSearch Il servizio mappa gli shard per ogni indice tra i nodi di dati del cluster. Questo assicura che le partizioni primarie e di replica per l'indice risiedano su nodi di dati diversi. La prima replica garantisce la presenza di due copie dei dati nell'indice. Dovresti sempre usare almeno una replica. Le repliche aggiuntive forniscono ridondanza e capacità di lettura aggiuntive.

OpenSearch invia richieste di indicizzazione a tutti i nodi di dati che contengono frammenti che appartengono all'indice. Invia le richieste di indicizzazione prima ai nodi di dati che contengono partizioni primarie e poi ai nodi di dati che contengono partizioni di replica. Le richieste di ricerca vengono indirizzate dal nodo coordinatore a una partizione primaria o di replica per tutte le partizioni appartenenti all'indice.

Ad esempio, per un indice con cinque partizioni primarie e una replica, ogni richiesta di indicizzazione interessa 10 partizioni. Le richieste di ricerca, invece, vengono inviate a n partizioni, dove n è il numero di partizioni primarie. Per un indice con cinque partizioni primarie e una replica, ogni query di ricerca interessa cinque partizioni (primarie o di replica) da quell'indice.

Determinazione del numero di partizioni e di nodi di dati

Utilizza le seguenti best practice per determinare il numero di partizioni e nodi di dati per il tuo dominio.

Dimensione delle partizioni: la dimensione dei dati su disco è il risultato diretto della dimensione dei dati di origine e cambia man mano che indicizzi più dati. Il source-to-index rapporto può variare notevolmente, da 1:10 a 10:1 o più, ma in genere è di circa 1:1,10. Puoi utilizzare questo rapporto per prevedere la dimensione dell'indice su disco. Puoi inoltre indicizzare alcuni dati e recuperare le dimensioni effettive dell'indice per determinare il rapporto per il tuo carico di lavoro. Una volta ottenuta una dimensione dell'indice prevista, imposta un numero di partizioni in modo che ogni partizione abbia una dimensione compresa tra 10 e 30 GiB (per i carichi di lavoro di ricerca) o tra 30 e 50 GiB (per i carichi di lavoro dei log). 50 GiB dovrebbe essere il massimo; assicurati di pianificare la crescita.

Conteggio partizioni: la distribuzione delle partizioni sui nodi di dati ha un grande impatto sulle prestazioni di un dominio. Quando hai indici con più shard, prova a fare in modo che gli shard contino un multiplo del conteggio dei nodi di dati. Ciò aiuta a garantire che le partizioni siano distribuite uniformemente tra i nodi di dati e previene i nodi ad accesso frequente. Ad esempio, se disponi di 12 partizioni primarie, il numero di nodi di dati deve essere 2, 3, 4, 6 o 12. Tuttavia, il numero delle partizioni è secondario rispetto alla dimensione delle partizioni; se disponi di 5 GiB di dati, dovresti comunque usare una singola partizione.

Partizioni per nodo di dati: il numero totale di partizioni che un nodo può contenere è proporzionale alla memoria heap JVM (Java Virtual Machine) del nodo. Punta a 25 partizioni o meno per GiB di memoria heap. Ad esempio, un nodo con 32 GiB di memoria heap non deve contenere più di 800 partizioni. Sebbene la distribuzione degli shard possa variare in base ai modelli di carico di lavoro, esiste un limite di 1.000 shard per nodo per Elasticsearch e da OpenSearch 1,1 a 2,15 e 4.000 per 2.17 e versioni successive. OpenSearch L'API cat/allocation fornisce una visualizzazione rapida del numero di partizioni e dello storage totale delle partizioni tra i nodi di dati.

Rapporto partizione/CPU: quando una partizione è coinvolta in una richiesta di indicizzazione o ricerca, utilizza una vCPU per elaborare la richiesta. Come best practice, utilizza un punto di scala iniziale di 1,5 vCPU per partizione. Se il tuo tipo di istanza ha 8 vCPUs, imposta il conteggio dei nodi di dati in modo che ogni nodo non abbia più di sei shard. Si tratta di un'approssimazione. Assicurati di testare il carico di lavoro e dimensionare il cluster di conseguenza.

Per suggerimenti sul volume di storage, la dimensione delle partizioni e il tipo di istanza, consulta le seguenti risorse:

- the section called “Dimensionamento dei domini”
- the section called “Scala in petabyte”

Evitare l'asimmetria di storage

L'asimmetria di archiviazione si verifica quando uno o più nodi all'interno di un cluster detengono una percentuale maggiore di archiviazione per uno o più indici rispetto agli altri. L'utilizzo non uniforme della CPU, la latenza intermittente e irregolare e l'accodamento non uniforme tra i nodi di dati sono tutte indicazioni di asimmetria di archiviazione. Per determinare se sono presenti problemi di asimmetria, consulta le seguenti sezioni sulla risoluzione dei problemi:

- the section called “Asimmetria di partizioni e storage di nodi”
- the section called “Asimmetria di partizioni e storage di indici”

Stabilità

Le seguenti best practice si applicano al mantenimento di un dominio di OpenSearch servizio stabile e integro.

Resta aggiornato con OpenSearch

Aggiornamenti del software del servizio

OpenSearch Il servizio rilascia regolarmente aggiornamenti software che aggiungono funzionalità o migliorano in altro modo i tuoi domini. Gli aggiornamenti non modificano né la versione del motore Elasticsearch OpenSearch né quella del motore Elasticsearch. Ti consigliamo di pianificare un orario ricorrente per eseguire l'operazione dell'DescribeDomainAPI e di avviare un aggiornamento del software di servizio, se lo è. `UpdateStatus` `ELIGIBLE` Se non aggiorni il dominio entro un determinato periodo di tempo (in genere due settimane), il OpenSearch Servizio esegue automaticamente l'aggiornamento.

OpenSearch aggiornamenti di versione

OpenSearch Il servizio aggiunge regolarmente il supporto per le versioni gestite dalla community di. OpenSearch Effettua sempre l'upgrade alle OpenSearch versioni più recenti quando sono disponibili.

OpenSearch Il servizio aggiorna contemporaneamente entrambe le OpenSearch OpenSearch dashboard (o Elasticsearch e Kibana se il dominio utilizza un motore legacy). Se il cluster dispone di

nodi master dedicati, gli aggiornamenti vengono completati senza tempi di inattività. In caso contrario, il cluster potrebbe non rispondere per diversi secondi dopo l'aggiornamento mentre elegge un nodo principale. OpenSearch I dashboard potrebbero non essere disponibili durante alcuni o tutti gli upgrade.

Ci sono due modi per aggiornare un dominio:

- Aggiornamento in loco: questa opzione è più semplice perché si mantiene lo stesso cluster.
- Aggiornamento di snapshot/ripristino: questa opzione è ideale per testare nuove versioni su un nuovo cluster o per eseguire la migrazione tra cluster

Indipendentemente dal processo di aggiornamento utilizzato, ti consigliamo di mantenere un dominio destinato esclusivamente allo sviluppo e al test e di aggiornarlo alla nuova versione prima di aggiornare il dominio di produzione. Quando crei il dominio di test, scegli Development and testing (Sviluppo e test) per il tipo di implementazione. Assicurati di aggiornare tutti i client alle versioni compatibili subito dopo l'aggiornamento del dominio.

Migliora le prestazioni delle istantanee

Per evitare che l'istantanea rimanga bloccata durante l'elaborazione, il tipo di istanza per il nodo master dedicato deve corrispondere al numero di shard. Per ulteriori informazioni, consulta the section called “Scelta dei tipi di istanza per nodi principali dedicati”. Inoltre, ogni nodo non deve avere più dei 25 shard consigliati per GiB di memoria heap Java. Per ulteriori informazioni, consulta the section called “Scelta del numero di partizioni”.

Abilitare nodi principali dedicati

I nodi principali dedicati migliorano la stabilità del cluster. Un nodo principale dedicato esegue attività di gestione del cluster ma non conserva dati di indice né risponde a richieste del client. Questo offload delle attività di gestione del cluster aumenta la stabilità del dominio e consente di apportare modifiche alla configurazione senza avere tempi di inattività.

Abilita e utilizza tre nodi principali dedicati per una stabilità ottimale del dominio in tre zone di disponibilità. L'implementazione con Multi-AZ con Standby configura tre nodi master dedicati per te. Per suggerimenti sul tipo di istanza, consulta the section called “Scelta dei tipi di istanza per nodi principali dedicati”.

Esecuzione dell'implementazione in più zone di disponibilità

Per prevenire la perdita di dati e ridurre al minimo i tempi di inattività del cluster in caso di interruzione del servizio, puoi distribuire i nodi tra due o tre zone di disponibilità nella stessa Regione AWS. La migliore pratica consiste nell'implementare Multi-AZ with Standby, che configura tre zone di disponibilità, con due zone attive e una che funge da standby, e con due shard di replica per indice. Questa configurazione consente a OpenSearch Service di distribuire gli shard di replica su shard primari diversi dai corrispondenti shard primari. AZs Non sono previsti costi di trasferimento dei dati tra zone di disponibilità per le comunicazioni dei cluster tra zone di disponibilità.

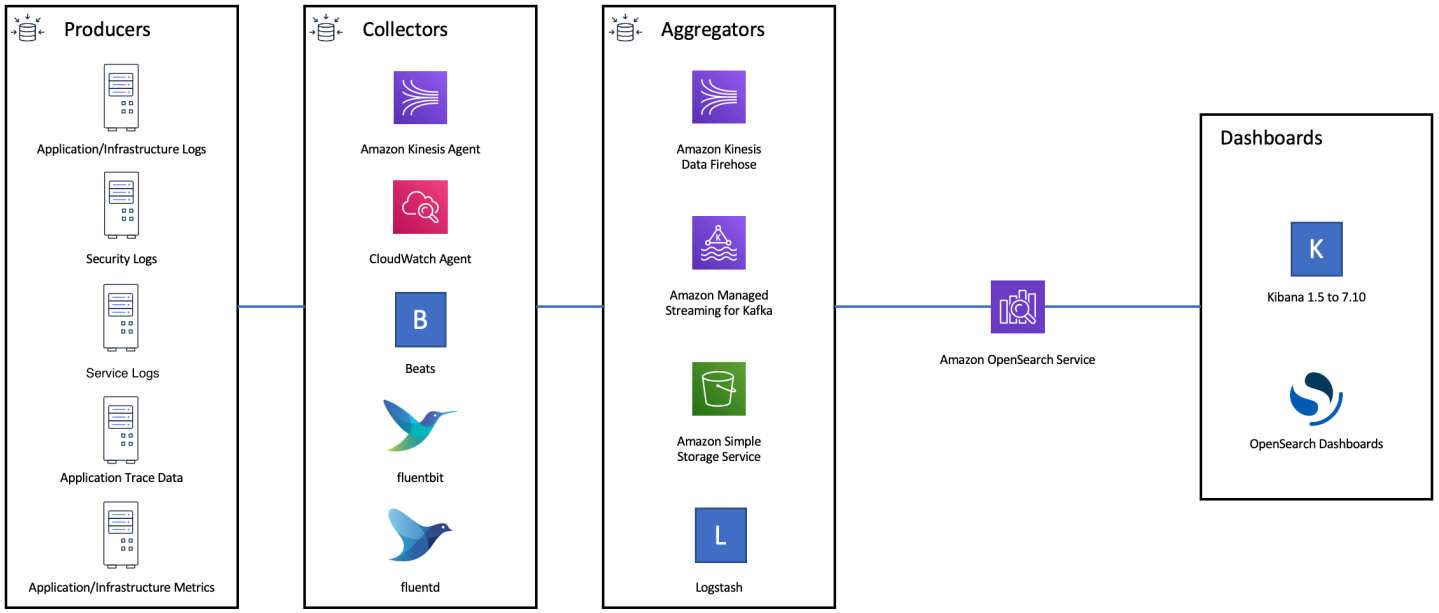
Le zone di disponibilità sono più posizioni isolate all'interno di ogni regione . Con una configurazione a due zone di disponibilità, perdere una zona di disponibilità significa perdere metà della capacità totale del dominio. Il passaggio a tre zone di disponibilità riduce ulteriormente l'impatto della perdita di una singola zona di disponibilità.

Controllo del flusso di importazione e del buffering

Consigliamo di limitare il numero complessivo delle richieste utilizzando l'operazione API _bulk. È più efficace inviare una richiesta `_bulk` contenente 5.000 documenti anziché inviare 5.000 richieste contenenti un singolo documento.

Per una stabilità operativa ottimale, a volte è necessario limitare o addirittura sospendere il flusso upstream delle richieste di indicizzazione. Limitare la percentuale di richieste di indicizzazione è un meccanismo importante per gestire picchi imprevisti o occasionali nelle richieste che potrebbero altrimenti sovraccaricare il cluster. Valuta la possibilità di creare un meccanismo di controllo del flusso nella tua architettura upstream.

Il diagramma seguente mostra più opzioni di componenti per un'architettura di importazione dei log. Configura il livello di aggregazione per avere spazio sufficiente per il buffering dei dati in ingresso per picchi di traffico improvvisi e brevi interventi di manutenzione del dominio.



Creare mappature per i carichi di lavoro di ricerca

Per i carichi di lavoro di ricerca, crea mappature che definiscono il modo in cui OpenSearch archivia e indicizza i documenti e i relativi campi. Imposta `dynamic` su `strict` per evitare l'aggiunta accidentale di nuovi campi.

```
PUT my-index
{
  "mappings": {
    "dynamic": "strict",
    "properties": {
      "title": { "type" : "text" },
      "author": { "type" : "integer" },
      "year": { "type" : "text" }
    }
  }
}
```

Utilizzare modelli di indici

È possibile utilizzare un modello di indice per spiegare OpenSearch come configurare un indice al momento della creazione. Configura i modelli di indice prima di creare gli indici. Quindi, quando crei un indice, l'indice eredita le impostazioni e le mappature dal modello. Puoi applicare più di un modello a un singolo indice, quindi puoi specificare le impostazioni in un modello e le mappature in un altro.

Questa strategia consente un modello per le impostazioni comuni su più indici e modelli separati per impostazioni e mappature più specifiche.

Le seguenti impostazioni sono particolarmente utili per la configurazione nei modelli:

- Numero di partizioni primarie e di replica
- Intervallo di aggiornamento (con quale frequenza aggiornare e rendere disponibili per la ricerca le modifiche recenti all'indice)
- Controllo dinamico delle mappature
- Mappature di campi esplicite

Il seguente modello di esempio contiene ognuna di queste impostazioni:

```
{
   "index_patterns":[
      "index-*"
   ],
   "order": 0,
   "settings": {
      "index": {
         "number_of_shards": 3,
         "number_of_replicas": 1,
         "refresh_interval": "60s"
      }
   },
   "mappings": {
      "dynamic": false,
      "properties": {
         "field_name1": {
            "type": "keyword"
         }
      }
   }
}
```

Anche se cambiano raramente, avere impostazioni e mappature definite centralmente OpenSearch è più semplice da gestire rispetto all'aggiornamento di più client upstream.

Gestire gli indici con Index State Management

Se gestisci log o dati di serie temporali, ti consigliamo di utilizzare Index State Management (ISM). ISM consente di automatizzare attività regolari di gestione del ciclo di vita degli indici. Con ISM, puoi creare policy che attivano i rollover degli alias degli indici, eseguire snapshot degli indici, spostare gli indici tra i livelli di archiviazione ed eliminare gli indici precedenti. Puoi persino usare le operazioni di rollover di ISM come strategia alternativa di gestione del ciclo di vita dei dati per evitare asimmetrie di partizioni.

In primo luogo, configura una policy ISM. Per un esempio, consulta the section called “Policy di esempio”. Quindi, collega la policy a uno o più indici. Se includi un campo modello ISM nella policy, OpenSearch Service applica automaticamente la policy a qualsiasi indice che corrisponda al modello specificato.

Rimuovere gli indici inutilizzati

Esamina regolarmente gli indici nel tuo cluster e identifica quelli che non sono in uso. Esegui uno snapshot di questi indici in modo che siano archiviati in S3, quindi eliminali. Quando rimuovi gli indici inutilizzati, riduci il numero di partizioni e si consente una distribuzione dell'archiviazione e un utilizzo delle risorse più bilanciati tra i nodi. Anche quando sono inattivi, gli indici consumano alcune risorse durante le attività interne di manutenzione degli indici.

Anziché eliminare manualmente gli indici inutilizzati, puoi utilizzare ISM per eseguire automaticamente un'istantanea ed eliminare gli indici dopo un certo periodo di tempo.

Utilizzare più domini per un'elevata disponibilità

Per ottenere un'elevata disponibilità oltre un tempo di attività del 99,9% in più regioni, prendi in considerazione l'utilizzo di due domini. Per set di dati piccoli o che cambiano lentamente, puoi configurare la replica tra cluster in modo da mantenere un modello attivo-passivo. In questo modello è scritto solo il dominio leader, da cui possono essere letti entrambi i domini. Per set di dati più grandi e dati che cambiano rapidamente, configura la distribuzione doppia nella pipeline di importazione, in modo che tutti i dati vengano scritti indipendentemente in entrambi i domini in un modello attivo-attivo.

Progetta le tue applicazioni upstream e downstream tenendo presente il failover. Assicurati di testare il processo di failover insieme ad altri processi di ripristino di emergenza.

Performance

Le seguenti best practice si applicano all'ottimizzazione dei domini per ottenere prestazioni ottimali.

Ottimizzare le dimensioni e la compressione delle richieste in blocco

Il dimensionamento in blocco dipende dai dati, dall'analisi e dalla configurazione del cluster, ma un buon punto di partenza è 3-5 MiB per ciascuna richiesta in blocco.

Invia richieste e ricevi risposte dai tuoi OpenSearch domini utilizzando la compressione gzip per ridurre la dimensione del payload di richieste e risposte. Puoi usare la compressione gzip con il client OpenSearch Python o includendo le seguenti intestazioni dal lato client:

- `'Accept-Encoding': 'gzip'`
- `'Content-Encoding': 'gzip'`

Per ottimizzare le dimensioni delle richieste in blocco, inizia con una richiesta in blocco di 3 MiB. Quindi, aumentane lentamente la dimensione finché le prestazioni di indicizzazione non smettono di migliorare.

> **Note**
>
> Per abilitare la compressione gzip sui domini che eseguono Elasticsearch versione 6.x, devi impostare `http_compression.enabled` a livello di cluster. Questa impostazione è vera per impostazione predefinita nelle versioni 7.x di Elasticsearch e in tutte le versioni di. OpenSearch

Ridurre le dimensioni delle risposte alle richieste in blocco

Per ridurre la dimensione delle OpenSearch risposte, escludi i campi non necessari con il parametro. `filter_path` Assicurati di non escludere i campi necessari per identificare o provare nuovamente le richieste non riuscite. Per maggiori informazioni ed esempi, consulta the section called “Riduzione delle dimensioni della risposta”.

Ottimizzare gli intervalli di aggiornamento

OpenSearch gli indici alla fine hanno una consistenza di lettura. Un'operazione di aggiornamento rende disponibili per la ricerca tutti gli aggiornamenti eseguiti su un indice. L'intervallo di

aggiornamento predefinito è di un secondo, il che significa che OpenSearch esegue un aggiornamento ogni secondo durante la scrittura di un indice.

Meno frequentemente si aggiorna un indice (intervallo di aggiornamento più elevato), migliori sono le prestazioni complessive di indicizzazione. L'aumento dell'intervallo di aggiornamento comporta un ritardo maggiore tra l'aggiornamento dell'indice e la disponibilità dei nuovi dati per la ricerca. Imposta l'intervallo di aggiornamento al livello più alto che puoi tollerare per migliorare le prestazioni complessive.

Consigliamo di impostare il parametro `refresh_interval` per tutti gli indici su 30 secondi o più.

Abilitare la regolazione automatica

Auto-Tune utilizza i parametri di prestazioni e utilizzo del OpenSearch cluster per suggerire modifiche alle dimensioni delle code, alle dimensioni della cache e alle impostazioni della macchina virtuale Java (JVM) sui nodi. Queste modifiche facoltative migliorano la velocità e la stabilità del cluster. È possibile ripristinare le impostazioni predefinite OpenSearch del servizio in qualsiasi momento. La regolazione automatica è abilitata per impostazione predefinita sui nuovi domini a meno che non venga esplicitamente disabilitata.

Consigliamo di abilitare la regolazione automatica su tutti i domini e di impostare una finestra di manutenzione ricorrente o di rivedere periodicamente i relativi suggerimenti.

Sicurezza

Le seguenti best practice si applicano alla protezione dei domini.

Abilitare il controllo granulare degli accessi

Il controllo granulare degli accessi consente di controllare chi può accedere a determinati dati all'interno di un dominio di servizio. OpenSearch Rispetto al controllo degli accessi generalizzato, il controllo granulare degli accessi fornisce a ciascun cluster, indice, documento e campo la propria policy specifica per l'accesso. I criteri di accesso possono essere basati su una serie di fattori, tra cui il ruolo della persona che richiede l'accesso e l'azione che intende eseguire sui dati. Ad esempio, potresti concedere a un utente l'accesso per scrivere su un indice, mentre a un altro potrebbe essere concesso l'accesso solo per leggere i dati sull'indice senza apportare alcuna modifica.

Il controllo granulare degli accessi consente ai dati con requisiti di accesso diversi di esistere nello stesso spazio di storage senza incorrere in problemi di sicurezza o conformità.

Si consiglia di abilitare il controllo granulare degli accessi sui domini.

Distribuire domini all'interno di un VPC

Posizionare il dominio del OpenSearch servizio all'interno di un cloud privato virtuale (VPC) consente una comunicazione sicura tra il OpenSearch Servizio e altri servizi all'interno del VPC, senza la necessità di un gateway Internet, un dispositivo NAT o una connessione VPN. Tutto il traffico rimane in modo sicuro all'interno del Cloud. AWS Grazie a loro isolamento logico, i domini che si trovano all'interno di un VPC hanno un ulteriore livello di sicurezza rispetto ai domini che usano gli endpoint pubblici.

Consigliamo di creare domini all'interno di un VPC.

Applicare una policy di accesso restrittiva

Anche se il tuo dominio è implementato all'interno di un VPC, è meglio implementare la sicurezza a più livelli. Assicurati di controllare la configurazione delle tue attuali policy di accesso.

Applica una politica di accesso restrittiva basata sulle risorse ai tuoi domini e segui il principio del privilegio minimo quando concedi l'accesso all'API di configurazione e alle operazioni dell'API. OpenSearch Come regola generale, evita di utilizzare il principale utente anonimo`"Principal": {"AWS": "*" }` nelle policy di accesso.

Esistono tuttavia alcune situazioni in cui è accettabile utilizzare una policy di accesso aperto, ad esempio quando abiliti il controllo degli accessi granulare. Una policy di accesso aperto può consentirti di accedere al dominio nei casi in cui la firma delle richieste è difficile o impossibile, ad esempio da determinati client e strumenti.

Abilitare la crittografia dei dati a riposo

OpenSearch I domini di servizio offrono la crittografia dei dati inattivi per impedire l'accesso non autorizzato ai dati. Encryption at rest utilizza AWS Key Management Service (AWS KMS) per archiviare e gestire le chiavi di crittografia e l'algoritmo Advanced Encryption Standard con chiavi a 256 bit (AES-256) per eseguire la crittografia.

Se il dominio archivia i dati sensibili, abilita la crittografia dei dati a riposo.

Abilita la crittografia node-to-node

Node-to-node la crittografia fornisce un ulteriore livello di sicurezza oltre alle funzionalità di sicurezza predefinite di OpenSearch Service. Implementa Transport Layer Security (TLS) per tutte le

comunicazioni tra i nodi che vengono forniti all'interno. OpenSearch Node-to-nodecrittografia, tutti i dati inviati al dominio del OpenSearch servizio tramite HTTPS rimangono crittografati in transito mentre vengono distribuiti e replicati tra i nodi.

Se il tuo dominio memorizza dati sensibili, abilita node-to-node la crittografia.

Monitora con AWS Security Hub CSPM

Monitora l'utilizzo del OpenSearch Servizio in relazione alle migliori pratiche di sicurezza utilizzando AWS Security Hub CSPM. Security Hub CSPM utilizza i controlli di sicurezza per valutare le configurazioni delle risorse e gli standard di sicurezza per aiutarti a rispettare vari framework di conformità. Per ulteriori informazioni sull'utilizzo di Security Hub CSPM per valutare le risorse OpenSearch del servizio, vedere Amazon OpenSearch Service i controlli nella Guida per l'AWS Security Hub utente.

Ottimizzazione dei costi

Le seguenti best practice si applicano all'ottimizzazione e al risparmio sui costi del OpenSearch Servizio.

Utilizzare tipi di istanza di ultima generazione

OpenSearch Service adotta sempre nuovi tipi di istanze Amazon EC2 che offrono prestazioni migliori a un costo inferiore. Si consiglia di utilizzare sempre istanze di ultima generazione.

Non utilizzare istanze T2 o `t3.small` per i domini di produzione in quanto possono diventare instabili in presenza di carichi pesanti sostenuti. Le istanze `r6g.large` sono un'opzione per piccoli carichi di lavoro di produzione (sia come nodi di dati che come nodi principali dedicati).

Utilizzo dei volumi gp3 di Amazon EBS più recenti

OpenSearch i nodi di dati richiedono uno storage a bassa latenza e ad alto throughput per fornire indicizzazione e query rapide. Utilizzando i volumi gp3 Amazon EBS, ottieni prestazioni di base (IOPS e velocità di trasmissione effettiva) più elevate a un costo inferiore del 9,6% rispetto al tipo di volume gp2 Amazon EBS offerto in precedenza. È possibile fornire IOPS evelocità di trasmissione effettiva aggiuntivi indipendentemente dalle dimensioni del volume tramite gp3. Questi volumi sono inoltre più stabili rispetto ai volumi della generazione precedente in quanto non utilizzano crediti di espansione. Il tipo di volume gp3 raddoppia anche i limiti di dimensione del volume del tipo di per-data-node volume

gp2. Con questi volumi più grandi, è possibile ridurre il costo dei dati passivi aumentando la quantità di spazio di archiviazione per nodo di dati.

Utilizzo UltraWarm e conservazione a freddo dei dati di registro delle serie temporali

Se li utilizzi OpenSearch per l'analisi dei log, trasferisci i dati in una cella frigorifera per ridurre i costi. UltraWarm Utilizza Index State Management (ISM) per migrare i dati tra livelli di storage e gestire la conservazione dei dati.

UltraWarmoffre un modo conveniente per archiviare grandi quantità di dati di sola lettura in Service. OpenSearch UltraWarm utilizza Amazon S3 per lo storage, il che significa che i dati sono immutabili e ne è necessaria solo una copia. Paghi solo per lo spazio di archiviazione equivalente alla dimensione delle partizioni primarie negli indici. Le latenze per le UltraWarm query crescono con la quantità di dati S3 necessari per soddisfare la query. Dopo che i dati sono stati memorizzati nella cache dei nodi, le query sugli indici hanno prestazioni simili alle query sugli UltraWarm indici caldi.

Lo storage a freddo è supportato anche da S3. Quando è necessario interrogare dati non aggiornati, è possibile collegarli in modo selettivo ai nodi esistenti. UltraWarm I cold data comportano gli stessi costi di storage gestito UltraWarm, ma gli oggetti in cold storage non consumano le risorse dei UltraWarm nodi. Pertanto, la conservazione a freddo offre una notevole capacità di archiviazione senza influire sulle dimensioni o sul numero dei UltraWarm nodi.

UltraWarm diventa conveniente quando si dispone di circa 2,5 TiB di dati da migrare dallo storage a caldo. Monitora il tasso di riempimento e pianifica di spostare gli indici UltraWarm prima di raggiungere quel volume di dati.

Rivedere i suggerimenti per le istanze riservate

Prendi in considerazione l'acquisto di istanze riservate (RIs) dopo aver stabilito una buona base per quanto riguarda le prestazioni e il consumo di elaborazione. Gli sconti partono dal 30% circa per prenotazioni di 1 anno senza anticipo e possono aumentare fino al 50% per tutti gli impegni anticipati di 3 anni.

Dopo aver osservato un funzionamento stabile per almeno 14 giorni, consulta Accedere ai consigli di prenotazione nella Guida per l'AWS Cost Management utente. L'intestazione Amazon OpenSearch Service mostra consigli di acquisto del RI specifici e risparmi previsti.

CloudWatch Allarmi consigliati per Amazon Service OpenSearch

CloudWatch gli allarmi eseguono un'azione quando una CloudWatch metrica supera un valore specificato per un certo periodo di tempo. Ad esempio, potresti voler AWS inviarti un'e-mail se lo stato di salute del cluster dura più `red` di un minuto. Questa sezione include alcuni allarmi consigliati per Amazon OpenSearch Service e come rispondere ad essi.

Puoi distribuire automaticamente questi allarmi utilizzando. CloudFormationPer uno stack di esempio, consulta il relativo repository. GitHub

> **Note**
>
> Se distribuisci lo CloudFormation stack, gli `KMSKeyInaccessible` allarmi `KMSKeyError` and esisteranno in `Insufficient Data` uno stato perché queste metriche vengono visualizzate solo se un dominio riscontra un problema con la sua chiave di crittografia.

Per ulteriori informazioni sulla configurazione degli allarmi, consulta Creating Amazon CloudWatch Alarms nella Amazon CloudWatch User Guide.

Allarme	Problema
`ClusterStatus.red` il massimo è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	Almeno una partizioni primaria e le relative repliche non sono assegnate a un nodo. Per informazioni, consulta the section called “Cluster in stato rosso”.
`ClusterStatus.yellow` il massimo è >= 1 per 1 minuto, 5 volte consecutive	Almeno una partizione di replica non è allocata per un nodo. Per informazioni, consulta the section called “Stato giallo del cluster”.
`FreeStorageSpace` minimo è <= 20480 per 1 minuto, 1	Un nodo nel cluster è legato ai 20 GiB di spazio di archiviazione gratuito. Per informazioni, consulta the section called “Mancanza di spazio di archiviazione disponibile”. Tale valore viene espresso in MiB, perciò

Allarme	Problema
periodo di tempo consecutivo	anziché su 20480, consigliamo di impostarlo al 25% dello spazio di archiviazione per ogni nodo.
`ClusterIndexWritesBlocked` è >= 1 per 5 minuti, 1 periodo di tempo consecutivo	Il cluster sta bloccando le richieste di scrittura. Per informazioni, consulta the section called “ClusterBlockException”.
`Nodes` minimo è x per 1 giorno, 1 periodo di tempo consecutivo	x è il numero di nodi del cluster. Questo allarme indica che almeno un nodo nel cluster è stato irraggiungibile per un giorno. Per informazioni, consulta the section called “Nodi cluster con errori”.

`AutomatedSnapshotFailure` il massimo è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	Uno snapshot automatico ha restituito un errore. Questo errore è spesso il risultato di uno stato del cluster rosso. Per informazioni, consulta the section called “Cluster in stato rosso”.
	
	Per un riepilogo di tutti gli snapshot automatici e alcune informazioni sui fallimenti, è possibile provare una delle seguenti richieste:
	
	`GET domain_endpoint /_snapshot/cs-automated/_all`
	`GET domain_endpoint /_snapshot/cs-automated-enc/_all`

`CPUUtilization` o `WarmCPUUtilization` massimo è >= 80% per 15 minuti, 3 periodi di tempo consecutivi	A volte può verificarsi un utilizzo della CPU al 100%, ma un uso elevato e sostenuto può rappresentare un problema. Consigliamo di utilizzare tipi di istanza più grandi o aggiungere istanze.
`JVMMemoryPressure` il massimo è >= 95% per 1 minuto, 3 volte consecutive	Sui cluster potrebbero verificarsi errori di esaurimento della memoria nel caso in cui l'utilizzo aumenti. Prendi in considerazione la possibilità di scalare verticalmente. OpenSearch Il servizio utilizza metà della RAM di un'istanza per l'heap Java, fino a una dimensione dell'heap di 32 GiB. Puoi scalare le istanze verticalmente fino a 64 GiB di RAM e poi scalare orizzontalmente aggiungendo le istanze.

Allarme	Problema
`OldGenJVM MemoryPressure` il massimo è >= 80% per 1 minuto, 3 volte consecutive	
`MasterCPU Utilization` il massimo è >= 50% per 15 minuti, 3 periodi di tempo consecutivi	Potresti utilizzare tipi di istanza di dimensioni maggiori per i tuoi nodi master dedicati. A causa del loro ruolo nella stabilità del cluster e nelle distribuzioni blu/verde, i nodi master dedicati devono avere un utilizzo ridotto della CPU rispetto ai nodi di dati.
`MasterJVM MemoryPressure` il massimo è >= 95% per 1 minuto, 3 volte consecutive	
`MasterOld GenJVMMem oryPressure` il massimo è >= 80% per 1 minuto, 3 volte consecutive	
`KMSKeyError` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	La chiave di AWS KMS crittografia utilizzata per crittografare i dati inattivi nel dominio è disabilitata. Riabilitala per ripristinare le normali operazioni. Per ulteriori informazioni, consultare the section called “Crittografia dei dati a riposo”.
`KMSKeyIna ccessible` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	La chiave di AWS KMS crittografia utilizzata per crittografare i dati archiviati nel dominio è stata eliminata o ha revocato le sue concessioni al Servizio. OpenSearch Non è possibile recuperare i domini che sono in questo stato. Se hai uno snapshot manuale, puoi utilizzarlo per migrare a un nuovo dominio. Per ulteriori informazioni, consulta the section called “Crittografia dei dati a riposo”.

Allarme	Problema
`shards.active` è >= 30000 per 1 minuto, 1 periodo di tempo consecutivo	Il numero totale di partizioni primarie e di replica attive è maggiore di 30.000. È possibile che gli indici vengano ruotati troppo frequentemente. Prendi in considerazione l'utilizzo di ISM per rimuovere gli indici una volta raggiunta un'età specifica.
Allarmi `5xx` >= 10% di `OpenSearchRequests`	Uno o più nodi di dati potrebbero essere sovraccarichi o le richieste non vengono completate entro il periodo di timeout inattivo. Considera il passaggio a tipi di istanza più grandi o di aggiungere più nodi al cluster. Conferma che stai seguendo le best practice per l'architettura di partizioni e cluster.
`MasterReachableFromNode` il massimo è < 1 per 5 minuti, 1 volta consecutiva	Questo avviso indica che il nodo principale è stato arrestato o non è raggiungibile. Questi errori sono in genere il risultato di un problema di connettività di rete o di AWS dipendenza.
`ThreadpoolWriteQueue` medio è >= 100 per 1 minuto, 1 periodo di tempo consecutivo	Il cluster sta riscontrando un'elevata concorrenza di indicizzazione. Esamina e controlla le richieste di indicizzazione o aumenta le risorse del cluster.
`ThreadpoolSearchQueue` medio è >= 500 per 1 minuto, 1 periodo di tempo consecutivo	Il cluster sta riscontrando un'elevata concorrenza di ricerca. Considera il dimensionamento del cluster. È inoltre possibile aumentare le dimensioni della coda di ricerca, ma un aumento eccessivo può causare errori di memoria.
`ThreadpoolSearchQueue` massimo è >= 5000 per 1 minuto, 1 periodo di tempo consecutivo	

Allarme	Problema
L'aumento di `Threadpool SearchRejected` SUM è >=1 {espressione matematica DIFF ()} per 1 minuto, 1 volta consecutiva	Questi allarmi ti informano di problemi di dominio che potrebbero influire sulle prestazioni e sulla stabilità.
L'aumento della `Threadpool WriteRejected` SOMMA è >=1 {espressione matematica DIFF ()} per 1 minuto, 1 volta consecutiva	

> **Note**
>
> Se si desidera soltanto visualizzare i parametri, consultare the section called “Monitoraggio dei parametri del cluster”.

Altri allarmi che potresti prendere in considerazione

Valuta la possibilità di configurare i seguenti allarmi a seconda delle funzionalità di OpenSearch servizio che utilizzi regolarmente.

Allarme	Problema
`WarmFreeStorageSpace` è >= 10%	Hai raggiunto il 10% del tuo accumulo di calore totale gratuito. `WarmFreeStorageSpace` misura la somma dello spazio di archiviazione caldo libero in MiB. UltraWarm utilizza Amazon S3 anziché dischi collegati.

Allarme	Problema
`HotToWarmMigrationQueueSize` è >= 20 per 1 minuto, 3 periodi di tempo consecutivi	Un numero elevato di indici passa contemporaneamente dallo storage a caldo a quello di storage. UltraWarm Considera il dimensionamento del cluster.
`HotToWarmMigrationSuccessLatency` è >= 1 giorno, 1 periodo di tempo consecutivo	Configura questo allarme in modo da ricevere una notifica se la latenza x di `HotToWarmMigrationSuccessCount` è superiore a 24 ore, se stai cercando di utilizzare indici giornalieri.
`WarmJVMMemoryPressure` il massimo è >= 95% per 1 minuto, 3 volte consecutive	Sui cluster potrebbero verificarsi errori di esaurimento della memoria nel caso in cui l'utilizzo aumenti. Considerare il dimensionamento verticale . OpenSearch Il servizio utilizza metà della RAM di un'istanza per l'heap Java, fino a una dimensione dell'heap di 32 GiB. Puoi scalare le istanze verticalmente fino a 64 GiB di RAM e poi scalare orizzontalmente aggiungendo le istanze.
`WarmOldGenJVMMemoryPressure` il massimo è >= 80% per 1 minuto, 3 volte consecutive	
`WarmToColdMigrationQueueSize` è >= 20 per 1 minuto, 3 periodi di tempo consecutivi	Un numero elevato di indici sta passando contemporaneamente dalla conservazione a freddo. UltraWarm Considera il dimensionamento del cluster.

Allarme	Problema
`HotToWarmMigrationFailureCount` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	Le migrazioni potrebbero non riuscire durante gli snapshot, le rilocazioni di partizioni o le fusioni forzate. Gli errori durante gli snapshot o il trasferimento di partizioni sono in genere dovuti a errori dei nodi o a problemi di connettività S3. La mancanza di spazio su disco è solitamente la causa sottostante degli errori di unioni forzate.
`WarmToColdMigrationFailureCount` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	Le migrazioni in genere falliscono quando i tentativi di migrazione dei metadati dell'indice nell'archiviazione a freddo non riescono. È possibile che si verifichino degli errori anche durante la rimozione dello stato del cluster di indice a caldo.
`WarmToColdMigrationLatency` è >= 1 giorno, 1 periodo di tempo consecutivo	Configura questo allarme in modo da ricevere una notifica se la latenza x di `WarmToColdMigrationSuccessCount` è superiore a 24 ore, se stai cercando di utilizzare indici giornalieri.
`AlertingDegraded` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	L'indice di avviso è rosso oppure uno o più nodi non sono pianificati.
`ADPluginUnhealthy` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	Il plugin di rilevamento delle anomalie non funziona correttamente a causa di alti tassi di errore o perché uno degli indici utilizzati è rosso.

Allarme	Problema
`AsynchronousSearchFailureRate` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	Almeno una ricerca asincrona non è riuscita nell'ultimo minuto, il che significa che il nodo coordinatore non è riuscito. Il ciclo di vita di una richiesta di ricerca asincrona viene gestito esclusivamente sul nodo del coordinatore, quindi se il coordinatore si interrompe, la richiesta non riesce.
`AsynchronousSearchStoreHealth` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	L'integrità dell'archivio delle risposte di ricerca asincrona nell'indice persistente è rossa. È possibile che si stiano memorizzando risposte asincrone di grandi dimensioni, che possono destabilizzare un cluster. Cerca di limitare le risposte di ricerca asincrone a 10 MB o meno.
`SQLUnhealthy` è >= 1 per 1 minuto, 3 periodi di tempo consecutivi	Il plugin SQL restituisce 5 codici di risposta xx o passa una query DSL non valida a. OpenSearch Risolvi i problemi relativi alle richieste che i client stanno facendo al plug-in.
`LTRStatus.red` è >= 1 per 1 minuto, 1 periodo di tempo consecutivo	Almeno uno degli indici necessari per eseguire il plug-in Learning to Rank (Imparare a classificare) ha partizioni primarie mancanti e non è funzionante.

Dimensionamento dei domini Amazon OpenSearch Service

Non esiste un metodo perfetto per dimensionare i domini di Amazon OpenSearch Service. Tuttavia, partendo da una comprensione delle tue esigenze di storage, del servizio e di OpenSearch se stesso, puoi fare una stima iniziale approfondita delle tue esigenze hardware. Questa stima può essere utilizzata come punto di partenza per l'aspetto più importante del dimensionamento dei domini: testarli con carichi di lavoro rappresentativi e monitorare le prestazioni.

Argomenti

- Calcolo dei requisiti di archiviazione

- Scelta del numero di partizioni
- Scelta del tipo di istanza e test

Calcolo dei requisiti di archiviazione

La maggior parte dei OpenSearch carichi di lavoro rientra in una delle due grandi categorie:

- Indice di lunga durata: scrivi codice che elabora i dati in uno o più OpenSearch indici e quindi aggiorna tali indici periodicamente man mano che i dati di origine cambiano. Alcuni esempi comuni riguardano la ricerca su siti Web, documenti ed e-commerce.
- Indici in sequenza: i dati fluiscono in modo continuo in un set di indici temporanei, con un periodo di indicizzazione e una finestra di conservazione, ad esempio un set di indici giornalieri che viene conservato per due settimane. Alcuni esempi comuni sono le analisi di log, l'elaborazione delle serie temporali e le analisi clickstream.

Per i carichi di lavoro dell'indice di lunga durata, è possibile esaminare i dati di origine sul disco e determinare facilmente la quantità di spazio di archiviazione che consuma. Se i dati provengono da più origini, devi aggiungere tali origini.

Per gli indici in sequenza, puoi moltiplicare la quantità di dati generati durante un periodo di tempo rappresentativo dal periodo di conservazione. Ad esempio, se generi 200 MiB di dati di log all'ora, questi corrispondono a 4,7 GiB al giorno, 66 GiB di dati in qualsiasi momento, se disponi di un periodo di retention di due settimane.

Le dimensioni dei dati di origine, tuttavia, sono solo un aspetto delle esigenze di archiviazione. È·necessario anche considerare quanto segue:

- Numero di repliche: ogni replica è una copia completa dello shard principale, la dimensione dell'archivio dell'indice mostra la dimensione occupata dallo shard primario e da quello di replica. Per impostazione predefinita, ogni OpenSearch indice ha una replica. Ne consigliamo almeno una per evitare la perdita di dati. Le repliche, inoltre, migliorano le prestazioni di ricerca, perciò potresti volerne di più se hai un carico di lavoro gravoso in lettura. Utilizzare `PUT /my-index/_settings` per aggiornare l'impostazione `number_of_replicas` per l'indice.
- OpenSearch sovraccarico di indicizzazione: la dimensione su disco di un indice varia. La dimensione totale dei dati di origine e dell'indice spesso è pari al 110% dell'origine, dove l'indice rappresenta fino al 10% dei dati di origine. Dopo l'indicizzazione dei dati, è possibile utilizzare

l'API `_cat/indices?v` e il valore `pri.store.size` per calcolare il sovraccarico esatto. `_cat/allocation?v` fornisce anche un riepilogo utile.

- Spazio riservato per il sistema operativo: per impostazione predefinita, Linux riserva il 5% del file system per l'utente `root` per i processi critici, il ripristino del sistema e per evitare problemi di frammentazione del disco.
- OpenSearch Sovraccarico del OpenSearch servizio: il servizio riserva il 20% dello spazio di archiviazione di ogni istanza (fino a 20 GiB) per fusioni di segmenti, log e altre operazioni interne.

  Data la dimensione massima di 20 GiB, la quantità totale di spazio riservato può variare notevolmente in funzione del numero di istanze nel tuo dominio. Ad esempio, un dominio può avere tre istanze `m6g.xlarge.search`, ognuna con 500 GiB di spazio di archiviazione, per un totale di 1,46 TiB. In questo caso, il totale di spazio riservato è solo 60 GiB. Un altro dominio può avere 10 istanze `m3.medium.search`, ognuna con 100 GiB di spazio di archiviazione, per un totale di 0,98 TiB. In questo caso, il totale di spazio riservato è di 200 GiB, anche se il primo dominio è il 50% più grande.

  Nella formula seguente, applichiamo una stima "nel peggiore dei casi" per un sovraccarico. Questa stima include spazio libero aggiuntivo per ridurre al minimo l'impatto degli errori dei nodi e delle interruzioni delle zone di disponibilità.

Riepilogando, se si dispone di 66 GiB di dati in qualsiasi momento e si desidera una replica, il requisito di archiviazione minimo è più vicino a 66* 2* 1,1/0,95/0,8 = 191 GiB. Puoi generalizzare il calcolo come indicato di seguito:

Dati di origine* (1+ numero di repliche) * (1+ sovraccarico di indicizzazione)/(1 - spazio riservato Linux)/(1 - sovraccarico del servizio) = requisito minimo di archiviazione OpenSearch

In alternativa, puoi utilizzare questa versione semplificata:

Dati di origine * (1 + Numero di repliche) * 1,45 = Requisito di minimo di archiviazione

Lo spazio di archiviazione insufficiente è una delle cause più comuni di instabilità del cluster. Pertanto quando scegli tipi di istanze, numero di istanze e volumi di archiviazione dovresti controllare i numeri.

Esistono altre considerazioni di archiviazione:

- Se i requisiti di archiviazione minimi superano 1 PB, consultare the section called “Scala in petabyte”.

- Se disponi di indici in sequenza e desideri utilizzare un'architettura a caldo/ad accesso frequente, consulta the section called “UltraWarm archiviazione”.

Scelta del numero di partizioni

Dopo aver individuato i requisiti di archiviazione, è possibile esaminare la strategia di indicizzazione. Per impostazione predefinita, in OpenSearch Service, ogni indice è suddiviso in cinque shard primari e una replica (per un totale di 10 shard). Questo comportamento è diverso da quello open source OpenSearch, che utilizza per impostazione predefinita uno shard primario e uno di replica. Poiché non è possibile modificare facilmente il numero di partizioni primarie per un indice esistente, è necessario decidere il numero di partizioni prima di indicizzare il primo documento.

L'obiettivo generale della scelta di un numero di partizioni è distribuire un indice in modo uniforme su tutti i nodi di dati del cluster. Tuttavia, queste partizioni non devono essere troppo grandi o troppo numerose. Una buona regola è cercare di mantenere le dimensioni delle partizioni tra 10 e 30 GiB per i carichi di lavoro in cui la latenza di ricerca è un obiettivo chiave delle prestazioni e 30-50 GiB per i carichi di lavoro pesanti in termini di scrittura, come l'analisi dei log.

Gli shard di grandi dimensioni possono rendere difficile il ripristino in caso di guasto, ma poiché ogni shard utilizza una certa quantità di CPU e memoria, avere troppi shard di piccole dimensioni può causare problemi di prestazioni ed errori di memoria insufficiente. OpenSearch In altre parole, gli shard devono essere sufficientemente piccoli da consentire all'istanza di OpenSearch Service sottostante di gestirli, ma non così piccoli da sovraccaricare inutilmente l'hardware.

Ad esempio, se disponi di 66 GiB di dati. Non si prevede che il numero aumenti nel corso del tempo e si desidera mantenere le dimensioni delle partizioni attorno a 30 GiB ciascuna. Il numero di partizioni perciò deve essere circa 66* 1,1/30 = 3. Puoi generalizzare il calcolo come indicato di seguito:

(Dati di origine + Spazio per crescere) * (1 + Gestione indicizzazione)/Dimensione desiderata della partizione = Numero approssimativo di partizioni primarie

Questa equazione aiuta a compensare la crescita dei dati nel tempo. Se si prevede che quegli stessi 66 GiB di dati quadruplicheranno l'anno successivo, il numero approssimativo di partizioni sarà (66+198) * 1,1/30 = 10. Ricorda, tuttavia, che non disponi ancora di questi 198 GiB di dati aggiuntivi. Verificare che la preparazione per il futuro non crei partizioni inutilmente piccole che consumano enormi quantità di CPU e memoria nel presente. In questo caso, 66* 1,1/10 partizioni = 7,26 GiB per partizione. Queste partizioni consumeranno risorse aggiuntive e sono inferiori all'intervallo di dimensione consigliato. Potreste prendere in considerazione l'idea di optare per middle-of-the-

road un approccio a sei shard, che vi lascerà con shard da 12 GiB oggi e shard da 48 GiB in futuro. Quindi, si potrebbe iniziare con tre partizioni e reindicizzare i dati quando le partizioni superano i 50 GiB.

Un problema molto meno comune comporta la limitazione del numero di partizioni per nodo. Se si dimensionano le partizioni in modo appropriato, in genere si esaurisce lo spazio su disco molto prima di rispettare questo limite. Ad esempio, un'istanza `m6g.large.search` ha una dimensione massima del disco di 512 GiB. Se si rimane al di sotto dell'80% di utilizzo del disco e si dimensionano le partizioni a 20 GiB, può ospitare circa 20 partizioni. Elasticsearch 7. x e versioni successive, e tutte le versioni OpenSearch fino alla 2.15, hanno un limite di 1.000 shard per nodo. Per regolare il numero massimo di partizioni per nodo, configura l'impostazione `cluster.max_shards_per_node`. Per la versione OpenSearch 2.17 e successive, OpenSearch Service supporta 1000 shard per ogni 16 GB di memoria heap JVM fino a un massimo di 4000 shard per nodo. Per un esempio, consulta Impostazioni del cluster. Per ulteriori informazioni sul numero di frammenti, vedere. the section called “Quote di conteggio condivise”

Il dimensionamento appropriato delle partizioni mantiene l'utente quasi sempre al di sotto di questo limite, ma è anche possibile considerare il numero di partizioni per ogni GiB di heap Java. Su un dato nodo, non avere più di 25 partizioni per GiB di heap Java. Ad esempio, un'istanza `m5.large.search` ha una memoria heap di 4 GiB, quindi ogni nodo non deve avere più di 100 partizioni. A quel numero di partizioni, ogni partizione ha una dimensione di circa 5 GiB, il che è ben al di sotto della nostra raccomandazione.

Scelta del tipo di istanza e test

Dopo aver calcolato i requisiti di archiviazione e scelto il numero di partizioni di cui si ha bisogno, è possibile iniziare a prendere decisioni sull'hardware. I requisiti hardware variano in modo significativo in base al carico di lavoro, ma possiamo comunque offrire alcune raccomandazioni di base.

In generale, i limiti di archiviazione per ogni tipo di istanza si mappano sulla quantità di CPU e memoria di cui si potrebbe aver bisogno per i carichi di lavoro leggeri. Ad esempio, un'istanza `m6g.large.search` dispone di dimensioni di volume EBS massime di 512 GiB, 2 core vCPU e 8 GiB di memoria. Se il cluster dispone di molte partizioni, esegue aggregazioni fiscali, aggiorna frequentemente i documenti o elabora un numero elevato di query, tali risorse potrebbero non essere sufficienti per le tue esigenze. Se ritieni che il tuo cluster rientri in una di queste categorie, prova a partire da una configurazione più vicina a 2 core vCPU e 8 GiB di memoria per ogni 100 GiB di archiviazione.

> **Tip**
>
> Per un riepilogo delle risorse hardware allocate a ciascun tipo di istanza, consulta i prezzi di Amazon OpenSearch Service.

Tuttavia, anche tali risorse potrebbero essere insufficienti. Alcuni OpenSearch utenti segnalano di aver bisogno di molte più risorse per soddisfare i propri requisiti. Trovare l'hardware appropriato per il tuo carico di lavoro significa fare una stima iniziale efficace, effettuare test con carichi di lavoro rappresentativi, apportare eventuali modifiche ed effettuare nuovamente il test.

Fase 1: Esecuzione di una stima iniziale

Per iniziare, consigliamo un minimo di tre nodi per evitare potenziali OpenSearch problemi, come uno stato cerebrale diviso (quando un'interruzione della comunicazione porta a un cluster con due nodi principali). Se si dispone di tre nodi master dedicati, consigliamo un minimo di due nodi di dati per la replica.

Fase 2: Calcolo dei requisiti di archiviazione per nodo

Se hai un requisito di archiviazione di 184 GiB e il numero minimo raccomandato di tre nodi, per trovare la quantità di spazio di archiviazione richiesta da ogni nodo puoi utilizzare l'equazione 184/3 = 61 GiB. In questo esempio, puoi selezionare tre istanze `m6g.large.search`, di cui ognuna utilizza un volume di archiviazione EBS di 90 GiB in modo da disporre di una rete di sicurezza e di spazio per la crescita nel corso del tempo. Questa configurazione fornisce 6 core vCPU e 24 GiB di memoria, quindi è ideale per i carichi di lavoro più leggeri.

Per un esempio più sostanziale, considerare un requisito di archiviazione di 14 TiB e un carico di lavoro pesante. In questo caso, è possibile scegliere di avviare il test con 2* 144 = 288 core vCPU e 8* 144 = 1152 GiB di memoria. Questi numeri portano a circa 18 istanze `i3.4xlarge.search`. Se non è necessaria un'archiviazione veloce locale, puoi anche testare 18 istanze `r6g.4xlarge.search`, ciascuna con un volume di archiviazione EBS di 1 TiB.

Se il cluster include centinaia di terabyte di dati, consulta the section called “Scala in petabyte”.

Fase 3: Esecuzione di test rappresentativi

Dopo aver configurato il cluster, puoi aggiungere gli indici utilizzando il numero di shard calcolato in precedenza, eseguire alcuni test rappresentativi sui client utilizzando un set di dati realistico e monitorare le CloudWatch metriche per vedere come il cluster gestisce il carico di lavoro.

Fase 4: Riuscita o iterazione

Se le prestazioni soddisfano le tue esigenze, i test hanno esito positivo e le CloudWatch metriche sono normali, il cluster è pronto per l'uso. Ricorda di impostare CloudWatch allarmi per rilevare un utilizzo non corretto delle risorse.

Se le prestazioni non sono accettabili, i test hanno esito negativo o `CPUUtilization` o `JVMMemoryPressure` sono elevati, potrebbe essere necessario scegliere un altro tipo di istanza (o aggiungere istanze) e continuare il test. Man mano che aggiungi istanze, riequilibra OpenSearch automaticamente la distribuzione degli shard in tutto il cluster.

Poiché è più semplice misurare le capacità in eccesso di un cluster con più potenza rispetto al disavanzo di uno con meno potenza, consigliamo di iniziare con un cluster di dimensioni maggiori rispetto alle tue esigenze. Successivamente, suggeriamo di testare e ridimensionare fino a un cluster efficiente che disponga di risorse aggiuntive per garantire operazioni stabili durante i periodi di maggiore attività.

I cluster di produzione o i cluster con stati complessi traggono vantaggio dai nodi master dedicati, che migliorano le prestazioni e l'affidabilità del cluster.

Scalabilità in petabyte in Amazon Service OpenSearch

I domini Amazon OpenSearch Service offrono storage collegato fino a 10 PB. Puoi configurare un dominio con 1000 tipi di `OR1.16xlarge.search` istanze, ciascuna con 36 TB di spazio di archiviazione. A causa dell'enorme differenza in scala, le raccomandazioni per i domini di queste dimensioni differiscono dalle nostre raccomandazioni generali. In questa sezione sono riportate le considerazioni sulla creazione di domini, sui costi, sull'archiviazione e sulle dimensioni delle partizioni.

Sebbene questa sezione faccia spesso riferimento ai tipi di `i3.16xlarge.search` istanze, puoi utilizzare diversi altri tipi di istanza per raggiungere 10 PB di storage totale del dominio.

Creazione di domini

> I domini di queste dimensioni superano il limite predefinito di 80 istanze per dominio. Per richiedere un aumento del limite di servizio fino a 1000 istanze per dominio, apri una richiesta presso il AWS Support Center.

Prezzi

Prima di creare un dominio di queste dimensioni, consulta la pagina dei prezzi di Amazon OpenSearch Service per assicurarti che i costi associati corrispondano alle tue aspettative. Esamina the section called “UltraWarm archiviazione” per vedere se un'architettura a caldo si adatta al tuo caso d'uso.

Archiviazione

I tipi di `i3` istanza sono progettati per fornire uno storage rapido e locale non volatile su memoria express (NVMe). Poiché questo storage locale tende a offrire vantaggi in termini di prestazioni rispetto ad Amazon Elastic Block Store, i volumi EBS non sono un'opzione quando si selezionano questi tipi di istanze in OpenSearch Service. Se si preferisce una archiviazione EBS, utilizzare un altro tipo di istanza, ad esempio `r6.12xlarge.search`.

Dimensioni e conteggio di partizioni

Una OpenSearch linea guida comune è quella di non superare i 50 GB per shard. Considerato il numero di partizioni necessarie per gestire domini di grandi dimensioni e le risorse disponibili per le istanze `i3.16xlarge.search`, consigliamo una dimensione della partizione pari a 100 GB.

Ad esempio, se si dispone di 450 TB di dati di origine e si desidere una replica, il requisito di archiviazione minimo è più vicino a 450 TB* 2* 1,1/0,95 = 1,04 PB. Per una spiegazione del calcolo, consulta the section called “Calcolo dei requisiti di archiviazione”. Anche se 1,04 PB/15 TB = 70 istanze, è possibile selezionare 90 o più istanze `i3.16xlarge.search` per garantire una rete di sicurezza per l'archiviazione, gestire i guasti dei nodi e tenere conto di alcune variazioni nella quantità di dati nel tempo. Ogni istanza aggiunge altri 20 GiB al requisito di archiviazione minimo, ma per dischi di queste dimensioni, questi 20 GiB sono quasi trascurabili.

Controllare il numero di shard è complicato. OpenSearch gli utenti spesso ruotano gli indici su base giornaliera e conservano i dati per una o due settimane. In questo caso, può essere utile distinguere tra partizioni "attive" e "inattive". Le partizioni attive sono attivamente oggetto di operazioni di lettura o scrittura. Le partizioni inattive potrebbero servire le richieste occasionali di lettura, ma sono fondamentalmente inattive. In generale, è consigliabile mantenere il numero di partizioni attive al di sotto di alcune migliaia. Se il numero di partizioni attive si avvicina a 10.000, potrebbero emergere rischi considerevoli in termini di prestazioni e stabilità.

Per calcolare il numero di partizioni primarie, utilizzare questa formula: 450.000 GB* 1,1/100 GB per partizione = 4.950 partizioni. Raddoppiando il numero per tenere conto delle repliche, raggiungiamo 9.900 partizioni, il che diventa motivo di grande preoccupazione se tutte le partizioni

sono attive. Tuttavia, se si ruotano gli indici e solo 1/7° o 1/14° delle partizioni è attivo in un dato giorno (1.414 o 707 partizioni, rispettivamente), il cluster potrebbe funzionare perfettamente. Come sempre, la fase più importante del dimensionamento e della configurazione del dominio è eseguire il test rappresentativo del client utilizzando un set di dati realistico.

Nodi coordinatori dedicati in Amazon Service OpenSearch

I nodi coordinatori dedicati in Amazon OpenSearch Service sono nodi specializzati che scaricano le attività di coordinamento dai nodi di dati. Queste attività includono la gestione delle richieste di ricerca e l'hosting OpenSearch di dashboard. Separando queste funzioni, i nodi coordinatori dedicati riducono il carico sui nodi di dati, il che consente loro di concentrarsi sull'archiviazione dei dati, sull'indicizzazione e sulle operazioni di ricerca. Ciò migliora le prestazioni complessive del cluster e l'utilizzo delle risorse.

Inoltre, i nodi coordinatori dedicati aiutano a ridurre il numero di indirizzi IP privati richiesti per le configurazioni VPC, il che porta a una gestione della rete più efficiente. Questa configurazione può comportare un miglioramento fino al 15% del throughput di indicizzazione e un miglioramento delle prestazioni delle query del 20%, a seconda delle caratteristiche del carico di lavoro.

Quando utilizzare i nodi coordinatori dedicati

I nodi coordinatori dedicati sono particolarmente utili nei seguenti scenari.

- Cluster di grandi dimensioni: in ambienti con un volume elevato di dati o query complesse, l'assegnazione delle attività di coordinamento su nodi dedicati può migliorare le prestazioni del cluster.
- Interrogazioni frequenti: i carichi di lavoro che comportano query di ricerca o aggregazioni frequenti, in particolare quelli con istogrammi di data complessi o aggregazioni multiple, traggono vantaggio da un'elaborazione più rapida delle query.
- Uso intensivo delle dashboard: le dashboard possono richiedere molte risorse. OpenSearch Affidare questa responsabilità a nodi coordinatori dedicati riduce il carico sui nodi di dati.

Architettura e comportamento

In un OpenSearch cluster, i nodi coordinatori dedicati gestiscono due responsabilità chiave.

- Gestione delle richieste: questi nodi ricevono le richieste di ricerca in entrata e le inoltrano ai nodi di dati appropriati, che archiviano i dati pertinenti. Quindi consolidano i risultati di più nodi di dati in un unico set di risultati globale, che viene restituito al client.
- Hosting di dashboard: i nodi Coordinator gestiscono le OpenSearch dashboard, il che allevia i nodi di dati dall'onere aggiuntivo di ospitare OpenSearch dashboard e gestire il traffico correlato.

Nei domini VPC, ai nodi coordinatori dedicati vengono assegnate Interfacce di rete elastiche (ENIs) anziché nodi dati. Questa disposizione consente di ridurre il numero di indirizzi IP privati necessari VPCs, migliorando così l'efficienza della rete. In genere, i nodi coordinatori dedicati rappresentano circa il 10% del totale dei nodi di dati.

Requisiti e limitazioni

I nodi coordinatori dedicati presentano i requisiti e le limitazioni seguenti.

- I nodi coordinatori dedicati sono supportati in tutte le OpenSearch versioni e nelle versioni di Elasticsearch da 6.8 a 7.10.
- Per abilitare i nodi coordinatori dedicati, il tuo dominio deve avere nodi master dedicati abilitati. Per ulteriori informazioni, consulta the section called “Nodi master dedicati”.
- Il provisioning di nodi coordinatori dedicati può comportare costi aggiuntivi. Tuttavia, la maggiore efficienza delle risorse e le migliori prestazioni giustificano l'investimento, in particolare per cluster di grandi dimensioni o complessi.

Fornitura di nodi coordinatori dedicati

Esegui i passaggi seguenti per effettuare il provisioning di nodi coordinatori dedicati in un dominio esistente. Assicurati che nel tuo dominio siano abilitati i nodi master dedicati prima di effettuare il provisioning dei nodi coordinatori.

Console

Per fornire nodi coordinatori dedicati in Console di gestione AWS

1. Accedi alla console di Amazon OpenSearch Service da https://console.aws.amazon.com/aos/casa.
2. Scegli Domini, quindi seleziona il dominio che desideri modificare.
3. Nella sezione Configurazione del cluster, scegli Modifica.

4. Scegli Abilita nodi coordinatori dedicati.
5. Seleziona il tipo di istanza e il numero di nodi coordinatori da fornire.
6. Scegli Save changes (Salva modifiche). L'aggiornamento del dominio potrebbe richiedere alcuni minuti.

AWS CLI

Per effettuare il provisioning di nodi coordinatori dedicati utilizzando il AWS CLI, usa il update-domain-configcomando. L'esempio seguente fornisce tre nodi `r6g.large.search` coordinatori in un dominio.

```
aws opensearch update-domain-config \
  --domain-name my-opensearch-domain \
  --cluster-config
 InstanceCount=3,InstanceType=r6g.large.search,DedicatedCoordinatorCount=3,ZoneAwarenessEnabled
```

Questo comando abilita i nodi coordinatori dedicati, imposta il tipo e il numero di istanze per i nodi coordinatori e abilita il riconoscimento delle zone per una maggiore disponibilità.

Best practice

Prendi in considerazione le seguenti best practice quando utilizzi nodi coordinatori dedicati.

- Utilizza istanze generiche per la maggior parte dei casi d'uso. Offrono un approccio equilibrato tra costi e prestazioni. Le istanze ottimizzate per la memoria sono ideali per carichi di lavoro che richiedono notevoli risorse di memoria, come quelli che implicano aggregazioni complesse o ricerche su larga scala.
- Un buon punto di partenza consiste nel fornire tra il 5% e il 10% dei nodi di dati come nodi coordinatori dedicati. Ad esempio, se il tuo dominio ha 90 nodi di dati di un particolare tipo di istanza, valuta la possibilità di fornire da 5 a 9 nodi di coordinazione dello stesso tipo di istanza.

> **Note**
>
> La disponibilità del tipo di istanza varia in base all'area geografica. Quando selezioni i tipi di istanza per i nodi coordinatori, verifica che il tipo di istanza scelto sia disponibile nella regione di destinazione. Puoi verificare la disponibilità del tipo di istanza nella console OpenSearch di servizio durante la creazione o la modifica del dominio.

- Per ridurre al minimo il rischio di un singolo punto di errore, fornisci almeno due nodi di coordinamento dedicati. Ciò garantisce che il cluster rimanga operativo anche in caso di guasto di un nodo.
- Se utilizzi la ricerca interregionale, fornisci nodi coordinatori dedicati nei domini di destinazione. I domini di origine in genere non gestiscono le attività di coordinamento.
- Per ambienti ad alta intensità di indicizzazione, prendi in considerazione le istanze ottimizzate per la CPU che corrispondono alle dimensioni dell'istanza dei nodi di dati per prestazioni ottimali.
- Per carichi di lavoro che richiedono molta memoria, utilizza un tipo di istanza leggermente più grande per i nodi coordinatori dedicati per gestire l'aumento delle richieste di memoria.
- Tieni traccia della CloudWatch metrica `CoordinatorCPUUtilization` Amazon. Se supera costantemente l'80%, potrebbe indicare che hai bisogno di nodi coordinatori più grandi o aggiuntivi per gestire il carico.
- Dimensiona i nodi coordinatori dedicati in modo che corrispondano ai nodi di dati. Ad esempio, inizia con 4 nodi di coordinamento generici quando usi nodi di dati 4xlarge.
- Utilizza più istanze più piccole anziché un numero inferiore di istanze più grandi per i nodi coordinatori, a meno che le tue richieste o risposte individuali non richiedano una quantità di memoria (in) estremamente elevata. GBs Ad esempio, scegli 12 istanze 4xl anziché 6 istanze generiche da 8 istanze di grandi dimensioni.

Consigli sui nodi in base alla dimensione del cluster

Utilizza le seguenti linee guida come punto di partenza per il provisioning di nodi coordinatori dedicati in base alle dimensioni del cluster. Modifica il numero e il tipo di nodi in base alle caratteristiche del carico di lavoro e alle metriche delle prestazioni.

Dimensione del cluster	Nodi coordinatori consigliati	Tipo di istanza
Piccoli (fino a 50 nodi)	3-5 nodi	Uso generale
Medio (50-100 nodi)	5-9 nodi	Memoria ottimizzata
Grande (oltre 100 nodi)	10-15 nodi	Memoria ottimizzata

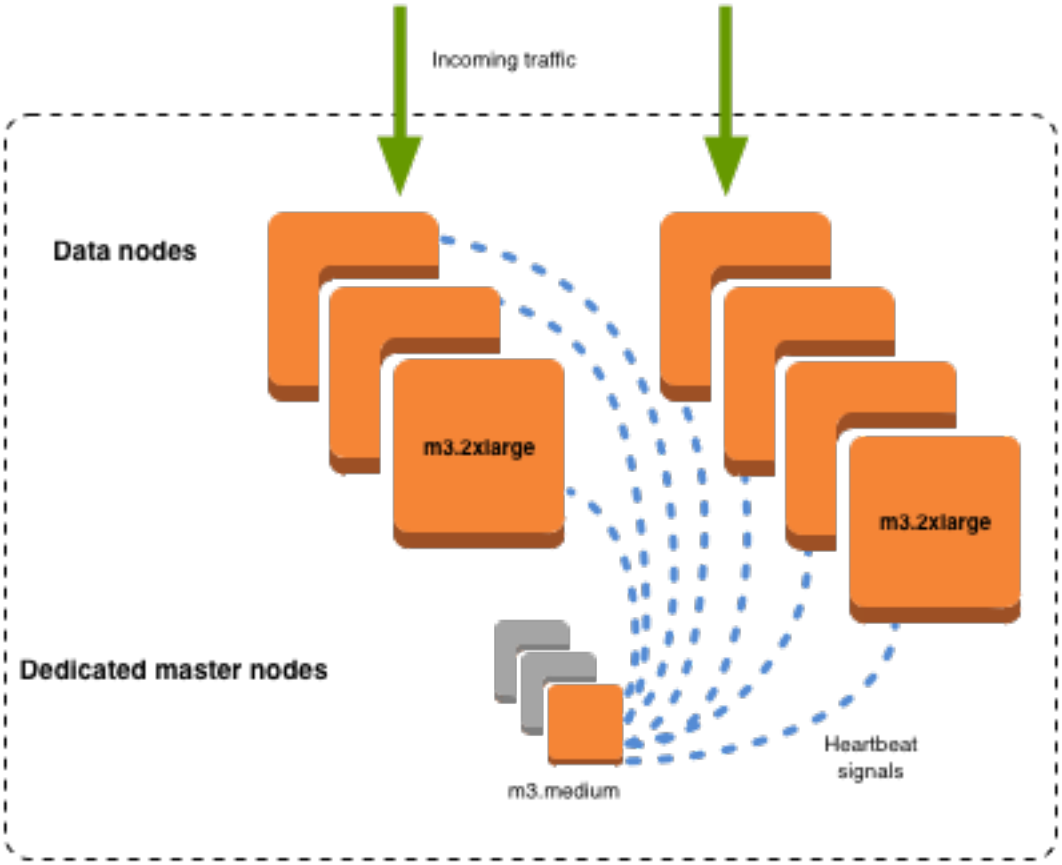
Nodi master dedicati in Amazon OpenSearch Service

Amazon OpenSearch Service utilizza nodi master dedicati per aumentare la stabilità del cluster. Un nodo master dedicato esegue task di gestione cluster ma non conserva dati né risponde a richieste di caricamento dei dati. L'offload dei task di gestione del cluster aumenta la stabilità del dominio. Come tutti gli altri tipi di nodo, per ogni nodo principale dedicato si paga una tariffa oraria.

I nodi master dedicati eseguono le seguenti attività di gestione del cluster:

- Monitora tutti i nodi del cluster.
- Monitora il numero di indici nel cluster.
- Monitora il numero di partizioni appartenenti a ciascun indice.
- Gestisci le informazioni di instradamento per i nodi del cluster.
- Aggiorna lo stato del cluster dopo aver apportato modifiche allo stato, ad esempio la creazione di un indice e l'aggiunta o la rimozione di nodi nel cluster.
- Replica le modifiche allo stato del cluster in tutti i nodi del cluster.
- Monitora l'integrità di tutti i nodi del cluster inviando segnali heartbeat, segnali periodici che monitorano la disponibilità dei nodi di dati nel cluster.

L'illustrazione seguente mostra un dominio OpenSearch di servizio con 10 istanze. Sette delle istanze sono nodi di dati e tre sono nodi master dedicati. È attivo solo uno dei nodi principali dedicati. I due nodi principali dedicati grigi attenderanno come backup nel caso in cui il nodo principale dedicato attivo riporti un errore. Tutte le richieste di caricamento dati vengono elaborate dai sette nodi di dati e su tutte le attività di gestione del cluster viene effettuato l'offload al nodo master dedicato attivo.



Scelta del numero di nodi principali dedicati

Si consiglia di utilizzare Multi-AZ con Standby, che aggiunge tre nodi master dedicati a ciascun dominio del servizio di produzione. OpenSearch Se utilizzi Multi-AZ senza Standby o Single-AZ, consigliamo comunque tre nodi master dedicati. Non scegliere mai un numero pari di nodi principali dedicati. Quando si sceglie il numero di nodi principali dedicati, tenere presente quanto segue:

- Un nodo master dedicato è esplicitamente vietato dal OpenSearch servizio perché non è disponibile alcun backup in caso di guasto. Se si prova a creare un dominio con un solo nodo principale dedicato, viene ricevuta un'eccezione di convalida.
- Se hai due nodi principali dedicati, allora il cluster non ha il quorum di nodi necessario per eleggere un nuovo nodo principale in caso di errore.

  Un quorum è il numero di nodi principali dedicati / 2 + 1 arrotondato per difetto al numero intero più vicino. In questo caso: 2 / 2 + 1 = 2. Poiché un nodo master dedicato non è andato a buon fine ed

esiste un solo backup, il cluster non dispone di un quorum e non è in grado di eleggere un nuovo master.

- Tre nodi master dedicati, il numero consigliato, offrono due nodi di backup in caso di guasto del nodo master e il quorum necessario (2) per eleggere un nuovo master.
- Quattro nodi principali dedicati non sono meglio di tre e possono causare problemi se utilizzi più zone di disponibilità.
  - Se un nodo master ha esito negativo, hai il quorum (3) per eleggere un nuovo master. Se due nodi hanno esito negativo, perdi il quorum, in modo analogo a quando disponi di tre nodi master dedicati.
  - In una configurazione a tre zone di disponibilità, due AZs hanno un nodo master dedicato e una zona AZ ne ha due. Se in quella zona si verifica un'interruzione, le due restanti AZs non hanno il quorum necessario (3) per eleggere un nuovo master.
- Avere cinque nodi master dedicati funziona così come tre e consente di perdere due nodi mantenendo un quorum. Ma poiché solo un nodo principale dedicato è attivo in un dato momento, questa configurazione implica pagare per quattro nodi inattivi. Molti clienti ritengono che questo livello di protezione failover sia eccessivo.

Se un cluster ha un numero pari di nodi idonei per il master e le versioni di Elasticsearch 7. OpenSearch x e versioni successive ignorano un nodo in modo che la configurazione di voto sia sempre un numero dispari. In questo caso, quattro nodi master dedicati sono essenzialmente equivalenti a tre (e due a uno).

> **Note**
>
> Se il cluster non ha il quorum necessario per eleggere un nuovo nodo master, le richieste di lettura e scrittura al cluster non andranno a buon fine. Questo comportamento è diverso da quello OpenSearch predefinito.

Scelta dei tipi di istanza per nodi principali dedicati

OpenSearch Quote di dominio e istanza del servizio

Sebbene i nodi master dedicati non elaborino le richieste di ricerca e di interrogazione, la loro dimensione è strettamente correlata alla dimensione dell'istanza e al numero di istanze, indici e

frammenti che possono gestire. Per i cluster di produzione, consigliamo almeno i seguenti tipi di istanza per nodi master dedicati.

Queste raccomandazioni sono basate su carichi di lavoro tipici e possono variare in base alle tue esigenze. I cluster con molte partizioni o mappature di campi possono trarre vantaggio da tipi di istanze di dimensioni maggiori. Per ulteriori informazioni, consulta la sezione CloudWatch Allarmi consigliati per Amazon OpenSearch Service per determinare se è necessario utilizzare un tipo di istanza più grande.

RAM	Max Node Support per Elasticsearch and OpenSearch Service da 1.x a 2.15	Max Shard Support per Elasticsearch and OpenSearch Service da 1.x a 2.15	Max Node Support for OpenSearch Service 2.17 e versioni successive	Max Shard Support for OpenSearch Service 2.17 e versioni successive
2 GB	Non applicabile	Non applicabile	10	1K
4 GB	Non applicabile	Non applicabile	10	5K
8 GB	10	10.000	30	15.000
16 GB	30	30K	60	30K
32 GB	75	40K	120	60 K
64 GB	125	75K	240	120 K
128 GB	200	75K	480	240 K
256 GB	Non applicabile	Non applicabile	1002	500 K

Riferimento generale per Amazon OpenSearch Service

Amazon OpenSearch Service supporta una varietà di istanze, operazioni, plug-in e altre risorse.

Argomenti

- Tipi di istanze supportati in Amazon OpenSearch Service
- Funzionalità per versione del motore in Amazon OpenSearch Service
- Plugin per versione del motore in Amazon Service OpenSearch
- Operazioni supportate in Amazon OpenSearch Service
- Quote OpenSearch del servizio Amazon
- Istanze riservate in Amazon Service OpenSearch
- Tecniche di ottimizzazione dei costi per Amazon OpenSearch Service
- Altre risorse supportate in Amazon OpenSearch Service

Tipi di istanze supportati in Amazon OpenSearch Service

Amazon OpenSearch Service supporta i seguenti tipi di istanze. Non tutte le regioni supportano tutti i tipi di istanze. Per i dettagli sulla disponibilità, consulta i prezzi OpenSearch di Amazon Service.

Per informazioni su quale tipo di istanza è appropriato per il tuo caso d'uso, consulta the section called “Dimensionamento dei domini”, the section called “Quote delle dimensioni dei volumi EBS” e the section called “Quote di rete”.

Tieni presente che le seguenti famiglie di istanze non supportano i controlli di crittografia VPC: T2, T3, R3, I3, C4, M4, R4, C5, M5, R5, C6g, M6g, R6g e R6gd. VPC Encryption Controls è una funzionalità di sicurezza e conformità che ti offre un controllo autorevole centralizzato per monitorare lo stato di crittografia dei flussi di traffico, ti aiuta a identificare le risorse che consentono la comunicazione in chiaro e infine ti offre meccanismi per applicare la crittografia in transito all'interno e all'interno di una regione. VPCs Per ulteriori dettagli, consulta la documentazione dei controlli di crittografia VPC.

Tipi di istanza della generazione attuale

Per prestazioni ottimali, ti consigliamo di utilizzare i seguenti tipi di istanza quando crei nuovi domini OpenSearch di servizio.

Tipo di istanza	Istanze	Restrizioni
i4i	i4i.large.search i4i.xlarge.search i4i.2xlarge.search i4i.4xlarge.search i4i.8xlarge.search i4i.12xlarge.search i4i.16xlarge.search i4i.24xlarge.search i4i.32xlarge.search	I tipi di istanza i4i richiedono Elasticsearch 5.1 o versione successiva o qualsiasi versione di, e non supportano lo storage di volumi EBS. OpenSearch
i7i	i7i.large.search i7i.xlarge.search i7i.2xlarge.search	I tipi di istanza i7i richiedono Elasticsearch 5.1 o versione successiva o qualsiasi versione di, e non supportano lo storage di volumi EBS. OpenSearch

Tipo di istanza	Istanze	Restrizioni
	i7i.4xlarge.search i7i.8xlarge.search i7i.12xlarge.search i7i.16xlarge.search	
i8g	i8g.large.search i8g.xlarge.search i8g.2xlarge.search i8g.4xlarge.search i8g.8xlarge.search i8g.12xlarge.search i8g.16xlarge.search	I tipi di istanza i8g richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e non supportano i volumi di archiviazione EBS. OpenSearch

Tipo di istanza	Istanze	Restrizioni
i8ge	i8ge.large.search i8ge.xlarge.search i8ge.2xlarge.search i8ge.3xlarge.search i8ge.6xlarge.search i8ge.12xlarge.search i8ge.18xlarge.search	I tipi di istanza i8ge richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e non supportano i volumi di archiviazione EBS. OpenSearch

Tipo di istanza	Istanze	Restrizioni
Graviton 3	c7g.large.search c7g.xlarge.search c7g.2xlarge.search c7g.4xlarge.search c7g.8xlarge.search c7g.12xlarge.search c7g.16xlarge.search m7g.medium.search m7g.large.search m7g.xlarge.search m7g.2xlarge.search m7g.4xlarge.search	Graviton3 supporta solo l'archiviazione gp3. Le istanze Graviton3 (c7g, m7g, r7g) supportano tutte le versioni e le versioni di Elasticsearch 7.9 e 7.10. OpenSearch

Tipo di istanza	Istanze	Restrizioni
	m7g.8xlarge.search m7g.12xlarge.search m7g.16xlarge.search r7g.medium.search r7g.large.search r7g.xlarge.search r7g.2xlarge.search r7g.4xlarge.ricerca r7g.8xlarge.search r7g.12xlarge.search r7g.16xlarge.ricerca r7gd.large.search r7gd.xlarge.search	

Tipo di istanza	Istanze	Restrizioni
	r7gd.2xlarge.search r7gd.4xlarge.search r7gd.8xlarge.search r7gd.12xlarge.search r7gd.16xlarge.search	

Tipo di istanza	Istanze	Restrizioni
OR1	or1.medium.search or1.large.search or1.xlarge.search or1.2xlarge.search or1.4xlarge.search or1.8xlarge.search or1.12xlarge.search or1.16xlarge.search	• I tipi di OR1 OpenSearch istanza richiedono la versione 2.11 o successiva. • OR1 le istanze sono compatibili solo con i nodi master di altri tipi di istanze Graviton (C6g, M6g, R6g).

Tipo di istanza	Istanze	Restrizioni
OR2	or2.medium.search or2.large.search or2.xlarge.search or2.2xlarge.search or2.4xlarge.search or2.8xlarge.search or2.12xlarge.search or2.16xlarge.search	• OR2 e i tipi di OM2 istanza richiedono la versione OpenSearch 2.11 o successiva e sono compatibili solo con i tipi di istanze dei nodi master basati su Graviton (C6g, M6g, R6g).

Tipo di istanza	Istanze	Restrizioni

OM2	`om2.large.search`
	`om2.xlarge.search`
	`om2.2xlarge.search`
	`om2.4xlarge.search`
	`om2.8xlarge.search`
	`om2.12xlarge.search`
	`om2.16xlarge.search`
		• OR2 e i tipi di OM2 istanza richiedono la versione OpenSearch 2.11 o successiva e sono compatibili solo con i tipi di istanze del nodo principale basati su Graviton (C6g, M6g, R6g).

Tipo di istanza	Istanze	Restrizioni

OI2	`oi2.large.search`
	`oi2.xlarge.search`
	`oi2.2xlarge.search`
	`oi2.4xlarge.search`
	`oi2.8xlarge.search`
	`oi2.12xlarge.search`
	`oi2.16xlarge.search`
	`oi2.24xlarge.search`
	`oi2.32xlarge.search`
		• I tipi di OI2 istanza richiedono la versione OpenSearch 2.11 o successiva.
		• Queste istanze sono dotate di storage e non richiedono uno NVMe storage EBS separato.
		• Queste istanze (fino a 8 volte più grandi) possono essere utilizzate anche come nodi caldi nella versione 3.3 o successiva. OpenSearch

Tipo di istanza	Istanze	Restrizioni

Im4gn	`im4gn.large.search`
	
	`im4gn.xlarge.search`
	
	`im4gn.2xlarge.search`
	
	`im4gn.4xlarge.search`
	
	`im4gn.8xlarge.search`
	
	`im4gn.16xlarge.search`
		• I tipi di istanza Im4gn richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e non supportano i volumi di archiviazione EBS. OpenSearch
		• Le istanze Im4gn sono compatibili solo con altri tipi di istanze Graviton (C6g, M6g, R6g, R6gd). Non è possibile combinare istanze Graviton e non Graviton nello stesso cluster.

T3	`t3.small.search`
	
	`t3.medium.search`
		• I tipi di istanza T3 richiedono Elasticsearch 5.6 o versione successiva o qualsiasi versione di. OpenSearch
		• Puoi utilizzare i tipi di istanza T3 solo se il tuo dominio viene fornito senza standby. Per ulteriori informazioni, consulta the section called “Multi-AZ senza Standby”.
		• Puoi utilizzare i tipi di istanze T3 solo se il numero di istanze per il tuo dominio è pari o inferiore a 10.
		• I tipi di istanze T3 non supportano UltraWarm storage, cold storage o Auto-Tune.

Tipo di istanza	Istanze	Restrizioni

c7i	`c7i.large.search`
	`c7i.xlarge.search`
	`c7i.2xlarge.search`
	`c7i.4xlarge.search`
	`c7i.8xlarge.search`
	`c7i.12xlarge.search`
	`c7i.16xlarge.search`
		• L'istanza c7i richiede Elasticsearch 5.1 o versione successiva o qualsiasi versione di e supporta solo volumi di archiviazione OpenSearch. GP3

Tipo di istanza	Istanze	Restrizioni

m7i	`m7i.large.search`
	`m7i.xlarge.search`
	`m7i.2xlarge.search`
	`m7i.4xlarge.search`
	`m7i.8xlarge.search`
	`m7i.12xlarge.search`
	`m7i.16xlarge.search`
		• L'istanza m7i richiede Elasticsearch 5.1 o versione successiva o qualsiasi versione di e supporta solo volumi di archiviazione OpenSearch. GP3

Tipo di istanza	Istanze	Restrizioni

r7i	`r7i.large.search`
	`r7i.xlarge.search`
	`r7i.2xlarge.search`
	`r7i.4xlarge.search`
	`r7i.8xlarge.search`
	`r7i.12xlarge.search`
	`r7i.16xlarge.search`
		• L'istanza r7i richiede Elasticsearch 5.1 o versione successiva o qualsiasi versione di e supporta solo volumi di archiviazione OpenSearch. GP3

Tipo di istanza	Istanze	Restrizioni
c8g	c8g.large.search c8g.xlarge.search c8g.2xlarge.search c8g.4xlarge.search c8g.8xlarge.search c8g.12xlarge.search c8g.16xlarge.search	• I tipi di istanza c8g richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e supportano solo volumi di archiviazione. OpenSearch GP3

Tipo di istanza	Istanze	Restrizioni

m8g	`m8g.medium.search`
	`m8g.large.search`
	`m8g.xlarge.search`
	`m8g.2xlarge.search`
	`m8g.4xlarge.search`
	`m8g.8xlarge.search`
	`m8g.12xlarge.search`
	`m8g.16xlarge.search`
		• I tipi di istanza m8g richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e supportano solo volumi di archiviazione. OpenSearch GP3

Tipo di istanza	Istanze	Restrizioni

r8g	`r8g.medium.search`
	`r8g.large.search`
	`r8g.xlarge.search`
	`r8g.2xlarge.search`
	`r8g.4xlarge.search`
	`r8g.8xlarge.search`
	`r8g.12xlarge.search`
	`r8g.16xlarge.search`
		• I tipi di istanza r8g richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e supportano solo volumi di archiviazione. OpenSearch GP3

Tipo di istanza	Istanze	Restrizioni

r8gd	`r8gd.medium.search`
	`r8gd.large.search`
	`r8gd.2xlarge.search`
	`r8gd.4xlarge.search`
	`r8gd.8xlarge.search`
	`r8gd.12xlarge.search`
	`r8gd.16xlarge.search`
		• I tipi di istanza r8gd richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e non supportano i volumi di archiviazione EBS. OpenSearch

Tipi di istanza di generazioni precedenti

OpenSearch Il servizio offre tipi di istanze di generazione precedente per gli utenti che hanno ottimizzato le proprie applicazioni in base a tali applicazioni e non hanno ancora effettuato l'aggiornamento. Si consiglia di utilizzare i tipi di istanza della generazione corrente per ottenere le migliori prestazioni, ma continuiamo a supportare i seguenti tipi di istanza della generazione precedente.

Tipo di istanza	Istanze	Restrizioni

C4	`c4.large.search`
	`c4.xlarge.search`
	`c4.2xlarge.search`
	`c4.4xlarge.search`
	`c4.8xlarge.search`	

I2	`i2.xlarge.search` `i2.2xlarge.search`	

M3	`m3.medium.search`
	`m3.large.search`
	`m3.xlarge.search`
	`m3.2xlarge.search`
		• I tipi di istanze M3 non supportano la crittografia dei dati a riposo, il controllo granulare degli accessi o la ricerca tra cluster.
		• I tipi di istanze M3 hanno restrizioni aggiuntive in base alla OpenSearch versione. Per ulteriori informazioni, consulta the section called “Tipo di istanza M3 non valido”.

M4	`m4.large.search`	

Tipo di istanza	Istanze	Restrizioni

	`m4.xlarge.search`
	`m4.2xlarge.search`
	`m4.4xlarge.search`
	`m4.10xlarge.search`	

R3	`r3.large.search`
	`r3.xlarge.search`
	`r3.2xlarge.search`
	`r3.4xlarge.search`
	`r3.8xlarge.search`
		I tipi di istanza R3 non supportano la crittografia dei dati a riposo o il controllo granulare degli accessi.

Tipo di istanza	Istanze	Restrizioni

R4	`r4.large.search`
	`r4.xlarge.search`
	`r4.2xlarge.search`
	`r4.4xlarge.search`
	`r4.8xlarge.search`
	`r4.16xlarge.search`	

i4g	i4g.large.search i4g.xlarge.search i4g.2xlarge.search i4g.4xlarge.ricerca i4g.8xlarge.search i4g.16xlarge.search	I tipi di istanze i4g richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e non supportano i volumi di archiviazione EBS. OpenSearch

Tipo di istanza	Istanze	Restrizioni

C5	`c5.large.search`
	`c5.xlarge.search`
	`c5.2xlarge.search`
	`c5.4xlarge.search`
	`c5.9xlarge.search`
	`c5.18xlarge.search`
		I tipi di istanza C5 richiedono Elasticsearch 5.1 o versione successiva o qualsiasi versione di. OpenSearch

Tipo di istanza	Istanze	Restrizioni

C6g	`c6g.large.search`
	`c6g.xlarge.search`
	`c6g.2xlarge.search`
	`c6g.4xlarge.search`
	`c6g.8xlarge.search`
	`c6g.12xlarge.search`
		• I tipi di istanza C6g richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di. OpenSearch
		• Le istanze C6g sono compatibili solo con altri tipi di istanze Graviton (Im4gn, M6g, R6g, R6gd). Non è possibile combinare istanze Graviton e non Graviton nello stesso cluster.

Tipo di istanza	Istanze	Restrizioni

I3	`i3.large.search`
	`i3.xlarge.search`
	`i3.2xlarge.search`
	`i3.4xlarge.search`
	`i3.8xlarge.search`
	`i3.16xlarge.search`	

M5	`m5.large.search`
	`m5.xlarge.search`
	`m5.2xlarge.search`
	`m5.4xlarge.search`
	`m5.12xlarge.search`
		I tipi di istanza M5 richiedono Elasticsearch 5.1 o versione successiva o qualsiasi versione di. OpenSearch

Tipo di istanza	Istanze	Restrizioni

M6g	`m6g.large.search`
	`m6g.xlarge.search`
	`m6g.2xlarge.search`
	`m6g.4xlarge.search`
	`m6g.8xlarge.search`
	`m6g.12xlarge.search`
		• I tipi di istanza M6g richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di. OpenSearch
		• Le istanze M6g sono compatibili solo con altri tipi di istanze Graviton (Im4gn, C6g, R6g, R6gd). Non è possibile combinare istanze Graviton e non Graviton nello stesso cluster.

Tipo di istanza	Istanze	Restrizioni

R5	`r5.large.search`
	`r5.xlarge.search`
	`r5.2xlarge.search`
	`r5.4xlarge.search`
	`r5.12xlarge.search`
		I tipi di istanza R5 richiedono Elasticsearch 5.1 o versione successiva o qualsiasi versione di. OpenSearch

Tipo di istanza	Istanze	Restrizioni

R6g	`r6g.large.search`
	`r6g.xlarge.search`
	`r6g.2xlarge.search`
	`r6g.4xlarge.search`
	`r6g.8xlarge.search`
	`r6g.12xlarge.search`
		• I tipi di istanza R6g richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di. OpenSearch
		• Le istanze R6g sono compatibili solo con altri tipi di istanze Graviton (Im4gn, C6g, M6g, R6gd). Non è possibile combinare istanze Graviton e non Graviton nello stesso cluster.

Tipo di istanza	Istanze	Restrizioni

R6gd	`r6gd.large.search`
	`r6gd.xlarge.search`
	`r6gd.2xlarge.search`
	`r6gd.4xlarge.search`
	`r6gd.8xlarge.search`
	`r6gd.12xlarge.search`
	`r6gd.16xlarge.search`
		• I tipi di istanza R6gd richiedono Elasticsearch 7.9 o versione successiva o qualsiasi versione di e non supportano i volumi di archiviazione EBS. OpenSearch
		• Le istanze R6gd sono compatibili solo con altri tipi di istanze Graviton (Im4gn, C6g, M6g, R6g). Non è possibile combinare istanze Graviton e non Graviton nello stesso cluster.

T2	`t2.micro.search`
	`t2.small.search`
	`t2.medium.search`
		• Puoi utilizzare i tipi di istanza T2 solo se il conteggio delle istanze per il tuo dominio è pari al massimo a 10.
		• Il tipo di istanza `t2.micro.search` supporta solo Elasticsearch 1.5 e 2.3.
		• I tipi di istanze T2 non supportano la crittografia dei dati inattivi, il controllo granulare degli accessi, l'archiviazione, la conservazione a freddo, la ricerca tra cluster o Auto-Tune. UltraWarm

> **Tip**
>
> Consigliamo di utilizzare tipi di istanze differenti per nodi principali dedicati e nodi di dati.

Funzionalità per versione del motore in Amazon OpenSearch Service

Molte funzionalità OpenSearch del servizio richiedono una OpenSearch versione minima o una versione precedente di Elasticsearch OSS. Se si soddisfa la versione minima di una funzionalità ma la funzionalità non è disponibile nel dominio, aggiornare il software di servizio del dominio.

Caratteristica	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
Architettura a più livelli	3.3	Non incluso
Fonte derivata	3.1	Non incluso
Indice Star-tree	2.17	Non incluso
Supporto di Amazon Q	2,17	Non incluso
Plugin personalizzati	2.15	Non incluso
Ricerca simultanea di segmenti	2.13	Non incluso
Generazione di interroga	2.13	Non incluso

Caratteristica	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
zioni in linguaggio naturale		
Ricerca semantica multimodale	2.11	Non incluso
Fonti di dati con interrogazione diretta	2.11	Non incluso
Connettori per l'apprendimento automatico	2.9	Non incluso
Ricerca semantica	2.9	Non incluso
Ricerca point-in-time	2.5	Non incluso
Notifications	2.3	Non incluso
ML Commons	1.3	Non incluso
Replica tra cluster	1.1	7.10
Trasformazioni degli indici	1.0	Non incluso

Caratteristica	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
Nodo coordinatore dedicato	1.0	6.8
Supporto per VPC	1.0	1.0
Richiedere HTTPS per tutto il traffico verso il dominio		
Supporto Multi-AZ		
Nodi master dedicati		
Pacchetti personalizzati		
Endpoint personalizzati		
Pubblicazione di log lenti		
Pubblicazione dei log di errori	1.0	5.1

Caratteristica	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
Crittografia dei dati a riposo		
Autenticazione Cognito per dashboard OpenSearch		
Aggiornamenti sul posto		
Snapshot automatiche orarie	1.0	5.3
Node-to-node crittografia	1.0	6.0
Supporto client REST Java di livello elevato		
Compressione delle richieste HTTP e delle risposte		

Caratteristica	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
Avviso	1.0	6.2
SQL	1.0	6,5
Funzionalità di ricerca tra cluster	1.0	6.7
Controllo granulare degli accessi		
autenticazione SAML per dashboard OpenSearch		
Regolazione automatica		
Reindicizzazione remota		
UltraWarm	1.0	6.8
Index State Management		

Caratteristica	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
k-NN per distanza euclidea	1.0	7.1
Rilevamento di anomalie	1.0	7.4
k-NN per similitudine del coseno	1.0	7.7
Learning to Rank		
Piped Processing Language (PPL)	1.0	7.9
OpenSearch Rapporti sulle dashboard		
OpenSearch Pannelli di controllo Trace Analytics		
Istanze Graviton basate su ARM		

Caratteristica	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
Archiviazione a freddo		
Distanza di marcia, distanza di norma L1 e scripting painless per k-NN	1.0	7.10
Ricerca asincrona		
Supporto curatore	Non incluso	5.1

Per informazioni sui plugin, che abilitano alcune di queste funzionalità e funzionalità aggiuntive, vedere the section called “Plug-in per versione motore”. Per informazioni sull'API OpenSearch per ogni versione, consulta the section called “Operazioni supportate”.

Plugin per versione del motore in Amazon Service OpenSearch

I domini Amazon OpenSearch Service sono preconfezionati con i plug-in della community. OpenSearch Il servizio distribuisce e gestisce automaticamente i plug-in per te, ma distribuisce plug-in diversi a seconda della versione OpenSearch o del sistema operativo Elasticsearch precedente che scegli per il tuo dominio.

La tabella seguente elenca i plugin per OpenSearch versione, nonché le versioni compatibili del precedente sistema operativo Elasticsearch. Include solo i plugin con cui è possibile interagire, non è esaustivo. OpenSearch Il servizio utilizza plug-in aggiuntivi per abilitare le funzionalità di base del servizio, come il plug-in S3 Repository per le istantanee e il plug-in OpenSearchPerformance

Analyzer per l'ottimizzazione e il monitoraggio. Per un elenco completo di tutti i plug-in in esecuzione sul dominio, effettuare la seguente richiesta:

```
GET _cat/plugins?v
```

Plug-in	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
HanLP	2.11	Non supportata
Analisi ebraica	2.11	Non supportata
Amazon Personalizza il posizionamento nelle ricerche	2.9	Non supportata
Ricerca neurale	2.9	Non supportata
Analisi della sicurezza	2.5	Non supportata
OpenSearch notifiche	2.3	Non supportata
ML Commons	1.3	Non supportata
Sudachi Analysis (consigliato per il giapponese)	1.3	Non supportata
STConvert	1.3	Non supportata

Plug-in	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
Analisi Pinyin	1.3	Non supportata
Analisi Nori	1.3	Non supportata
OpenSearch osservabilità	1.2	Non supportata
OpenSearch replica tra cluster	1.1	7.10
OpenSearch ricerca asincrona	1.0	7.10
IK (Chinese) Analysis	1.0	7.7
Vietnamese Analysis		
Thai analysis		
Learning to Rank		
OpenSearch rilevamento di anomalie	1.0	7.4
OpenSearch k-NN	1.0	7.1

Plug-in	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
OpenSearch Index State Management	1.0	6.8
OpenSearch sicurezza	1.0	6.7
OpenSearch SQL	1.0	6,5
OpenSearch avvisare	1.0	6.2
Ukrainian Analysis	1.0	5.3
Mapper Size	1.0	5.3
Mapper Murmur3	1.0	5.1
Ingest User Agent Processor	1.0	5.1
Ingest Attachment Processor	1.0	5.1
Stempel Polish Analysis	1.0	5.1

Plug-in	Versione minima richiesta OpenSearch	Versione minima richiesta di Elasticsearch
Smart Chinese Analysis	1.0	5.1
Analisi coreano Seunjeon	1.0	5.1
Phonetic Analysis	1.0	2.3
Japanese (kuromoji) Analysis	1.0	Incluso in tutti i domini
ICU Analysis	1.0	Incluso in tutti i domini

Plugin opzionali

Oltre ai plugin preinstallati di default, Amazon OpenSearch Service supporta diversi plugin opzionali per l'analisi del linguaggio. Puoi utilizzare Console di gestione AWS e AWS CLI per associare un plug-in a un dominio, dissociare un plug-in da un dominio ed elencare tutti i plug-in. Un pacchetto di plugin opzionale è compatibile con una OpenSearch versione specifica e può essere associato solo a domini con quella versione.

Nota che per il plugin Sudachi, quando riassoci un file di dizionario, questo non si riflette immediatamente sul dominio. Il dizionario si aggiorna quando viene eseguita la blue/green distribuzione successiva sul dominio come parte di una modifica della configurazione o di un altro aggiornamento. In alternativa, è possibile creare un nuovo pacchetto con i dati aggiornati, creare un nuovo indice utilizzando questo nuovo pacchetto, reindicizzare l'indice esistente nel nuovo indice e quindi eliminare il vecchio indice. Se preferisci utilizzare l'approccio di reindicizzazione, utilizza un alias di indice in modo da evitare interruzioni del traffico.

I plugin opzionali utilizzano il tipo di pacchetto. `ZIP-PLUGIN` Per ulteriori informazioni sui plugin opzionali, consulta. the section called “Pacchetti”

Operazioni supportate in Amazon OpenSearch Service

OpenSearch Il servizio supporta molte versioni OpenSearch e versioni precedenti di Elasticsearch OSS. Le sezioni seguenti mostrano le operazioni supportate dal OpenSearch servizio per ogni versione.

Argomenti

- Differenze significative tra le API

Differenze significative tra le API

Nuovo elenco APIs

Per supportare cluster di grandi dimensioni con un gran numero di indici e frammenti, abbiamo introdotto una nuova lista APIs con supporto per l'impaginazione, come _. list/indices and _list/shards L'API List recupera le statistiche su indici e frammenti in un formato impaginato. Ciò semplifica l'attività di elaborazione delle risposte che includono molti indici.

- `_list/indices`: _list/indices
- `_list/shards`: _lista/frammenti

Modifiche a quelle esistenti APIs

Per supportare cluster di grandi dimensioni, nell'`_cluster/stats`API abbiamo aggiunto il supporto per aggiungere filtri metrici aggiuntivi per supportare il recupero solo delle risposte alle statistiche pertinenti, ad esempio e. `_cluster/stats/<metric>/nodes/<node-filters>` `_cluster/stats/<metric>/<index_metric>/nodes/<node-filters>` Per i dettagli, consulta _cluster/stats.

Abbiamo aggiunto il supporto nell'`_cat/shards`API per l'annullamento delle attività specificando un parametro di richiesta. `cancel_after_time_interval` Per i dettagli, vedi _cat/shards.

Limitazione della dimensione della risposta per l'API _cat

Per supportare cluster di grandi dimensioni con un numero totale di istanze superiore a 200 tra data e warm node, abbiamo un limite di 10.000 al numero di indici restituiti da. `_cat/segments API` Se il numero di indici nella risposta supera questo limite, l'API restituisce un errore 429. Per evitare ciò, puoi specificare un filtro basato su un modello di indice nella tua query, ad esempio. `_cat/segments/<index-pattern>`

Impostazioni e statistiche

OpenSearch Il servizio accetta solo richieste `_cluster/settings` PUT all'API che utilizzano il modulo di impostazioni «flat». Rifiuta le richieste che utilizzano il modulo delle impostazioni espanse.

```
// Accepted
PUT _cluster/settings
{
  "persistent" : {
    "action.auto_create_index" : false
  }
}

// Rejected
PUT _cluster/settings
{
  "persistent": {
    "action": {
      "auto_create_index": false
    }
  }
}
```

Il client Java REST di alto livello utilizza il modulo espanso, quindi se è necessario inviare richieste di impostazioni, utilizzare il client di basso livello.

Prima di Elasticsearch 5.3, l'`_cluster/settings`API sui domini di OpenSearch servizio supportava solo il `PUT` metodo HTTP, non il metodo. `GET` OpenSearch e le versioni successive di Elasticsearch supportano il `GET` metodo, come mostrato nell'esempio seguente:

```
GET https://domain-name.region.es.amazonaws.com/_cluster/settings?pretty
```

Ecco un esempio di ritorno:

```
{
```

```
  "persistent": {
    "cluster": {
      "routing": {
        "allocation": {
          "cluster_concurrent_rebalance": "2",
          "node_concurrent_recoveries": "2",
          "disk": {
            "watermark": {
              "low": "1.35gb",
              "flood_stage": "0.45gb",
              "high": "0.9gb"
            }
          },
          "node_initial_primarirecoveries": "4"
        }
      }
    },
    "indices": {
      "recovery": {
        "max_bytper_sec": "40mb"
      }
    }
  }
}
```

Se confronti le risposte di un OpenSearch cluster open source e di OpenSearch Service per determinate impostazioni e statistiche APIs, potresti notare la mancanza di campi. OpenSearch Il servizio oscura alcune informazioni che rivelano le funzionalità interne del servizio, come il percorso dei dati del file system da cui proviene `_nodes/stats` o il nome e la versione del sistema operativo da. `_nodes`

Riduzione

L'API `_shrink` può impedire aggiornamenti, modifiche di configurazione ed eliminazioni di domini. Ne sconsigliamo l'utilizzo su domini che eseguono Elasticsearch versioni 5.3 o 5.1. Queste versioni contengono un baco che impedisce il ripristino snapshot di indici ridotti.

Se utilizzi l'`_shrink`API su altre o OpenSearch versioni di Elasticsearch, fai la seguente richiesta prima di iniziare l'operazione di compattazione:

```
PUT https://domain-name.region.es.amazonaws.com/source-index/_settings
{
```

```
  "settings": {
    "index.routing.allocation.require._name": "name-of-the-node-to-shrink-to",
    "index.blocks.read_only": true
  }
}
```

Quindi effettua le seguenti richieste al termine dell'operazione di riduzione:

```
PUT https://domain-name.region.es.amazonaws.com/source-index/_settings
{
  "settings": {
    "index.routing.allocation.require._name": null,
    "index.blocks.read_only": false
  }
}

PUT https://domain-name.region.es.amazonaws.com/shrunken-index/_settings
{
  "settings": {
    "index.routing.allocation.require._name": null,
    "index.blocks.read_only": false
  }
}
```

Nuova lista APIs

Per supportare cluster di grandi dimensioni con un numero enorme di indici e frammenti, abbiamo introdotto un nuovo elenco APIs con supporto per l'impaginazione, ad esempio e. `_list/indices` `_list/shards` L'API List recupera le statistiche su indici e frammenti in un formato impaginato. Ciò semplifica l'attività di elaborazione delle risposte che includono molti indici. Per ulteriori informazioni su`_list/indices`, consulta List indices. Per ulteriori informazioni su`_list/shards`, consulta List shards.

Modifiche a quelle esistenti APIs

Per supportare cluster di grandi dimensioni, abbiamo aggiunto il supporto in `_cluster/stats/<metric>/nodes/<node-filters>` and`_cluster/stats/<metric>/<index_metric>/nodes/<node-filters>`. Per ulteriori informazioni su`_cluster/stats`, consulta Cluster stats.

Limitazione della dimensione della risposta per _cat APIs

Per supportare cluster di grandi dimensioni con un numero totale di istanze superiore a 200 tra data e warm node, abbiamo un limite di 10.000 al numero di indici restituiti dall'API _cat/segments. Se il numero di indici nella risposta supera questo limite, l'API restituisce un errore. `429` Per evitare ciò, puoi specificare un filtro basato su un modello di indice nella tua query (ad esempio,). `_cat/segments/<index-pattern>`

Inoltre, il supporto per l'annullamento delle attività è ora disponibile per l'`_cat/shards`API per l'annullamento delle attività specificando il parametro di `cancel_after_time_interval` richiesta. Per ulteriori informazioni su questo argomento, consulta CAT shards.

Scelta dei tipi di istanza per i nodi master dedicati

La tabella seguente fornisce consigli per la scelta dei tipi di istanza appropriati per i nodi master dedicati:

RAM	Numero massimo di nodi supportati	Numero massimo di shard supportati
2 GB	10	1.000
4 GB	10	5.000
8 GB	30	15.000
16 GB	60	30.000
32 GB	120	60.000
64 GB	240	120.000
128 GB	480	240.000
256 GB	1002	500.000

OpenSearch versione 2.19

Per informazioni sulle operazioni della OpenSearch versione 2.19, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico. Per maggiori dettagli sulle modifiche di questa versione, consultate le note di rilascio della versione 2.19.

OpenSearch versione 2.17

Per la OpenSearch versione 2.17, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

> **Note**
>
> A partire dalla OpenSearch 2.17, l'`cluster.max_shards_per_node`impostazione non può essere modificata. A partire dalla versione OpenSearch 2.17, OpenSearch Service supporta 1000 shard per ogni 16 GB di memoria heap JVM fino a un massimo di 4000 shard per nodo.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_list`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_plugins/_asynchronous_search`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search/pipeline`
- `/_search/point_in_time`
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`

- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
  - `cluster.search.request.slowlog.level`
  - `cluster.search.request.slowlog.threshold.warn`
  - `cluster.search.request.slowlog.threshold.info`
  - `cluster.search.request.slowlog.threshold.debug`
  - `cluster.search.request.slowlog.threshold.trace`
  - `search.phase_took_enabled`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_dashboards`
- `/_plugins/_alerting`
- `/_plugins/_anomaly_detection`
- `/_plugins/_ism`
- `/_plugins/_ml`
- `/_plugins/_notifications`
- `/_plugins/_ppl`
- `/_plugins/_security`
- `/_plugins/_security_analytics`
- `/_plugins/_sm`
- `/_plugins/_sql`
- `/_percolate`
- `/_rank_eval`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare i valori al servizio. `scroll_id` OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Fate riferimento al `PUT` metodo. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e altre.
5. Per informazioni, consulta the section called “Riduzione”.

> **Note**
>
> Attualmente, la modifica della funzionalità di `cluster.max_shards_per_node` impostazione non è abilitata per i clienti con Multi-AZ (Availability Zone) con standby.

OpenSearch versione 2.15

Per la OpenSearch versione 2.15, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search/pipeline`

- /_analyze
- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
  - cluster.search.request.slowlog.level
  - cluster.search.request.slowlog.threshold.warn
  - cluster.search.request.slowlog.threshold.info
  - cluster.search.request.slowlog.threshold.debug
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_plugins/_asynchronous_search
- /_plugins/_alerting
- /_plugins/_anomaly_detection
- /_plugins/_ism
- /_plugins/_ml
- /_plugins/_notifications
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_security_analytics
- /_plugins/_sm
- /_plugins/_sql
- /_percolate
- /_rank_eval
- /_search/point_in_time
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

  - `cluster.search.request.slowlog.threshold.trace`
  - `search.phase_took_enabled`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_dashboards`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 2.13

Per la OpenSearch versione 2.13, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-*
- `/_delete_by_query` [1]
- `/_explain`
- `/_refresh`
- `/_reindex` [1]

name /_forcemerge , /*index-name* /update/*id* e /*index-name* /_close).
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
  - cluster.search.request.slowlog.level
  - cluster.search.request.slowlog.threshold.warn
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_ltr
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_plugins/_asynchronous_search
- /_plugins/_alerting
- /_plugins/_anomaly_detection
- /_plugins/_ism
- /_plugins/_ml
- /_plugins/_notifications
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_security_analytics
- /_plugins/_sm
- /_plugins/_sql
- /_percolate
- /_rank_eval
- /_render
- /_resolve/index
- /_rollover
- /_scripts [3]
- /_search[2]
- /_search/pipeline
- /_search/point_in_time
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

    - `cluster.search.request.slowlog.threshold.info`
    - `cluster.search.request.slowlog.threshold.debug`
    - `cluster.search.request.slowlog.threshold.trace`
    - `search.phase_took_enabled`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_dashboards`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 2.11

Per la OpenSearch versione 2.11, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_plugins/_asynchronous_search`
- `/_plugins/_alerting`
- `/_plugins/_anomaly_detection`
- `/_plugins/_ism`
- `/_plugins/_ml`
- `/_plugins/_notifications`
- `/_plugins/_ppl`
- `/_plugins/_security`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search/pipeline`
- `/_search/point_in_time`
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

- indices.breaker.total.limit
- cluster.max_shards_per_node
- /_cluster/state
- /_cluster/stats
- /_count
- /_dashboards
- /_plugins/_security_analytics
- /_plugins/_sm
- /_plugins/_sql
- /_percolate
- /_rank_eval

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 2.9

Per la OpenSearch versione 2.9, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio `/index-name /_forcemerge` ,
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_refresh`
- `/_reindex` [1]
- `/_render`

/*index-name* /update/*id* e /*index-name* /_close).
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_cluster/state
- /_cluster/stats
- /_count
- /_dashboards
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_ltr
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_plugins/_asynchronous_search
- /_plugins/_alerting
- /_plugins/_anomaly_detection
- /_plugins/_ism
- /_plugins/_ml
- /_plugins/_notifications
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_security_analytics
- /_plugins/_sm
- /_plugins/_sql
- /_percolate
- /_rank_eval
- /_resolve/index
- /_rollover
- /_scripts [3]
- /_search[2]
- /_search/pipeline
- /_search/point_in_time
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 2.7

Per OpenSearch 2.7, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search/point_in_time`
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]

- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
  - `cluster.max_shards_per_node`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_dashboards`
- `/_plugins/_asynchronous_search`
- `/_plugins/_alerting`
- `/_plugins/_anomaly_detection`
- `/_plugins/_ism`
- `/_plugins/_ml`
- `/_plugins/_notifications`
- `/_plugins/_ppl`
- `/_plugins/_security`
- `/_plugins/_security_analytics`
- `/_plugins/_sm`
- `/_plugins/_sql`
- `/_percolate`
- `/_rank_eval`
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.

4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called "Differenze significative tra le API". Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called "Riduzione".

OpenSearch versione 2.5

Per OpenSearch 2.5, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_plugins/_asynchronous_search`
- `/_plugins/_alerting`
- `/_plugins/_anomaly_detection`
- `/_plugins/_ism`
- `/_plugins/_ml`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search/point_in_time`
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_cluster/state
- /_cluster/stats
- /_count
- /_dashboards
- /_plugins/_notifications
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_security_analytics
- /_plugins/_sm
- /_plugins/_sql
- /_percolate
- /_rank_eval

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 2.3

Per OpenSearch 2.3, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
  - `cluster.max_shards_per_node`
- `/_cluster/state`
- `/_cluster/stats`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_plugins/_asynchronous_search`
- `/_plugins/_alerting`
- `/_plugins/_anomaly_detection`
- `/_plugins/_ism`
- `/_plugins/_ml`
- `_plugins/_notifications`
- `/_plugins/_ppl`
- `/_plugins/_security`
- `/_plugins/_sql`
- `/_percolate`
- `/_rank_eval`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

- `/_count`
- `/_dashboards`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 1.3

Per OpenSearch 1.3, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`

- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_cluster/state
- /_cluster/stats
- /_count
- /_dashboards
- /_msearch
- /_mtermvectors
- /_nodes
- /_plugins/_asynchronous_search
- /_plugins/_alerting
- /_plugins/_anomaly_detection
- /_plugins/_ism
- /_plugins/_ml
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_sql
- /_percolate
- /_rank_eval
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione /_tasks insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per /_search/scroll con un corpo del messaggio devono specificare il valore "Content-Length" nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei scroll_id valori, utilizzate il

corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.

3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 1.2

Per OpenSearch 1.2, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_plugins/_asynchronous_search`
- `/_plugins/_alerting`
- `/_plugins/_anomaly_detection`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]

- `action.auto_create_index`
- `action.search.shard_count.limit`
- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`
- `cluster.max_shards_per_node`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_dashboards`
- `/_plugins/_ism`
- `/_plugins/_ppl`
- `/_plugins/_security`
- `/_plugins/_sql`
- `/_percolate`
- `/_rank_eval`
- `/_validate`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 1.1

Per OpenSearch 1.1, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio `/index-name /_forcemerge`, `/index-name /update/id` e `/index-name /_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
- `/_delete_by_query`[1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_plugins/_asynchronous_search`
- `/_plugins/_alerting`
- `/_plugins/_anomaly_detection`
- `/_plugins/_ism`
- `/_plugins/_ppl`
- `/_plugins/_security`
- `/_plugins/_sql`
- `/_plugins/_transforms`
- `/_percolate`
- `/_refresh`
- `/_reindex`[1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts`[3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query`[1]
- `/_validate`

- `indices.breaker.total.limit`
- `cluster.max_shards_per_node`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_dashboards`
- `/_rank_eval`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

OpenSearch versione 1.0

Per la OpenSearch versione 1.0, OpenSearch Service supporta le seguenti operazioni. Per informazioni sulla maggior parte delle operazioni, consulta il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.

- Tutte le operazioni nel percorso dell'indice (ad esempio `/index-name /_forcemerge` ,
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_refresh`
- `/_reindex` [1]
- `/_render`

/*index-name* /update/*id* e /*index-name* /_close).

- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_cluster/state
- /_cluster/stats
- /_count
- /_dashboards
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_ltr
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_plugins/_asynchronous_search
- /_plugins/_alerting
- /_plugins/_anomaly_detection
- /_plugins/_ism
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_sql
- /_plugins/_transforms
- /_percolate
- /_rank_eval
- /_resolve/index
- /_rollover
- /_scripts [3]
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 7.10

Per Elasticsearch 7.10, OpenSearch Service supporta le seguenti operazioni.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_index_template` [6]
- `/_ingest/pipeline`
- `/_index_template`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`

- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_cluster/state
- /_cluster/stats
- /_count
- /_opendistro/_alerting
- /_opendistro/_asynchronous_search
- /_opendistro/_anomaly_detection
- /_opendistro/_ism
- /_opendistro/_ppl
- /_opendistro/_security
- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_plugins/_replication
- /_rank_eval
- /_tasks
- /_template [6]
- /_update_by_query [1]
- /_validate

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione /_tasks insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per /_search/scroll con un corpo del messaggio devono specificare il valore "Content-Length" nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei scroll_id valori, utilizzate il corpo della richiesta, non la stringa di query, per passare scroll_id i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

6. I modelli di indice legacy (`_template`) sono stati sostituiti da modelli componibili (`_index_template`) a partire da Elasticsearch 7.8. I modelli componibili hanno la precedenza sui modelli legacy. Se nessun modello componibile corrisponde a un determinato indice, un modello legacy può comunque corrispondere ed essere applicato. L'`_template`operazione funziona ancora nelle versioni successive di Elasticsearch OSS, ma le chiamate GET ai due tipi di modello restituiscono risultati diversi. OpenSearch

Elasticsearch versione 7.9

Per Elasticsearch 7.9, OpenSearch Service supporta le seguenti operazioni.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_index_template` [6]
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_opendistro/_anomaly_detection`
- `/_opendistro/_ism`
- `/_opendistro/_ppl`
- `/_opendistro/_security`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template` [6]
- `/_update_by_query` [1]
- `/_validate`

  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_cluster/state
- /_cluster/stats
- /_count

- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_rank_eval

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle OpenSearch operazioni generiche supportate dal OpenSearch servizio e non include le operazioni supportate specifiche del plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.
6. I modelli di indice legacy (`_template`) sono stati sostituiti da modelli componibili (`_index_template`) a partire da Elasticsearch 7.8. I modelli componibili hanno la precedenza sui modelli legacy. Se nessun modello componibile corrisponde a un determinato indice, un modello legacy può comunque corrispondere ed essere applicato. L'`_template`operazione funziona ancora nelle versioni successive di Elasticsearch OSS, ma le chiamate GET ai due tipi di modello restituiscono risultati diversi. OpenSearch

Elasticsearch versione 7.8

Per Elasticsearch 7.8, OpenSearch Service supporta le seguenti operazioni.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_index_template` [6]
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_opendistro/_anomaly_detection`
- `/_opendistro/_ism`
- `/_opendistro/_security`
- `/_opendistro/_sql`
- `/_percolate`
- `/_plugin/kibana`
- `/_rank_eval`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template` [6]
- `/_update_by_query` [1]
- `/_validate`

- `cluster.max_shards_per_node`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.
6. I modelli di indice legacy (`_template`) sono stati sostituiti da modelli componibili (`_index_template`) a partire da Elasticsearch 7.8. I modelli componibili hanno la precedenza sui modelli legacy. Se nessun modello componibile corrisponde a un determinato indice, un modello legacy può comunque corrispondere ed essere applicato. L'`_template`operazione funziona ancora nelle versioni successive di Elasticsearch OSS, ma le chiamate GET ai due tipi di modello restituiscono risultati diversi. OpenSearch

Elasticsearch versione 7.7

Per Elasticsearch 7.7, OpenSearch Service supporta le seguenti operazioni.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]

- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_ltr
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_opendistro/_alerting
- /_opendistro/_anomaly_detection
- /_opendistro/_ism
- /_opendistro/_security
- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_rank_eval
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione /_tasks insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per /_search/scroll con un corpo del messaggio devono specificare il valore "Content-Length" nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei scroll_id valori, utilizzate il corpo della richiesta, non la stringa di query, per passare scroll_id i valori al servizio. OpenSearch

3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 7.4

Per Elasticsearch 7.4, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` , */index-name* `/update/`*id* e */index-name* `/_close`).
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_opendistro/_anomaly_detection`
- `/_opendistro/_ism`
- `/_opendistro/_security`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

- indices.breaker.fielddata.limit
- indices.breaker.request.limit
- indices.breaker.total.limit
- cluster.max_shards_per_node

- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_rank_eval

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione /_tasks insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per /_search/scroll con un corpo del messaggio devono specificare il valore "Content-Length" nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei scroll_id valori, utilizzate il corpo della richiesta, non la stringa di query, per passare scroll_id i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 7.1

Per Elasticsearch 7.1, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* /_forcemerge e */index-name* /update/*id*) tranne */index-name* /_close.
- /_alias
- /_cluster/state
- /_cluster/stats
- /_count
- /_delete_by_query [1]
- /_explain
- /_refresh
- /_reindex [1]
- /_render
- /_rollover
- /_scripts [3]

- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_opendistro/_alerting
- /_opendistro/_ism
- /_opendistro/_security
- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_rank_eval
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione /_tasks insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per /_search/scroll con un corpo del messaggio devono specificare il valore "Content-Length" nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei scroll_id valori, utilizzate il corpo della richiesta, non la stringa di query, per passare scroll_id i valori al servizio. OpenSearch

3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 6.8

Per Elasticsearch 6.8, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` e */index-name* `/update/`*id*) tranne */index-name* `/_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_opendistro/_ism`
- `/_opendistro/_security`
- `/_opendistro/_sql`
- `/_percolate`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`
- `cluster.max_shards_per_node`
- `cluster.blocks.read_only`

- `/_plugin/kibana`
- `/_rank_eval`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 6.7

Per Elasticsearch 6.7, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` e

- `/_cluster/state`
- `/_cluster/stats`
- `/_count`

- `/_refresh`
- `/_reindex` [1]
- `/_render`

/*index-name* /update/*id*) tranne /*index-name* /_close.
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
  - cluster.max_shards_per_node
- /_delete_by_query [1]
- /_explain
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_opendistro/_alerting
- /_opendistro/_security
- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_rank_eval
- /_rollover
- /_scripts [3]
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione /_tasks insieme a queste altre operazioni per verificare il corretto completamento delle richieste.

2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 6.5

Per Elasticsearch 6.5, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` e */index-name* `/update/`*id*) tranne */index-name* `/_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]

- `action.auto_create_index`
- `action.search.shard_count.limit`
- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`
- `/_opendistro/_sql`
- `/_percolate`
- `/_plugin/kibana`
- `/_rank_eval`
- `/_validate`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 6.4

Per Elasticsearch 6.4, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio `/index-name /_forcemerge` e
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_refresh`
- `/_reindex` [1]
- `/_render`

/*index-name* /update/*id*) tranne /*index-name* /_close.
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (eccetto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings per numerose proprietà[4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
- /_delete_by_query [1]
- /_explain
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_opendistro/_alerting
- /_percolate
- /_plugin/kibana
- /_rank_eval
- /_rollover
- /_scripts [3]
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione /_tasks insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per /_search/scroll con un corpo del messaggio devono specificare il valore "Content-Length" nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei scroll_id valori, utilizzate

il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch

3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 6.3

Per Elasticsearch 6.3, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio `/index-name /_forcemerge` e `/index-name /update/id`) tranne `/index-name /_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_percolate`
- `/_plugin/kibana`
- `/_rank_eval`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

- `action.search.shard_count.limit`
- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 6.2

Per Elasticsearch 6.2, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` e */index-name* `/update/`*id*) tranne */index-name* `/_close`.
- `/_alias`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]

- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_percolate`
- `/_plugin/kibana`
- `/_rank_eval`
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.

4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called "Differenze significative tra le API". Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called "Riduzione".

Elasticsearch versione 6.0

Per Elasticsearch 6.0, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio `/index-name /_forcemerge` e `/index-name /update/id`) tranne `/index-name /_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 5.6

Per Elasticsearch 5.6, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` e */index-name* `/update/`*id*) tranne */index-name* `/_close`.
- `/_alias`
- `/_aliases`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search` [2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink` [5]

- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_refresh`
- `/_reindex` [1]
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo PUT. Per informazioni sul metodo GET, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di

Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.

5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 5.5

Per Elasticsearch 5.5, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` e */index-name* `/update/`*id*) tranne */index-name* `/_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[4]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

- `indices.breaker.request.limit`
- `indices.breaker.total.limit`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Per considerazioni sull'uso degli script, consulta the section called “Altre risorse supportate”.
4. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
5. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 5.3

Per Elasticsearch 5.3, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` e */index-name* `/update/`*id*) tranne */index-name* `/_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_render`
- `/_rollover`
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[4]
- `/_snapshot`
- `/_stats`
- `/_status`

- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà[3]:
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_refresh`
- `/_reindex` [1]
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al servizio. OpenSearch
3. Riferimento al metodo `PUT`. Per informazioni sul metodo `GET`, consulta the section called “Differenze significative tra le API”. Questo elenco si riferisce solo alle operazioni generiche di Elasticsearch supportate da OpenSearch Service e non include le operazioni supportate dai plug-in per il rilevamento delle anomalie, ISM e così via.
4. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 5.1

Per Elasticsearch 5.1, Service supporta le seguenti operazioni. OpenSearch

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forcemerge` e */index-name* `/update/`*id*) tranne */index-name* `/_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` per numerose proprietà (solo PUT):
  - `action.auto_create_index`
  - `action.search.shard_count.limit`
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_search` [2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink` [3]
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

1. Le modifiche di configurazione del cluster potrebbero interrompere queste operazioni prima del loro completamento. È consigliabile usare l'operazione `/_tasks` insieme a queste altre operazioni per verificare il corretto completamento delle richieste.
2. Le richieste DELETE per `/_search/scroll` con un corpo del messaggio devono specificare il valore `"Content-Length"` nell'intestazione HTTP. La maggior parte dei client aggiunge questa intestazione di default. Per evitare problemi con = i caratteri nei `scroll_id` valori, utilizzate il corpo della richiesta, non la stringa di query, per passare `scroll_id` i valori al OpenSearch servizio.
3. Per informazioni, consulta the section called “Riduzione”.

Elasticsearch versione 2.3

Per Elasticsearch 2.3, OpenSearch Service supporta le seguenti operazioni.

- Tutte le operazioni nel percorso dell'indice (ad esempio */index-name* `/_forceme rge` e */index-name* `/_recovery`) tranne */index-name* `/_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cache/clear` (solo indice)
- `/_cat` (eccetto `/_cat/nodeattrs`)
- `/_cluster/health`
- `/_cluster/settings` per numerose proprietà (solo PUT):
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
  - `threadpool.get.queue_size`
- `/_cluster/stats`
- `/_count`
- `/_flush`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_refresh`
- `/_render`
- `/_search`
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_template`

- `threadpool.bulk.queue_size`
- `threadpool.index.queue_size`
- `threadpool.percolate.queue_size`
- `threadpool.search.queue_size`
- `threadpool.suggest.queue_size`

Elasticsearch versione 1.5

Per Elasticsearch 1.5, OpenSearch Service supporta le seguenti operazioni.

- Tutte le operazioni nel percorso dell'indice, ad esempio */index-name* `/_optimize` e */index-name* `/_warmer`, tranne */index-name* `/_close`.
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat`
- `/_cluster/health`
- `/_cluster/settings` per numerose proprietà (solo PUT):
  - `indices.breaker.fielddata.limit`
  - `indices.breaker.request.limit`
  - `indices.breaker.total.limit`
  - `threadpool.get.queue_size`
  - `threadpool.bulk.queue_size`
  - `threadpool.index.queue_size`
- `/_cluster/stats`
- `/_count`
- `/_flush`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_plugin/kibana3`
- `/_plugin/migration`
- `/_refresh`
- `/_search`
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_template`

- `threadpool.percolate.queue_size`
- `threadpool.search.queue_size`
- `threadpool.suggest.queue_size`

Quote OpenSearch del servizio Amazon

Il tuo AWS account ha delle quote predefinite, precedentemente denominate limiti, per ogni servizio. AWS Salvo diversa indicazione, ogni quota si applica a una Regione specifica.

Per visualizzare le quote per i domini e le istanze del OpenSearch servizio, Amazon OpenSearch Serverless e Amazon OpenSearch Ingestion, consulta le quote di Amazon OpenSearch Service nel. Riferimenti generali di AWS

Per visualizzare le quote per OpenSearch Service in Console di gestione AWS, apri la console Service Quotas. Nel riquadro di navigazione, scegli AWS servizi e seleziona Amazon OpenSearch Service. Per richiedere un aumento delle quote, consultare Richiesta di aumento delle quote nella Guida dell'utente di Service Quotas.

Quote di archiviazione Warm Node

La tabella seguente elenca i tipi di istanze warm node e la quantità massima di storage che ogni tipo può utilizzare. Per OI2 le istanze OpenSearch Optimized, l'80% dello storage locale è disponibile come cache e lo storage warm indirizzabile massimo è 5 volte la dimensione di archiviazione della cache.

Ad esempio, se un' OI2 istanza ha 468 GB di storage locale:

- Archiviazione cache = 375 GB (80% della memoria locale)
- Memoria calda indirizzabile massima = 1875 GB (cache 5 x 375 GB)

Tipo di istanza	Archiviazione delle istanze (GB)	Dimensione della cache (GB)	Warm Storage massimo indirizzabile (GB)
oi2.large.search	468	375	1875

Tipo di istanza	Archiviazione delle istanze (GB)	Dimensione della cache (GB)	Warm Storage massimo indirizzabile (GB)
oi2.xlarge.search	937	750	3750
oi2.2xlarge.search	1875	1500	7500
oi2.4xlarge.search	3750	3000	15000
oi 2.8 x large.search	7500	6000	30000

UltraWarm quote di archiviazione

La tabella seguente elenca i tipi di UltraWarm istanza e la quantità massima di storage che ogni tipo può utilizzare. Per ulteriori informazioni su UltraWarm, vederethe section called “UltraWarm archiviazione”.

Tipo di istanza	Spazio di archiviazione massimo
ultrawarm1.medium.search	1,5 TiB
ultrawarm1.large.search	20 TiB

Numero di nodi dati per AZ

La tabella seguente elenca il numero totale di nodi di dati per l'implementazione AZ. Il limite complessivo indica il numero di nodi di dati per limite, incluso il numero di nodi di dati caldi e caldi. Lo storage può essere utilizzato da ciascun tipo.

Configurazione AZ	Limite di numero di nodi caldi	Limite di conteggio dei nodi caldi	Limite complessivo (caldo+caldo)
1 - AZ	334	250	334

Configurazione AZ	Limite di numero di nodi caldi	Limite di conteggio dei nodi caldi	Limite complessivo (caldo+caldo)
2 - AZ	668	500	668
3 - AZ	1002	750	1002

Limite totale di nodi per famiglia di istanze

La tabella seguente elenca il limite totale di nodi per famiglia di istanze.

Famiglia di istanze	ElasticSearch OpenSearch fino a 2,15	OpenSearch 2.17 e versioni successive	Limite predefinito
T2	10	10	10
T3	10	10	10
M3, C4, M4, R4, C5, M5, R5, I2, I3	10	200	80
Gravitone 2, Gravitone 3, Gravitone 4	200	400	80
C7i, R7i, M7i, i4i, i4g, i8g, i7i, i8ge	200	400	80
OR1.medium.search OR1.ricerca di grandi dimensioni OR2.medium.search OR2.ricerca di grandi dimensioni OM2.large.ricerca OI2.large.ricerca	200	400	80
OR1.xlarge.search e versioni successive	200	1002	80

Famiglia di istanze	ElasticSearch OpenSearch fino a 2,15	OpenSearch 2.17 e versioni successive	Limite predefinito
OR2.xlarge.search e versioni successive OM2.xlarge.search e versioni successive OI2.xlarge.search e versioni successive			
Ultra caldo 1	150	750	150

Quote delle dimensioni dei volumi EBS

La tabella seguente mostra le dimensioni minime e massime dei volumi EBS per ogni tipo di istanza supportato dal servizio. OpenSearch Per informazioni sui tipi di istanza che includono lo storage delle istanze e dettagli hardware aggiuntivi, consulta i prezzi OpenSearch di Amazon Service.

- Se scegli lo storage magnetico nel tipo di volume EBS quando crei il tuo dominio, la dimensione massima del volume è di 100 GiB per tutti i tipi di istanza tranne t2.small e t2.medium e tutte le istanze Graviton (M6g, C6g, R6g e R6gd), che non supportano lo storage magnetico. Per le dimensioni massime elencate nella seguente tabella, selezionare una delle opzioni SSD.
- Alcuni tipi di istanze di generazioni precedenti includono l'archiviazione dell'istanza, ma supportano anche l'archiviazione EBS. Se si sceglie l'archiviazione EBS per uno di questi tipi di istanze, tenere presente che i volumi di archiviazione non sono cumulativi. È possibile utilizzare un volume EBS o l'archiviazione dell'istanza, non entrambi.
- Per i volumi EBS gp3, MB/s i limiti massimi supportati sono 80.000 IOPS e 2.000 velocità effettiva. Tieni presente che i limiti effettivi per il tuo dominio potrebbero essere inferiori a seconda del tipo di istanza. Per i limiti massimi specifici dell'istanza, consulta Limiti per tipo di istanza.

Tipo di istanza	Dimensione minima EBS	Dimensione massima EBS (gp2)	Dimensione massima EBS (gp3)
t2.micro.search	10 GiB	35 GiB	N/D
t2.small.search	10 GiB	35 GiB	N/D
t2.medium.ricerca	10 GiB	35 GiB	N/D
t3.small.search	10 GiB	100 GiB	100 GiB
t3.medium.search	10 GiB	200 GiB	200 GiB
m3.medium.ricerca	10 GiB	100 GiB	N/D
m3.large.search	10 GiB	512 GiB	N/D
m3.xlarge.search	10 GiB	512 GiB	N/D
m3.2 x large.search	10 GiB	512 GiB	N/D
m4.large.search	10 GiB	512 GiB	N/D
m4.xlarge.search	10 GiB	1 TiB	N/D
m4.2xlarge.search	10 GiB	1,5 TiB	N/D
m4.4xlarge.search	10 GiB	1,5 TiB	N/D
m4.10xlarge.search	10 GiB	1,5 TiB	N/D
m5.large.search	10 GiB	512 GiB	1 TiB
m5.xlarge.search	10 GiB	1 TiB	2 TiB
m5.2xlarge.search	10 GiB	1,5 TiB	3 TiB
m5.4xlarge.search	10 GiB	3 TiB	6 TiB
m5.12xlarge.search	10 GiB	9 TiB	18 TiB

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
m6g.large.search	10 GiB	512 GiB	1 TiB
m6g.xlarge.search	10 GiB	1 TiB	2 TiB
m6g.2xlarge.search	10 GiB	1,5 TiB	3 TiB
m6g.4xlarge.ricerca	10 GiB	3 TiB	6 TiB
m6g.8xlarge.search	10 GiB	6 TiB	12 TiB
m6g.12xlarge.ricerca	10 GiB	9 TiB	18 TiB
c4.large.search	10 GiB	100 GiB	N/D
c4.xlarge.search	10 GiB	512 GiB	N/D
c4.2 xlarge.search	10 GiB	1 TiB	N/D
c4.4xlarge.search	10 GiB	1,5 TiB	N/D
c4.8xlarge.search	10 GiB	1,5 TiB	N/D
c5.large.search	10 GiB	256 GiB	256 GiB
c5.xlarge.search	10 GiB	512 GiB	512 GiB
c5.2xlarge.search	10 GiB	1 TiB	1 TiB
c5.4xlarge.search	10 GiB	1,5 TiB	1,5 TiB
c5.9xlarge.search	10 GiB	3,5 TiB	3,5 TiB
c5.18xlarge.search	10 GiB	7 TiB	7 TiB
c6g.large.search	10 GiB	256 GiB	256 GiB
c6g.xlarge.search	10 GiB	512 GiB	512 GiB

Tipo di istanza	Dimensione minima EBS	Dimensione massima EBS (gp2)	Dimensione massima EBS (gp3)
c6g.2xlarge.search	10 GiB	1 TiB	1 TiB
c6g.4xlarge.search	10 GiB	1,5 TiB	1,5 TiB
c6g.8xlarge.search	10 GiB	3 TiB	3 TiB
c6g.12xlarge.search	10 GiB	4,5 TiB	4,5 TiB
r3.large.search	10 GiB	512 GiB	N/D
r3.xlarge.search	10 GiB	512 GiB	N/D
r3.2xlarge.search	10 GiB	512 GiB	N/D
r3.4xlarge.search	10 GiB	512 GiB	N/D
r3.8xlarge.search	10 GiB	512 GiB	N/D
r4.large.search	10 GiB	1 TiB	N/D
r4.xlarge.search	10 GiB	1,5 TiB	N/D
r4.2xlarge.search	10 GiB	1,5 TiB	N/D
r4.4xlarge.search	10 GiB	1,5 TiB	N/D
r4.8xlarge.search	10 GiB	1,5 TiB	N/D
r4.16xlarge.search	10 GiB	1,5 TiB	N/D
r5.large.search	10 GiB	1 TiB	2 TiB
r5.xlarge.search	10 GiB	1,5 TiB	3 TiB
r5.2xlarge.search	10 GiB	3 TiB	6 TiB
r5.4xlarge.search	10 GiB	6 TiB	12 TiB

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
r5.12xlarge.search	10 GiB	12 TiB	24 TiB
r6g.large.search	10 GiB	1 TiB	2 TiB
r6g.xlarge.search	10 GiB	1,5 TiB	3 TiB
r6g.2xlarge.search	10 GiB	3 TiB	6 TiB
r6g.4xlarge.search	10 GiB	6 TiB	12 TiB
r6g.8xlarge.search	10 GiB	8 TiB	16 TiB
r6g.12xlarge.search	10 GiB	12 TiB	24 TiB
r6gd.large.search	N/D	N/D	N/D
r6gd.xlarge.search	N/D	N/D	N/D
r6gd.2xlarge.search	N/D	N/D	N/D
r6gd.4xlarge.search	N/D	N/D	N/D
r6gd.8xlarge.search	N/D	N/D	N/D
r6gd.12xlarge.search	N/D	N/D	N/D
r6gd.16xlarge.search	N/D	N/D	N/D
i2.xlarge.search	10 GiB	512 GiB	N/D
i2.2xlarge.search	10 GiB	512 GiB	N/D
i3.large.search	N/D	N/D	N/D
i3.xlarge.search	N/D	N/D	N/D
i3.2 x large.search	N/D	N/D	N/D

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
i3.4xlarge.search	N/D	N/D	N/D
i3.8xlarge.search	N/D	N/D	N/D
i3.16xlarge.search	N/D	N/D	N/D
o 1.medium.search	20 GiB	N/D	768 GiB
o 1.large.search	20 GiB	N/D	15,32 GiB
o 1.xlarge.search	20 GiB	N/D	3 TiB
o 1,2 x large.search	20 GiB	N/D	6 TiB
o 1,4 x large.search	20 GiB	N/D	12 TiB
o 1,8 x large.search	20 GiB	N/D	16 TiB
o 1,12 xlarge.search	20 GiB	N/D	24 TiB
o 1,16 x large.search	20 GiB	N/D	36 TiB
o 2.medium.search	20 GiB	N/D	768 GiB
o 2.large.search	20 GiB	N/D	15,32 GiB
o 2.xlarge.search	20 GiB	N/D	3 TiB
o 2,2 x large.search	20 GiB	N/D	6 TiB
o 2,4xlarge.search	20 GiB	N/D	12 TiB
o 2,8 x large.search	20 GiB	N/D	16 TiB
o 2,12 xlarge.search	20 GiB	N/D	24 TiB
o 2,16 x large.search	20 GiB	N/D	36 TiB

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
om 2. Large.search	20 GiB	N/D	768 GiB
om2.xlarge.search	20 GiB	N/D	2 TiB
om 2.2 x large.search	20 GiB	N/D	3 TiB
om 2.4xlarge.search	20 GiB	N/D	6 TiB
om 2.8xlarge.search	20 GiB	N/D	12 TiB
om2.12xlarge.search	20 GiB	N/D	18 TiB
om 2.16xlarge.search	20 GiB	N/D	24 TiB
im4gn.large.search	N/D	N/D	N/D
im4gn.xlarge.search	N/D	N/D	N/D
im4gn.2xlarge.search	N/D	N/D	N/D
im4gn.4xlarge.search	N/D	N/D	N/D
im4gn.8xlarge.search	N/D	N/D	N/D
im4gn.16xlarge.search	N/D	N/D	N/D
c7g.large.search	10 GiB	N/D	256 GiB
C7g.xlarge.search	10 GiB	N/D	512 GiB
c7g.2xlarge.search	10 GiB	N/D	1 TiB
c7g.4xlarge.search	10 GiB	N/D	1,5 TiB
c7g.8xlarge.search	10 GiB	N/D	3 TiB
C7G.12XLarge.Ricerca	10 GiB	N/D	4,5 TiB

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
C7G.16xlarge.search	10 GiB	N/D	6 TiB
m7g.medium.search	10 GiB	N/D	512 GiB
m7g.large.search	10 GiB	N/D	768 GiB
m7g.xlarge.search	10 GiB	N/D	2 TiB
M7G.2XLarge.Ricerca	10 GiB	N/D	3 TiB
m7g.4xlarge.search	10 GiB	N/D	6 TiB
m7g.8xlarge.search	10 GiB	N/D	12 TiB
M7G.12XLarge.Ricerca	10 GiB	N/D	18 TiB
M7G.16xLarge.Ricerca	10 GiB	N/D	24 TiB
R7g.medium.search	10 GiB	N/D	768 GiB
R7g.large.search	10 GiB	N/D	1,5 TiB
R7g.xlarge.search	10 GiB	N/D	3 TiB
R7G.2XLarge.Ricerca	10 GiB	N/D	6 TiB
R7g.4xlarge.ricerca	10 GiB	N/D	12 TiB
R7g.8xlarge.search	10 GiB	N/D	16 TiB
R7g.12xlarge.ricerca	10 GiB	N/D	24 TiB
R7g.16xlarge.ricerca	10 GiB	N/D	36 TiB
R7gd.large.cerca	N/D	N/D	N/D
r7gd.xlarge.search	N/D	N/D	N/D

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
r7gd.2xlarge.ricerca	N/D	N/D	N/D
r7gd.4xlarge.search	N/D	N/D	N/D
r7gd.8xlarge.search	N/D	N/D	N/D
r7gd.12xlarge.ricerca	N/D	N/D	N/D
R7GD.16xlarge.search	N/D	N/D	N/D
i4i.large.search	N/D	N/D	N/D
i4i.xlarge.search	N/D	N/D	N/D
i4i.2xlarge.search	N/D	N/D	N/D
i4i.4xlarge.search	N/D	N/D	N/D
i4i.8xlarge.search	N/D	N/D	N/D
i4i.12xlarge.search	N/D	N/D	N/D
i4i.16xlarge.search	N/D	N/D	N/D
i4i.24xlarge.search	N/D	N/D	N/D
i4i.32xlarge.search	N/D	N/D	N/D
i4g.large.search	N/D	N/D	N/D
i4g.xlarge.search	N/D	N/D	N/D
i4g.2xlarge.search	N/D	N/D	N/D
i4g.4xlarge.ricerca	N/D	N/D	N/D
i4g.8xlarge.search	N/D	N/D	N/D

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
i4g.16xlarge.search	N/D	N/D	N/D
i8g.large.search	N/D	N/D	N/D
i8g.xlarge.search	N/D	N/D	N/D
i8g.2xlarge.search	N/D	N/D	N/D
i8g.4xlarge.search	N/D	N/D	N/D
i8g.8xlarge.search	N/D	N/D	N/D
i8g.12xlarge.search	N/D	N/D	N/D
i8g.16xlarge.search	N/D	N/D	N/D
i7i.large.search	N/D	N/D	N/D
i7i.xlarge.search	N/D	N/D	N/D
i7i.2xlarge.search	N/D	N/D	N/D
i7i.4xlarge.search	N/D	N/D	N/D
i7i.8xlarge.search	N/D	N/D	N/D
i7i.12xlarge.search	N/D	N/D	N/D
i7i.16xlarge.search	N/D	N/D	N/D
i8ge.large.search	N/D	N/D	N/D
i8ge.xlarge.search	N/D	N/D	N/D
i8ge.2xlarge.search	N/D	N/D	N/D
i8ge.3xlarge.search	N/D	N/D	N/D

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
i8ge.6xlarge.search	N/D	N/D	N/D
i8ge.12xlarge.search	N/D	N/D	N/D
i8ge.18xlarge.search	N/D	N/D	N/D
c7i.large.search	10 GiB	N/D	256 GiB
c7i.xlarge.search	10 GiB	N/D	512 GiB
c7i.2xlarge.search	10 GiB	N/D	1 TiB
c7i.4xlarge.search	10 GiB	N/D	1,5 TiB
c7i.8xlarge.search	10 GiB	N/D	3 TiB
c7i.12xlarge.search	10 GiB	N/D	4,5 TiB
c7i.16xlarge.search	10 GiB	N/D	6 TiB
m7i.large.search	10 GiB	N/D	768 GiB
m7i.xlarge.search	10 GiB	N/D	2 TiB
m7i.2xlarge.search	10 GiB	N/D	3 TiB
m7i.4xlarge.search	10 GiB	N/D	6 TiB
m7i.8xlarge.search	10 GiB	N/D	12 TiB
m7i.12xlarge.search	10 GiB	N/D	18 TiB
m7i.16xlarge.search	10 GiB	N/D	24 TiB
r7i.large.search	10 GiB	N/D	1,5 TiB
r7i.xlarge.search	10 GiB	N/D	3 TiB

Tipo di istanza	Dimensione minima EBS	Dimensione massima EBS (gp2)	Dimensione massima EBS (gp3)
r7i.2xlarge.search	10 GiB	N/D	6 TiB
r7i.4xlarge.search	10 GiB	N/D	12 TiB
r7i.8xlarge.search	10 GiB	N/D	16 TiB
r7i.12xlarge.search	10 GiB	N/D	24 TiB
r7i.12xlarge.search	10 GiB	N/D	36 TiB
c8g.large	10 GiB	N/D	256 GiB
c8g.xlarge	10 GiB	N/D	512 GiB
c8g.2xlarge	10 GiB	N/D	1 TiB
c8g.4xlarge	10 GiB	N/D	1,5 TiB
c8g.8xlarge	10 GiB	N/D	3 TiB
c8g.12xlarge	10 GiB	N/D	4,5 TiB
c8g.16xlarge	10 GiB	N/D	6 TiB
m8g.medium	10 GiB	N/D	512 GiB
m8g.large	10 GiB	N/D	768 GiB
m8g.xlarge	10 GiB	N/D	2 TiB
m8g.2xlarge	10 GiB	N/D	3 TiB
m8g.4xlarge	10 GiB	N/D	6 TiB
m8g.8xlarge	10 GiB	N/D	12 TiB
m8g.12xlarge	10 GiB	N/D	18 TiB

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
m8g.16xlarge	10 GiB	N/D	24 TiB
r8g.medium	10 GiB	N/D	768 GiB
r8g.large	10 GiB	N/D	15,32 GiB
r8g.xlarge	10 GiB	N/D	3 TiB
r8g.2xlarge	10 GiB	N/D	6 TiB
r8g.4xlarge	10 GiB	N/D	12 TiB
r8g.8xlarge	10 GiB	N/D	16 TiB
r8g.12xlarge	10 GiB	N/D	24 TiB
r8g.16xlarge	10 GiB	N/D	36 TiB
r8gd.medium	N/D	N/D	N/D
r8gd.large	N/D	N/D	N/D
r8gd.xlarge	N/D	N/D	N/D
r8gd.2xlarge	N/D	N/D	N/D
r8gd.4xlarge	N/D	N/D	N/D
r8gd.8xlarge	N/D	N/D	N/D
r8gd.12xlarge	N/D	N/D	N/D
r8gd.16xlarge	N/D	N/D	N/D
oi2.large.search	N/D	N/D	N/D
oi2.xlarge.search	N/D	N/D	N/D

Tipo di istanza	Dimensione minima EBS	Dimension e massima EBS (gp2)	Dimensione massima EBS (gp3)
oi2.2xlarge.search	N/D	N/D	N/D
oi2.4xlarge.search	N/D	N/D	N/D
oi2.8 xlarge.search	N/D	N/D	N/D
oi2.12xlarge.search	N/D	N/D	N/D
oi2.16xlarge.search	N/D	N/D	N/D

Quote di rete

La tabella seguente mostra la dimensione massima dei payload delle richieste HTTP.

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
t2.micro.search	10 MiB
t2.small.search	10 MiB
t2.medium.ricerca	10 MiB
t3.small.search	10 MiB
t3.medium.search	10 MiB
m3.medium.ricerca	10 MiB
m3.large.search	10 MiB
m3.xlarge.search	100 MiB
m3.2 x large.search	100 MiB
m4.large.search	10 MiB

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
m4.xlarge.search	100 MiB
m4.2xlarge.search	100 MiB
m4.4xlarge.search	100 MiB
m4.10xlarge.search	100 MiB
m5.large.search	10 MiB
m5.xlarge.search	100 MiB
m5.2xlarge.search	100 MiB
m5.4xlarge.search	100 MiB
m5.12xlarge.search	100 MiB
m6g.large.search	10 MiB
m6g.xlarge.search	100 MiB
m6g.2xlarge.search	100 MiB
m6g.4xlarge.ricerca	100 MiB
m6g.8xlarge.search	100 MiB
m6g.12xlarge.ricerca	100 MiB
c4.large.search	10 MiB
c4.xlarge.search	100 MiB
c4.2 xlarge.search	100 MiB
c4.4xlarge.search	100 MiB
c4.8xlarge.search	100 MiB

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
c5.large.search	10 MiB
c5.xlarge.search	100 MiB
c5.2xlarge.search	100 MiB
c5.4xlarge.search	100 MiB
c5.9xlarge.search	100 MiB
c5.18xlarge.search	100 MiB
c6g.large.search	10 MiB
c6g.xlarge.search	100 MiB
c6g.2xlarge.search	100 MiB
c6g.4xlarge.search	100 MiB
c6g.8xlarge.search	100 MiB
c6g.12xlarge.search	100 MiB
r3.large.search	10 MiB
r3.xlarge.search	100 MiB
r3.2xlarge.search	100 MiB
r3.4xlarge.search	100 MiB
r3.8xlarge.search	100 MiB
r4.large.search	100 MiB
r4.xlarge.search	100 MiB
r4.2 xlarge.search	100 MiB

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
r4.4xlarge.search	100 MiB
r4.8xlarge.search	100 MiB
r4.16xlarge.search	100 MiB
r5.large.search	100 MiB
r5.xlarge.search	100 MiB
r5.2xlarge.search	100 MiB
r5.4xlarge.search	100 MiB
r5.12xlarge.search	100 MiB
r6g.large.search	100 MiB
r6g.xlarge.search	100 MiB
r6g.2xlarge.search	100 MiB
r6g.4xlarge.search	100 MiB
r6g.8xlarge.search	100 MiB
r6g.12xlarge.search	100 MiB
r6gd.large.search	100 MiB
r6gd.xlarge.search	100 MiB
r6gd.2xlarge.search	100 MiB
r6gd.4xlarge.search	100 MiB
r6gd.8xlarge.search	100 MiB
r6gd.12xlarge.search	100 MiB

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
r6gd.16xlarge.search	100 MiB
i2.xlarge.search	100 MiB
i2.2xlarge.search	100 MiB
i3.large.search	100 MiB
i3.xlarge.search	100 MiB
i3.2 x large.search	100 MiB
i3.4xlarge.search	100 MiB
i3.8xlarge.search	100 MiB
i3.16xlarge.search	100 MiB
o 1.medium.search	10 MiB
o 1.large.search	100 MiB
o 1.xlarge.search	100 MiB
o 1.2xlarge.search	100 MiB
o 1.4xlarge.search	100 MiB
o 1.8xlarge.search	100 MiB
o 1.12xlarge.search	100 MiB
o 1,16 x large.search	100 MiB
o 2.medium.search	100 MiB
o 2.large.search	100 MiB
o 2.xlarge.search	100 MiB

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
o 2.2xlarge.search	100 MiB
o 2.4xlarge.search	100 MiB
o 2.8xlarge.search	100 MiB
o 2.12xlarge.search	100 MiB
o 2.16xlarge.search	100 MiB
om2.large.search	10 MiB
om2.xlarge.search	100 MiB
om2.2xlarge.search	100 MiB
om2.4xlarge.search	100 MiB
om2.8xlarge.search	100 MiB
om2.12xlarge.search	100 MiB
om2.16xlarge.search	100 MiB
im4gn.large.search	100 MiB
im4gn.xlarge.search	100 MiB
im4gn.2xlarge.search	100 MiB
im4gn.4xlarge.search	100 MiB
im4gn.8xlarge.search	100 MiB
im4gn.16xlarge.search	100 MiB
i4i.large.search	N/D
i4i.xlarge.search	N/D

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
i4i.2xlarge.search	N/D
i4i.4xlarge.search	N/D
i4i.8xlarge.search	N/D
i4i.12xlarge.search	N/D
i4i.16xlarge.search	N/D
i4i.24xlarge.search	N/D
i4i.32xlarge.search	N/D
i4g.large.search	N/D
i4g.xlarge.search	N/D
i4g.2xlarge.search	N/D
i4g.4xlarge.ricerca	N/D
i4g.8xlarge.search	N/D
i4g.16xlarge.search	N/D
i8g.large.search	N/D
i8g.xlarge.search	N/D
i8g.2xlarge.search	N/D
i8g.4xlarge.search	N/D
i8g.8xlarge.search	N/D
i8g.12xlarge.search	N/D
i8g.16xlarge.search	N/D

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
i7i.large.search	100 MiB
i7i.xlarge.search	100 MiB
i7i.2xlarge.search	100 MiB
i7i.4xlarge.search	100 MiB
i7i.8xlarge.search	100 MiB
i7i.12xlarge.search	100 MiB
i7i.16xlarge.search	100 MiB
i8ge.large.search	100 MiB
i8ge.xlarge.search	100 MiB
i8ge.2xlarge.search	100 MiB
i8ge.3xlarge.search	100 MiB
i8ge.6xlarge.search	100 MiB
i8ge.12xlarge.search	100 MiB
i8ge.18xlarge.search	100 MiB
c7i.large.search	100 MiB
c7i.xlarge.search	100 MiB
c7i.2xlarge.search	100 MiB
c7i.4xlarge.search	100 MiB
c7i.8xlarge.search	100 MiB
c7i.12xlarge.search	100 MiB

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
c7i.16xlarge.search	100 MiB
m7i.large.search	100 MiB
m7i.xlarge.search	100 MiB
m7i.2xlarge.search	100 MiB
m7i.4xlarge.search	100 MiB
m7i.8xlarge.search	100 MiB
m7i.12xlarge.search	100 MiB
m7i.16xlarge.search	100 MiB
r7i.large.search	100 MiB
r7i.xlarge.search	100 MiB
r7i.2xlarge.search	100 MiB
r7i.4xlarge.search	100 MiB
r7i.8xlarge.search	100 MiB
r7i.12xlarge.search	100 MiB
r7i.16xlarge.search	100 MiB
c8g.large	100 MiB
c8g.xlarge	100 MiB
c8g.2xlarge	100 MiB
c8g.4xlarge	100 MiB
c8g.8xlarge	100 MiB

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
c8g.12xlarge	100 MiB
c8g.16xlarge	100 MiB
m8g.medium	100 MiB
m8g.large	100 MiB
m8g.xlarge	100 MiB
m8g.2xlarge	100 MiB
m8g.4xlarge	100 MiB
m8g.8xlarge	100 MiB
m8g.12xlarge	100 MiB
m8g.16xlarge	100 MiB
r8g.medium	100 MiB
r8g.large	100 MiB
r8g.xlarge	100 MiB
r8g.2xlarge	100 MiB
r8g.4xlarge	100 MiB
r8g.8xlarge	100 MiB
r8g.12xlarge	100 MiB
r8g.16xlarge	100 MiB
r8gd.medium	100 MiB
r8gd.large	100 MiB

Tipo di istanza	Dimensione massima dei payload delle richieste HTTP
r8gd.xlarge	100 MiB
r8gd.2xlarge	100 MiB
r8gd.4xlarge	100 MiB
r8gd.8xlarge	100 MiB
r8gd.12xlarge	100 MiB
r8gd.16xlarge	100 MiB
oi2.large.search	100 MiB
oi2.xlarge.search	100 MiB
oi2.2xlarge.search	100 MiB
oi2.4xlarge.search	100 MiB
oi2.8 xlarge.search	100 MiB
oi2.12xlarge.search	100 MiB
oi2.16xlarge.search	100 MiB

Quote di dimensione condivise

La sezione seguente elenca le dimensioni massime degli shard per varie famiglie di istanze.

Tipo di istanza	Multi-AZ senza Standby	Multi-AZ con Standby
R5, C5, M5, C7i, M7i, R7i	N/D	65 GiB
I3, i4i, i4g, i8g, i7i	N/D	65 GiB

Tipo di istanza	Multi-AZ senza Standby	Multi-AZ con Standby
R6g, C6g, M6g, R6gd, C7g,M7g, R7g, R7gd, C8g, M8g, R8g, R8gd	N/D	65 GiB
OR1, OR2, OM2, OI2	N/D	N/D
Im4gn, i8ge	N/D	65 GiB

Per richiedere un aumento della quota, contatta il AWS Supporto.

Quote di conteggio condivise

La sezione seguente elenca il numero massimo di frammenti per le versioni. OpenSearch

Versione del motore	Limite	Note
Elasticsearch da 1.5 a 6.x	Nessun limite predefinito	
Elasticsearch 7.x	1000	Il limite predefinito può essere modificato tramite cluster. Impostazione max_shard s_per_node.
OpenSearch da 1.x a 2.15	1000	Il limite predefinito può essere modificato tramite cluster. Impostazione max_shard s_per_node.
OpenSearch 2.17 e versioni successive	1000 per ogni 16 GB di heap fino a un massimo di 4000	Il limite predefinito non può essere modificato.

Quota dei processi Java

OpenSearch Il servizio limita la memoria dei processi Java al 50% della memoria totale disponibile (massimo fino a 32 GiB). Il limite massimo di 32 GB non si applica a r7g e alle istanze ottimizzate. OpenSearch

Per richiedere l'aumento del limite, contatta AWS il Supporto.

Gli utenti esperti possono anche specificare la percentuale dell'heap utilizzata per i dati sul campo. Per ulteriori informazioni, vedere Impostazioni avanzate del cluster e OutOfMemoryErrorJVM.

Quota della policy di dominio

OpenSearch Il servizio limita le politiche di accesso sui domini a 100 KiB.

Istanze riservate in Amazon Service OpenSearch

Le istanze riservate (RIs) di Amazon OpenSearch Service offrono sconti significativi rispetto alle istanze on demand standard. Le istanze stesse sono identiche; si RIs tratta solo di uno sconto di fatturazione applicato alle istanze On-Demand del tuo account. Per applicazioni di lunga durata con utilizzo prevedibile, RIs può offrire notevoli risparmi nel tempo.

OpenSearch Il servizio RIs richiede una durata di uno o tre anni e prevede tre opzioni di pagamento che influiscono sul tasso di sconto:

- Nessun pagamento anticipato: non è previsto alcun pagamento anticipato. Paghi una tariffa oraria scontata per ogni ora entro il termine.
- Pagamento anticipato parziale: si paga una parte dei costi in anticipo e una tariffa oraria scontata per ogni ora entro il termine.
- Pagamento anticipato totale: di paga l'intero costo in anticipo. Non paghi una tariffa oraria per il termine.

In linea generale, un pagamento anticipato maggiore implica uno sconto maggiore. Non è possibile annullare le istanze riservate: quando vengono prenotate, ci si impegna a pagare per l'intero termine. I pagamenti anticipati non sono rimborsabili.

RIs non sono flessibili; si applicano solo al tipo esatto di istanza prenotato. Ad esempio, una prenotazione per otto istanze `c5.2xlarge.search` non si applica a sedici istanze

`c5.xlarge.search` o quattro istanze `c5.4xlarge.search`. Tuttavia, gli account collegati che fanno parte di un'organizzazione AWS Organizations possono beneficiare di qualsiasi richiesta di sconto non utilizzata dell'account proprietario del RI, purché i tipi di istanza, la regione, la famiglia e le dimensioni corrispondano. Per ulteriori informazioni, consulta i prezzi e le domande frequenti di Amazon OpenSearch Service.

Acquisto di istanze riservate (console)

La console ti consente di visualizzare le istanze riservate esistenti e acquistarne di nuove.

Per acquistare una prenotazione

1. Vai sul sito https://aws.amazon.com e scegli Sign In to the Console (Accedi alla console).
2. In Analytics, scegli Amazon OpenSearch Service.
3. Scegli Locazioni di istanze riservate dal pannello di navigazione.

   In questa pagina, è possibile visualizzare le prenotazioni esistenti. Se si dispone di numerose prenotazioni, è possibile filtrarle per identificarle più facilmente e visualizzare una determinata prenotazione.

   > **Tip**
   >
   > Se il collegamento Locazioni di istanze riservate non è visibile, crea un dominio nella Regione AWS.

4. Scegli Ordina istanza riservata.
5. Fornisci un nome descrittivo univoco.
6. Scegli un tipo di istanza e il numero di istanze. Per le linee guida, consulta the section called “Dimensionamento dei domini”.
7. Scegliere la durata del termine e l'opzione di pagamento. Esaminare attentamente le informazioni di pagamento.
8. Seleziona Next (Successivo).
9. Rivedere attentamente il riepilogo dell'acquisto. Le istanze riservate acquistate non sono rimborsabili.
10. Scegliere Ordina.

Acquisto di istanze riservate (AWS CLI)

AWS CLI Dispone di comandi per visualizzare le offerte, acquistare una prenotazione e visualizzare le prenotazioni. Il comando e la risposta di esempio seguenti mostrano le offerte relative a un determinato elemento: Regione AWS

```
aws opensearch describe-reserved-instance-offerings --region us-east-1
{
  "ReservedInstanceOfferings": [
    {
      "FixedPrice": x,
      "ReservedInstanceOfferingId": "1a2a3a4a5-1a2a-3a4a-5a6a-1a2a3a4a5a6a",
      "RecurringCharges": [
        {
          "RecurringChargeAmount": y,
          "RecurringChargeFrequency": "Hourly"
        }
      ],
      "UsagePrice": 0.0,
      "PaymentOption": "PARTIAL_UPFRONT",
      "Duration": 31536000,
      "InstanceType": "m4.2xlarge.search",
      "CurrencyCode": "USD"
    }
  ]
}
```

Per la spiegazione di ogni valore restituito, consultare la tabella riportata di seguito.

Campo	Description
`FixedPrice`	Il costo anticipato della prenotazione.
`ReservedInstanceOfferingId`	ID dell'offerta. Annotare questo valore se si desidera prenotare l'offerta.
`RecurringCharges`	La tariffa oraria per la prenotazione.
`UsagePrice`	Un campo legacy. Per OpenSearch Service, questo valore è sempre 0.

Campo	Description
`PaymentOption`	Nessun pagamento anticipato, pagamento anticipato parziale o pagamento anticipato totale
`Duration`	Durata del termine in secondi: • 31536000 secondi corrispondono a 1 anno. • 94608000 secondi corrispondono a tre anni.
`InstanceType`	Il tipo di istanza per la prenotazione. Per informazioni sulle risorse hardware allocate a ciascun tipo di istanza, consulta i prezzi di Amazon OpenSearch Service.
`CurrencyCode`	La valuta per `FixedPrice` e `RecurringChargeAmount` .

Nell'esempio successivo viene acquistata una prenotazione:

```
aws opensearch purchase-reserved-instance-offering --reserved-instance-offering-
id 1a2a3a4a5-1a2a-3a4a-5a6a-1a2a3a4a5a6a --reservation-name my-reservation --instance-
count 3 --region us-east-1
{
  "ReservationName": "my-reservation",
  "ReservedInstanceId": "9a8a7a6a-5a4a-3a2a-1a0a-9a8a7a6a5a4a"
}
```

Infine, è possibile elencare le prenotazioni per una determinata regione utilizzando l'esempio seguente:

```
aws opensearch describe-reserved-instances --region us-east-1
{
  "ReservedInstances": [
    {
      "FixedPrice": x,
      "ReservedInstanceOfferingId": "1a2a3a4a5-1a2a-3a4a-5a6a-1a2a3a4a5a6a",
      "ReservationName": "my-reservation",
```

```
      "PaymentOption": "PARTIAL_UPFRONT",
      "UsagePrice": 0.0,
      "ReservedInstanceId": "9a8a7a6a-5a4a-3a2a-1a0a-9a8a7a6a5a4a",
      "RecurringCharges": [
        {
          "RecurringChargeAmount": y,
          "RecurringChargeFrequency": "Hourly"
        }
      ],
      "State": "payment-pending",
      "StartTime": 1522872571.229,
      "InstanceCount": 3,
      "Duration": 31536000,
      "InstanceType": "m4.2xlarge.search",
      "CurrencyCode": "USD"
    }
  ]
}
```

> **Note**
>
> `StartTime` è l'Unix epoch, che è il numero di secondi trascorsi dalla mezzanotte UTC del 1° gennaio 1970. Ad esempio, l'epoch 1522872571 sono le 20:09:31 UTC del 4 aprile 2018. È possibile utilizzare convertitori online.

Per ulteriori informazioni sui comandi utilizzati negli esempi precedenti, consultare Riferimento ai comandi AWS CLI.

Acquisto di istanze riservate ()AWS SDKs

AWS SDKs (eccetto Android e iOS SDKs) supportano tutte le operazioni definite nell'Amazon OpenSearch Service API Reference, incluse le seguenti:

- `DescribeReservedInstanceOfferings`
- `PurchaseReservedInstanceOffering`
- `DescribeReservedInstances`

Questo script di esempio utilizza il client Python OpenSearchServicedi basso livello per AWS SDK per Python (Boto3) l'acquisto di istanze riservate. È necessario fornire un valore per `instance_type`:

```
import boto3
from botocore.config import Config

# Build the client using the default credential configuration.
# You can use the CLI and run 'aws configure' to set access key, secret
# key, and default region.

my_config = Config(
    # Optionally lets you specify a region other than your default.
    region_name='us-east-1'
)

client = boto3.client('opensearch', config=my_config)

instance_type = '' # e.g. m4.2xlarge.search


def describe_RI_offerings(client):
    """Gets the Reserved Instance offerings for this account"""

    response = client.describe_reserved_instance_offerings()
    offerings = (response['ReservedInstanceOfferings'])
    return offerings


def check_instance(offering):
    """Returns True if instance type is the one you specified above"""

    if offering['InstanceType'] == instance_type:
        return True

    return False


def get_instance_id():
    """Iterates through the available offerings to find the ID of the one you
 specified"""

    instance_type_iterator = filter(
```

```
        check_instance, describe_RI_offerings(client))
    offering = list(instance_type_iterator)
    id = offering[0]['ReservedInstanceOfferingId']
    return id


def purchase_RI_offering(client):
    """Purchase Reserved Instances"""

    response = client.purchase_reserved_instance_offering(
        ReservedInstanceOfferingId = get_instance_id(),
        ReservationName = 'my-reservation',
        InstanceCount = 1
    )
    print('Purchased reserved instance offering of type ' + instance_type)
    print(response)


def main():
    """Purchase Reserved Instances"""
    purchase_RI_offering(client)
```

Per ulteriori informazioni sull'installazione e l'utilizzo di AWS SDKs, vedete AWS Software Development Kit.

Analisi dei costi

Cost Explorer è uno strumento gratuito che è possibile utilizzare per visualizzare i dati di spesa per gli ultimi 13 mesi. L'analisi di questi dati consente di identificare le tendenze e capire se sono RIs adatte al caso d'uso. Se lo avete già fatto RIs, potete raggrupparli per opzione di acquisto e mostrare i costi ammortizzati per confrontare tale spesa con la spesa per le istanze on demand. È anche possibile impostare i budget di utilizzo per verificare di usare appieno le istanze riservate. Per ulteriori informazioni, consultare Analisi dei costi con Cost Explorer nella Guida per l'utente di AWS Billing .

Tecniche di ottimizzazione dei costi per Amazon OpenSearch Service

Di seguito sono riportate alcune delle tecniche più comunemente utilizzate per ottimizzare i costi durante l'utilizzo di Amazon OpenSearch Service, sia Managed Clusters che Serverless. Poiché

ogni carico di lavoro è unico, valuta queste strategie rispetto ai tuoi modelli di utilizzo specifici e convalidale in un ambiente di test prima di applicarle alla produzione.

Argomenti

- Ottimizzazione dei costi per i cluster gestiti da Amazon OpenSearch Service
- Ottimizzazione dei costi per Amazon OpenSearch Service Serverless

Ottimizzazione dei costi per i cluster gestiti da Amazon OpenSearch Service

Fonte derivata: ignora la memorizzazione del campo _source

Derived Source è una funzionalità di ottimizzazione dello storage che elimina il sovraccarico legato alla memorizzazione del campo: `_source`

- OpenSearch memorizza ogni documento inserito due volte: una volta nel `_source` campo (documento non elaborato) e una volta come campi indicizzati per la ricerca.
- Il `_source` campo da solo può occupare una notevole quantità di spazio di archiviazione, spesso il 30-50% dello spazio di archiviazione totale dell'indice.
- Con Derived Source, si evita l'archiviazione `_source` e la si ricostruisce invece dinamicamente da campi indicizzati quando necessario (durante le operazioni di ricerca, get, mget, reindex o aggiornamento).
- Si tratta di un opt-in, abilitato al momento della creazione dell'indice utilizzando le impostazioni dell'indice composito.
- Disponibile in tutte le regioni in cui è supportata la versione OpenSearch 3.1 o successiva.

Ideale per: carichi di lavoro analitici e di registrazione in cui non è necessario recuperare frequentemente il documento non elaborato originale ma è comunque necessario cercare e aggregare più campi.

Per ulteriori informazioni, consulta la documentazione open source e l'annuncio What's New.

OR1 / OR2 / OM2 instances: famiglie di istanze OpenSearch ottimizzate

OR1 e le OM2 istanze più recenti OR2 utilizzano Amazon S3 per lo storage di repliche tramite la replica di segmenti:

- OR2: Throughput di indicizzazione fino al 26% superiore rispetto al 70% in più rispetto alle istanze R7g. OR1
- OM2: Throughput di indicizzazione fino al 15% superiore rispetto alle istanze M7g, il 66% in più. OR1
- Entrambe utilizzano la stessa architettura: EBS locale per lo storage principale e S3 per la durabilità.
- Elimina i costi di storage delle repliche: repliche archiviate in S3 (durabilità 11 nove secondi) anziché in costosi volumi EBS.
- Miglioramento del rapporto prezzo/prestazioni fino al 30% rispetto alle istanze della generazione precedente.
- Supporta Shallow Snapshot v2, ossia istantanee quasi istantanee senza sovraccarico. I/O

Ideale per: carichi di lavoro di analisi operativa e analisi dei log che richiedono un'elevata indicizzazione.

Per ulteriori informazioni, consulta l'argomento relativo all'annuncio OR2 e alle OM2 novità e alla guida per gli the section called “Istanze ottimizzate per OpenSearch ” sviluppatori.

Indici cumulativi: aggregano i dati storici delle serie temporali

Gli indici cumulativi riepilogano e comprimono i dati delle serie temporali precedenti in intervalli di tempo più approssimativi, riducendo drasticamente il volume di archiviazione:

- Dati IoT/sensori: conserva i dati al secondo in hot storage per periodi recenti; aggrega i riepiloghi orari o giornalieri per i dati più vecchi.
- Metriche di sistema: conserva le metriche dettagliate degli ultimi 30 giorni; aggrega i dati più vecchi in riepiloghi orari o giornalieri.
- Dati di registro: conserva tutti i dettagli per la finestra di risoluzione dei problemi attiva (ad esempio, 1 settimana); mantieni i modelli di errore riepilogati per i periodi precedenti.
- Combinalo con le politiche ISM per automatizzare il rollup e la migrazione a più livelli in un'unica politica del ciclo di vita.
- Maggiori risparmi nell'aggregazione da secondi a ore rispetto a secondi a minuti.

Per ulteriori informazioni, consulta il post sul blog Index Rollups.

Index State Management: automatizza l'intero ciclo di vita dei dati

Le policy ISM automatizzano lo spostamento degli indici attraverso livelli di storage e azioni del ciclo di vita:

- Migrazione automatica degli indici: da Hot to Cold to Cold to Delete, in base UltraWarm all'età, alle dimensioni o al numero di documenti.
- Attiva i rollup prima delle transizioni di livello per ridurre il volume di dati.
- Imposta politiche di rollover (ad esempio, quando un indice raggiunge i 50 GB o ha 30 giorni) per controllare la crescita dell'indice.
- Automatizza l'unione forzata sugli indici di sola lettura per recuperare spazio di archiviazione dai documenti eliminati.
- Combinalo con i rollup per il massimo risparmio su set di dati di grandi serie temporali.

Istanze riservate: assicurati sconti prevedibili

Per carichi di lavoro analitici stabili e prevedibili, le istanze riservate offrono sconti significativi rispetto ai prezzi on demand:

- Termini di impegno di 1 o 3 anni senza opzioni di pagamento anticipato, anticipato parziale o completamente anticipato.
- Ideale per nodi di dati di livello avanzato e nodi master dedicati che funzionano continuamente.
- Usa il calcolatore AWS dei prezzi per stimare i risparmi prima di impegnarti.
- Le istanze riservate sono uno sconto di fatturazione applicato alle istanze on demand: non sono necessarie modifiche all'infrastruttura.

Tipi di istanze e numero corretti

Linee guida chiave tratte da OpenSearch Well-Architected Lens e best practice per il corretto dimensionamento:

- Utilizza sempre istanze di ultima generazione (ad esempio, le istanze Graviton3 offrono prestazioni migliori fino al 25% rispetto alle istanze basate su Graviton2).
- Utilizza i volumi EBS gp3 anziché gp2: prestazioni migliori a costi inferiori senza costi aggiuntivi.
- Abbina il tipo di istanza al carico di lavoro: ottimizzata per la memoria per le ricerche intensive, ottimizzata per il calcolo per le attività di indicizzazione più impegnative.

- Valuta i nodi dedicati alla gestione dei cluster: necessari solo per 3 o più nodi di dati; evita di sovradimensionare i nodi master con dimensioni eccessive.
- Monitora i CloudWatch parametri per rilevare l'over-provisioning: CPU sostenuta inferiore al 40%, JVM inferiore al 50% e storage inferiore al 50% sono segnali di spreco.
- Intervalli ottimali: CPU 60-80%, JVM 65-85%, storage 70-85% per carichi di lavoro sostenuti.

Per ulteriori informazioni, consulta il post sul blog Right-Sizing Best Practices.

Ottimizzazione degli shard: evita lo sharding eccessivo

L'over-sharding è un fattore di costo nascosto: troppi shard di piccole dimensioni sprecano CPU, memoria e heap JVM:

- Dimensioni degli shard consigliate: 10-50 GiB per shard a seconda del carico di lavoro.
- Non più di 25 shard per GiB di heap Java, non più di 1.000 shard per nodo di dati.
- Utilizza le politiche di rollover di ISM per controllare la crescita degli indici ed evitare la proliferazione illimitata degli shard.
- Riduci il numero di repliche laddove la durabilità lo consente (OR1/OR2 le istanze eliminano completamente la necessità di repliche).
- Utilizza l'unione forzata sugli indici di sola lettura per ridurre il numero di segmenti e recuperare spazio di archiviazione.

Per ulteriori informazioni, consulta Quanti frammenti mi servono?

Zero-ETL/Interrogazione diretta con Amazon S3

Per i dati che vengono interrogati molto raramente ma che devono rimanere accessibili, Direct Query (zero-ETL con S3) consente di interrogare i dati S3 direttamente senza importarli: OpenSearch

- Nessun costo di inserimento: i dati rimangono in S3.
- Nessun costo di storage di livello superiore per i dati di archiviazione.
- Pay-per-query modello di calcolo.
- Supporta OpenSearch dashboard per la visualizzazione.
- La latenza di secondi o minuti è accettabile, non per i casi d'uso in tempo reale.

Campionamento e compressione al momento dell'ingestione

Riduci i costi prima ancora che i dati raggiungano: OpenSearch

- Campionamento: inserisci solo un sottoinsieme rappresentativo di flussi di log ad alto volume (ad esempio, il 10% dei log di debug).
- Compressione dell'indice: abilita il codec di compressione migliore per ridurre l'ingombro di archiviazione.
- Filtraggio dei campi: elimina i campi ad alta cardinalità e con basso valore prima dell'indicizzazione (ad esempio, tracce di stack non elaborate per i vecchi log).
- Politiche di conservazione: definisci finestre di conservazione massime in linea con i requisiti di conformità: non archiviare mai i dati a tempo indeterminato.

Evita i costi di assistenza prolungati: rimani aggiornato sulle versioni del motore

Amazon OpenSearch Service addebita una tariffa fissa per l'ora di istanza normalizzata per le versioni del motore in Extended Support:

- Il mantenimento di versioni precedenti e non supportate comporta costi aggiuntivi oltre ai costi delle istanze.
- Esegui l'upgrade alle versioni attualmente supportate per evitare i costi di Extended Support.

Tag di allocazione dei costi e monitoraggio CloudWatch

La governance proattiva dei costi previene gli sprechi:

- Applica i tag di allocazione dei costi ai OpenSearch domini per un monitoraggio dettagliato dei costi per team o carico di lavoro.
- Imposta CloudWatch allarmi per l'utilizzo dello storage, la pressione della JVM e la CPU per rilevare tempestivamente l'over-provisioning.
- Usa AWS Cost Explorer per identificare i domini con un utilizzo costantemente basso.
- Evaluate Auto-Tune: regola automaticamente le dimensioni dell'heap JVM e altre impostazioni per migliorare le prestazioni e ridurre lo spreco di risorse.

Ottimizzazione dei costi per Amazon OpenSearch Service Serverless

Ricerca vettoriale ottimizzata per il disco (raccolte vettoriali)

La ricerca vettoriale ottimizzata per disco è una delle tecniche di riduzione dei costi più potenti per i carichi di lavoro vettoriali. Esegue la ricerca vettoriale a una frazione del costo della modalità in memoria, mantenendo solo i vettori compressi nella RAM e archiviando vettori ad alta precisione su disco.

Come funziona:

- In modalità standard (`in_memory`), l'intero grafico HNSW viene caricato nella RAM, il che diventa proibitivamente costoso su larga scala.
- In `on_disk` modalità, solo i vettori compressi (quantizzati) vengono conservati in memoria per la generazione candidata; i vettori a piena precisione vengono recuperati dal disco solo per la fase di rescoring finale (ricerca a due fasi).
- Ciò riduce drasticamente il fabbisogno di RAM mantenendo al contempo un'elevata qualità di ricerca.
- `on_disk`La modalità predefinita utilizza una quantizzazione binaria 32x, riducendo i requisiti di memoria del 97% rispetto alla modalità in memoria.
- Supporta i livelli di compressione: 2x (FP16), 4x (byte), 8x, 16x, 32x (binario).
- Latenza P90 nell'intervallo 100-200 ms, adatta per carichi di lavoro che non richiedono tempi di risposta a una cifra di millisecondi.

Benchmark per il risparmio sui costi:

Set di dati	Richiama @100	Latenza P90	Riduzione dei costi
Coerenza TREC-RAG	0,94	104 ms	83%
Tasb-1M	0,96	7 ms	67%
Marco-1M	0,99	7 ms	67%

Ideale per: pipeline RAG, ricerca semantica, recupero di documenti e qualsiasi carico di lavoro vettoriale in cui una latenza P90 di 100-200 ms è accettabile e la riduzione dei costi è una priorità.

> **Note**
>
> Per applicare questa modifica ai dati indicizzati esistenti, è necessario reindicizzare. È possibile utilizzare uno strumento di pipeline esterno, ad esempio per reindicizzare i dati in un nuovo indice.

Per ulteriori informazioni, consultate il post del blog Disk-Optimized Vector Search, il post sul blog Quantization Techniques e. the section called “Lavorare con le raccolte di ricerca vettoriale”

Vector Auto-Optimize (raccolte vettoriali)

L'ottimizzazione automatica valuta automaticamente le configurazioni degli indici vettoriali e consiglia il miglior compromesso tra qualità di ricerca, latenza e costo della memoria, senza richiedere competenze vettoriali:

- Fornisce consigli di ottimizzazione in meno di un'ora.
- Integrato con pipeline di ingestione vettoriale.
- Può essere combinato con l'indicizzazione accelerata da GPU per database vettoriali su scala miliardaria.

Per ulteriori informazioni, consulta il post sul blog Auto-Optimize.

Indicizzazione vettoriale accelerata da GPU (raccolte vettoriali)

L'accelerazione GPU trasferisce l'indicizzazione vettoriale HNSW alla modalità serverless, riducendo drasticamente i tempi e i costi dell'OCU legati alla creazione di indici vettoriali di grandi dimensioni: GPUs

- Tempi di creazione degli indici da 6,4 a 13,8 volte più rapidi rispetto all'indicizzazione basata solo su CPU.
- Costi OCU di indicizzazione inferiori fino al 75% per carichi di lavoro vettoriali con elevati livelli di scrittura.
- GPUs sono collegati dinamicamente: paghi solo quando l'accelerazione GPU è attiva. OCUs
- Consente di creare database vettoriali su scala miliardaria in meno di un'ora.
- Caricato separatamente come accelerazione vettoriale. OCUs

Ideale per: ingestione vettoriale iniziale su larga scala o frequenti scenari di riqualificazione dei modelli in cui la ricostruzione degli indici è costosa.

Per ulteriori informazioni, consulta il post sul blog GPU Acceleration.

Imposta i limiti massimi dell'OCU (tutti i tipi di raccolta)

Amazon OpenSearch Service Serverless si ridimensiona automaticamente in OCUs base alla richiesta. Senza un limite, i costi possono aumentare inaspettatamente. Imposta un limite massimo di OCU a livello di account o per gruppo di raccolta per evitare una scalabilità incontrollata. Il sistema è scalabile fino a questo limite durante i picchi di carico, ma non lo supererà.

Valori consentiti:

- Livello di account: qualsiasi valore fino a 1.700 OCUs (non limitato a multipli di 16).
- Gruppi di raccolta: 1, 2, 4, 8, 16 e multipli di 16 fino a 1.696. OCUs

Monitora le CloudWatch metriche (`OCUUtilization`) per ridimensionare correttamente le impostazioni massime dell'OCU nel tempo.

> **Note**
>
> Se l'utilizzo raggiunge il limite massimo dell'OCU, le prestazioni potrebbero peggiorare notevolmente anche se i costi sono contenuti. Il raggiungimento del limite massimo non risolve la causa alla base del picco dell'OCU. Per le raccolte vettoriali, la causa principale sono in genere i vettori in memoria, che devono essere affrontati direttamente ottimizzando l'indicizzazione vettoriale, riducendo la dimensione dell'indice o regolando i compromessi tra richiamo e precisione.

> **Tip**
>
> Iniziate con una OCU massima conservativa e aumentatela solo quando mostrate un utilizzo prolungato vicino al limite massimo. CloudWatch Se raggiungi costantemente il limite massimo, analizza la causa principale, in particolare l'utilizzo dei vettori in memoria per le raccolte vettoriali, anziché limitarti a innalzare il limite.

Per ulteriori informazioni, consulta the section called “Gestione dei limiti di capacità”.

Ottimizza il periodo di conservazione (raccolte di serie temporali)

Le policy relative al ciclo di vita dei dati eliminano automaticamente i dati dalle raccolte di serie temporali dopo un periodo di conservazione specificato, impedendo una crescita illimitata dello storage. Solo le raccolte di serie temporali supportano le politiche relative al ciclo di vita dei dati, mentre la ricerca e le raccolte vettoriali no.

Il conteggio OCU per le raccolte di serie temporali dipende direttamente dalla quantità di dati recenti che devono essere conservati nell'archiviazione locale. Le raccolte di serie temporali conservano solo la parte più recente dei dati nell'archiviazione OCU locale; i dati più vecchi vengono trasferiti su S3 e il numero di scale di conseguenza: OCUs

OCUs required = max (minimo OCUs, OCUs necessario per conservare i dati entro la finestra di conservazione)

Configurazione delle politiche del ciclo di vita dei dati:

- Imposta periodi di conservazione da 24 ore a 3.650 giorni per indice o modello di indice.
- Amazon OpenSearch Service Serverless elimina automaticamente i dati nel miglior modo possibile (in genere entro 48 ore o il 10% del periodo di conservazione).
- Le regole possono essere applicate a livello di raccolta, a livello di modello di indice o a livello di indice individuale: le regole più specifiche hanno la precedenza.

Esempio di dimensionamento:

- 1 TiB/day ingestione con una conservazione di 30 giorni = circa 1 TiB di hot data = 20 OCUs per l'indicizzazione + 20 per la ricerca. OCUs
- Riduzione a una conservazione di 7 giorni = circa 233 GiB di hot data = circa OCUs 4 per l'indicizzazione + 4 per la ricerca. OCUs

Una conservazione più breve significa meno dati «hot» nello storage locale, meno dati OCUs necessari e una minore spesa di elaborazione. Allinea i periodi di conservazione agli effettivi requisiti aziendali e di conformità: per impostazione predefinita, non conservare i dati a tempo indeterminato.

Per ulteriori informazioni, consulta the section called “Utilizzo di politiche relative al ciclo di vita dei dati”.

Evita di archiviare dati non necessari (tutti i tipi di raccolta)

La riduzione diretta del volume di dati acquisiti riduce sia i costi di elaborazione (OCUs) che quelli di archiviazione:

- Filtra i campi al momento dell'inserimento: utilizza le pipeline per eliminare i campi di basso valore prima che raggiungano la raccolta.
- Evita di importare dati duplicati o ridondanti: deduplica a livello di pipeline.
- Usa le mappature degli indici appropriate: disabilita l'indicizzazione sui campi archiviati ma mai ricercati (). `"index": false`
- Per le raccolte di ricerca: evitate di archiviare blob binari di grandi dimensioni o testo non elaborato che gonfia lo spazio di archiviazione senza valore di ricerca.

Gruppi di raccolta per carichi di lavoro multi-tenant (tutti i tipi di raccolta)

I gruppi di raccolta consentono raccolte multiple con chiavi KMS diverse per condividere le risorse OCU all'interno dello stesso limite di sicurezza, riducendo drasticamente i costi per le architetture multi-tenant. Applicabile ai clienti che utilizzano più chiavi KMS per tenant o per raccolta:

- In precedenza, ogni chiave KMS unica richiedeva una chiave KMS dedicata, il che rendeva OCUs proibitivo l'isolamento per tenant.
- Con i gruppi di raccolta, i tenant con chiavi di crittografia separate possono condividere la capacità dell'OCU.
- Risparmi sui costi fino al 90% per un gran numero di carichi di lavoro degli inquilini più piccoli.
- Supporta sia i valori minimi OCUs (base garantita, nessun avvio a freddo) che quelli massimi OCUs (limite di costo).
- Le raccolte con diversi tipi di accesso alla rete (pubblico e VPC) possono coesistere nello stesso gruppo.
- CloudWatch le metriche forniscono visibilità per gruppo sul consumo di risorse e sulla latenza.

Ideale per: provider SaaS, piattaforme multi-tenant o qualsiasi carico di lavoro con molte piccole raccolte che richiedono ciascuna la propria chiave KMS.

Per ulteriori informazioni, consulta il post sul blog Collection Groups, l'annuncio What's New e. the section called “Gruppi di raccolta”

Altre risorse supportate in Amazon OpenSearch Service

Questo argomento descrive le risorse aggiuntive supportate da Amazon OpenSearch Service.

bootstrap.memory_lock

OpenSearch Il servizio `bootstrap.memory_lock` abilita l'accesso`opensearch.yml`, che blocca la memoria JVM e impedisce al sistema operativo di scambiarla su disco. Ciò vale per tutti i tipi di istanze supportati, tranne che per le seguenti:

- `t2.micro.search`
- `t2.small.search`
- `t2.medium.search`
- `t3.small.search`
- `t3.medium.search`

Modulo di scripting

OpenSearch Il servizio supporta lo scripting per Elasticsearch 5. domini x e successivi. Il servizio non supporta lo scripting per le versioni 1.5 o 2.3.

Le opzioni di scripting supportate includono le seguenti:

- Painless
- Lucene Expressions
- Mustache

Per i domini Elasticsearch 5.5 e versioni successive e per tutti i OpenSearch domini, OpenSearch Service supporta gli script archiviati utilizzando l'endpoint. `_scripts` I domini Elasticsearch 5.3 e 5.1 supportano solo gli script in linea.

Trasporto TLS

OpenSearch Il servizio supporta HTTP sulla porta 80 e HTTPS sulla porta 443, ma non supporta il trasporto TLS.

Tutorial Amazon OpenSearch Service

Questo capitolo include diversi start-to-finish tutorial per lavorare con Amazon OpenSearch Service, tra cui come migrare al servizio, creare una semplice applicazione di ricerca e creare una visualizzazione nelle dashboard. OpenSearch

Argomenti

- Tutorial: creazione e ricerca di documenti in Amazon OpenSearch Service
- Tutorial: migrazione ad Amazon Service OpenSearch
- Tutorial: creazione di un'applicazione di ricerca con Amazon OpenSearch Service
- Tutorial: visualizzazione delle chiamate all'assistenza clienti con OpenSearch Service e dashboard OpenSearch

Tutorial: creazione e ricerca di documenti in Amazon OpenSearch Service

In questo tutorial, imparerai come creare e cercare un documento in Amazon OpenSearch Service. Aggiungi dati a un indice sotto forma di documento JSON. OpenSearch Il servizio crea un indice attorno al primo documento aggiunto.

Questo tutorial spiega come effettuare richieste HTTP per creare documenti, generare automaticamente un ID per un documento ed eseguire ricerche di base e avanzate sui documenti.

> **Note**
>
> Questo tutorial utilizza un dominio con accesso aperto. Per il massimo livello di sicurezza, ti consigliamo di inserire il dominio all'interno di un cloud privato virtuale (VPC).

Prerequisiti

Di seguito sono elencati i requisiti per questo tutorial:

- Devi avere un Account AWS.
- È necessario disporre di un dominio OpenSearch di servizio attivo.

Aggiunta di un documento a un indice

Per aggiungere un documento a un indice, puoi utilizzare qualsiasi strumento HTTP, come Postman, CURL o OpenSearch la console Dashboards. Questi esempi presuppongono che tu stia utilizzando la console per sviluppatori in Dashboards. OpenSearch Se utilizzi uno strumento diverso, modificalo di conseguenza fornendo l'URL completo e le credenziali, se necessario.

Per aggiungere un documento a un indice

1. Vai all'URL delle OpenSearch dashboard per il tuo dominio. Puoi trovare l'URL nella dashboard del dominio nella console di OpenSearch servizio. L'URL segue il seguente formato:

   ```
   domain-endpoint/_dashboards/
   ```

2. Accedi utilizzando il nome utente e la password principali.
3. Apri il pannello di navigazione a sinistra e scegli Strumenti di sviluppo.
4. Il verbo HTTP per creare una nuova risorsa è PUT, e si utilizza per creare un nuovo documento e indice. Immettere il seguente comando nella console:

   ```
   PUT fruit/_doc/1
   {
     "name":"strawberry",
     "color":"red"
   }
   ```

   La richiesta PUT crea un indice denominato frutta e aggiunge un singolo documento all'indice con un ID pari a 1. Viene generata la risposta seguente:

   ```
   {
     "_index" : "fruit",
     "_type" : "_doc",
     "_id" : "1",
     "_version" : 1,
     "result" : "created",
     "_shards" : {
       "total" : 2,
       "successful" : 2,
       "failed" : 0
     },
     "_seq_no" : 0,
   ```

```
  "_primary_term" : 1
}
```

Creazione generata automaticamente IDs

OpenSearch Il servizio può generare automaticamente un ID per i tuoi documenti. Il comando da generare IDs utilizza una richiesta POST anziché una richiesta PUT e non richiede alcun ID di documento (rispetto alla richiesta precedente).

Inserisci la seguente richiesta nella console degli sviluppatori:

```
POST veggies/_doc
{
  "name":"beet",
  "color":"red",
  "classification":"root"
}
```

Questa richiesta crea un indice denominato verdure e aggiunge il documento all'indice. Viene generata la risposta seguente:

```
{
  "_index" : "veggies",
  "_type" : "_doc",
  "_id" : "3WgyS4IB5DLqbRIvLxtF",
  "_version" : 1,
  "result" : "created",
  "_shards" : {
    "total" : 2,
    "successful" : 2,
    "failed" : 0
  },
  "_seq_no" : 0,
  "_primary_term" : 1
}
```

Nota quel `_id` campo aggiuntivo nella risposta, che indica che un ID è stato creato automaticamente.

> **Note**
>
> Non fornisci nulla dopo `_doc` nell'URL, dove normalmente va l'ID. Poiché stai creando un documento con un ID generato, non ne fornisci ancora uno. È riservato agli aggiornamenti.

Aggiornamento di un documento con un comando POST

Per aggiornare un documento, si utilizza un comando `POST` HTTP con il numero di ID.

In primo luogo, creare un documento con un ID di `42`:

```
POST fruits/_doc/42
{
  "name":"banana",
  "color":"yellow"
}
```

Quindi usa quell'ID per aggiornare il documento:

```
POST fruits/_doc/42
{
  "name":"banana",
  "color":"yellow",
  "classification":"berries"
}
```

Questo comando aggiorna il documento con il nuovo campo `classification`. Viene generata la risposta seguente:

```
{
  "_index" : "fruits",
  "_type" : "_doc",
  "_id" : "42",
  "_version" : 2,
  "result" : "updated",
  "_shards" : {
    "total" : 2,
    "successful" : 2,
    "failed" : 0
  },
```

```
  "_seq_no" : 1,
  "_primary_term" : 1
}
```

> **Note**
>
> Se si tenta di aggiornare un documento che non esiste, OpenSearch Service crea il documento.

Esecuzione di operazioni in blocco

Puoi utilizzare il l'operazione API `POST _bulk` per eseguire più azioni su uno o più indici in una richiesta. I comandi di azione in blocco assumono il formato seguente:

```
POST /_bulk
<action_meta>\n
<action_data>\n
<action_meta>\n
<action_data>\n
```

Ogni azione richiede due righe di JSON. Innanzitutto, fornisci la descrizione o i metadati dell'azione. Nella riga successiva, fornisci i dati. Ogni parte è separata da una nuova riga (\n). Una descrizione dell'azione per un inserto potrebbe essere simile alla seguente:

```
{ "create" : { "_index" : "veggies", "_type" : "_doc", "_id" : "7" } }
```

E la riga successiva che contiene i dati potrebbe essere simile alla seguente:

```
{ "name":"kale", "color":"green", "classification":"leafy-green" }
```

Nel loro insieme, i metadati e i dati rappresentano una singola azione in un'operazione di blocco. È possibile eseguire molte operazioni in un'unica richiesta, ad esempio:

```
POST /_bulk
{ "create" : { "_index" : "veggies", "_id" : "35" } }
{ "name":"kale", "color":"green", "classification":"leafy-green" }
{ "create" : { "_index" : "veggies", "_id" : "36" } }
{ "name":"spinach", "color":"green", "classification":"leafy-green" }
```

```
{ "create" : { "_index" : "veggies", "_id" : "37" } }
{ "name":"arugula", "color":"green", "classification":"leafy-green" }
{ "create" : { "_index" : "veggies", "_id" : "38" } }
{ "name":"endive", "color":"green", "classification":"leafy-green" }
{ "create" : { "_index" : "veggies", "_id" : "39" } }
{ "name":"lettuce", "color":"green", "classification":"leafy-green" }
{ "delete" : { "_index" : "vegetables", "_id" : "1" } }
```

Si noti che l'ultima azione è `delete`. Non ci sono dati che seguono l'azione `delete`.

Ricerca di documenti

Ora che i dati sono presenti nel cluster, è possibile cercarli. Ad esempio, potresti voler cercare tutte le verdure con la radice o ottenere il numero di tutte le verdure a foglia verde o trovare il numero di errori registrati all'ora.

Ricerche base

Una ricerca di base è simile a questa:

```
GET veggies/_search?q=name:l*
```

La richiesta genera una risposta JSON che contiene il documento sulla lattuga.

Ricerca avanzata

È possibile eseguire ricerche più avanzate fornendo le opzioni di query come JSON nel corpo della richiesta:

```
GET veggies/_search
{
  "query": {
    "term": {
      "name": "lettuce"
    }
  }
}
```

Questo esempio produce anche una risposta JSON con il documento sulla lattuga.

Ordinamento

È possibile eseguire più query di questo tipo utilizzando l'ordinamento. Innanzitutto, è necessario ricreare l'indice, poiché la mappatura automatica dei campi ha scelto tipi che non possono essere ordinati per impostazione predefinita. Inviare le richieste seguenti per eliminare e ricreare l'indice:

```
DELETE /veggies

PUT /veggies
{
   "mappings":{
      "properties":{
         "name":{
            "type":"keyword"
         },
         "color":{
            "type":"keyword"
         },
         "classification":{
            "type":"keyword"
         }
      }
   }
}
```

Quindi ripopolare l'indice con i dati:

```
POST /_bulk
{ "create" : { "_index" : "veggies", "_id" : "7"  } }
{ "name":"kale", "color":"green", "classification":"leafy-green" }
{ "create" : { "_index" : "veggies", "_id" : "8" } }
{ "name":"spinach", "color":"green", "classification":"leafy-green" }
{ "create" : { "_index" : "veggies", "_id" : "9" } }
{ "name":"arugula", "color":"green", "classification":"leafy-green" }
{ "create" : { "_index" : "veggies", "_id" : "10" } }
{ "name":"endive", "color":"green", "classification":"leafy-green" }
{ "create" : { "_index" : "veggies", "_id" : "11" } }
{ "name":"lettuce", "color":"green", "classification":"leafy-green" }
```

Ora puoi cercare con un ordinamento. Questa richiesta aggiunge un ordinamento crescente in base alla classificazione:

```
GET veggies/_search
{
```

```
  "query" : {
    "term": { "color": "green" }
  },
  "sort" : [
      "classification"
  ]
}
```

Risorse correlate

Per maggiori informazioni, consulta le seguenti risorse:

- Nozioni di base
- Indicizzazione dei dati
- Ricerca di dati

Tutorial: migrazione ad Amazon Service OpenSearch

Gli snapshot degli indici sono un modo popolare per migrare da un cluster Elasticsearch legacy OpenSearch o autogestito ad Amazon Service. OpenSearch In generale, il processo consiste nei seguenti passaggi:

1. Acquisire uno snapshot del cluster esistente e caricare lo snapshot in un bucket Amazon S3.
2. Crea un dominio di servizio. OpenSearch
3. Concedi al OpenSearch Servizio le autorizzazioni per accedere al bucket e assicurati di disporre delle autorizzazioni per lavorare con le istantanee.
4. Ripristina l'istantanea nel dominio del servizio. OpenSearch

Questa procedura dettagliata fornisce passaggi più dettagliati e opzioni alternative, ove applicabile.

Acquisizione e caricamento dello snapshot

Sebbene sia possibile utilizzare il plug-in repository-s3 per scattare istantanee direttamente su S3, è necessario installare il plug-in su ogni nodo, modificare `opensearch.yml` (o `elasticsearch.yml` se si utilizza un cluster Elasticsearch), riavviare ogni nodo, aggiungere le credenziali e infine scattare l'istantanea. AWS Il plugin è una grande opzione per l'uso continuo o per la migrazione di cluster più grandi.

Per i cluster più piccoli, un approccio monouso consiste nello scattare un'istantanea del file system condivisa e quindi utilizzarla per caricarla su S3. AWS CLI Se si dispone già di uno snapshot, andare al passaggio 4.

Per acquisire uno snapshot e caricarlo in Amazon S3

1. Aggiungi l'impostazione `path.repo` a `opensearch.yml` (o `Elasticsearch.yml`) su tutti i nodi e quindi riavvia ogni nodo.

```
path.repo: ["/my/shared/directory/snapshots"]
```

2. Registra un repository di snapshot; questa operazione è necessaria prima di acquisire uno snapshot. Un repository è solo un percorso di storage: un file system condiviso, Amazon S3, File system distribuito Hadoop (HDFS) ecc. In questo caso, utilizzeremo un file system condiviso ("fs"):

```
PUT _snapshot/my-snapshot-repo-name
{
  "type": "fs",
  "settings": {
    "location": "/my/shared/directory/snapshots"
  }
}
```

3. Acquisire lo snapshot:

```
PUT _snapshot/my-snapshot-repo-name/my-snapshot-name
{
  "indices": "migration-index1,migration-index2,other-indices-*",
  "include_global_state": false
}
```

4. Installare AWS CLI ed eseguire `aws configure` per aggiungere le credenziali.
5. Passare alla directory snapshot. Quindi eseguire i seguenti comandi per creare un nuovo bucket S3 e caricare il contenuto della directory snapshot in quel bucket:

```
aws s3 mb s3://amzn-s3-demo-bucket --region us-west-2
aws s3 sync . s3://amzn-s3-demo-bucket --sse AES256
```

A seconda delle dimensioni dello snapshot e della velocità della connessione Internet, questa operazione può richiedere un po' di tempo.

Creare un dominio

Sebbene la console sia il modo più semplice per creare un dominio, in questo caso il terminale è già aperto e installato. AWS CLI Modificare il seguente comando per creare un dominio che si adatti alle proprie esigenze:

```
aws opensearch create-domain \
  --domain-name migration-domain \
  --engine-version OpenSearch_1.0 \
  --cluster-config InstanceType=c5.large.search,InstanceCount=2 \
  --ebs-options EBSEnabled=true,VolumeType=gp2,VolumeSize=100 \
  --node-to-node-encryption-options Enabled=true \
  --encryption-at-rest-options Enabled=true \
  --domain-endpoint-options EnforceHTTPS=true,TLSSecurityPolicy=Policy-Min-
TLS-1-2-2019-07 \
  --advanced-security-options
 Enabled=true,InternalUserDatabaseEnabled=true,MasterUserOptions='{MasterUserName=master-
user,MasterUserPassword=master-user-password}' \
  --access-policies '{"Version": "2012-10-17",       "Statement":
[{"Effect":"Allow","Principal": {"AWS": "arn:aws:iam::aws-
region:user/UserName"},"Action":["es:ESHttp*"],"Resource":"arn:aws:es:aws-
region:111122223333:domain/migration-domain/*"}]}' \
  --region aws-region
```

Così com'è, il comando crea un dominio accessibile a Internet con due nodi di dati, ciascuno con 100 GiB di archiviazione. Consente inoltre il controllo granulare degli accessi con l'autenticazione di base HTTP e tutte le impostazioni di crittografia. Usa la console OpenSearch di servizio se hai bisogno di una configurazione di sicurezza più avanzata, come un VPC.

Prima di eseguire il comando, modificare il nome di dominio, le credenziali dell'utente master e il numero di account. Specificate la stessa Regione AWS che avete usato per il bucket S3 e una OpenSearch/Elasticsearch versione compatibile con la vostra istantanea.

> **Important**
>
> Gli snapshot sono compatibili con le versioni successive e solo con una versione principale. Ad esempio, non è possibile ripristinare un'istantanea da un file 1. OpenSearch cluster x

su un Elasticsearch 7. cluster x, solo un OpenSearch 1. x o 2. cluster x. Anche la versione minore conta. Non è possibile ripristinare un'istantanea da un cluster 5.3.3 autogestito su un dominio di servizio OpenSearch 5.3.2. Ti consigliamo di scegliere la versione più recente di Elasticsearch supportata dalla OpenSearch tua istantanea. Per una tabella delle versioni compatibili, consultare the section called “Utilizzo di uno snapshot per migrare i dati”.

Fornire le autorizzazioni al bucket S3.

Nella console AWS Identity and Access Management (IAM), crea un ruolo con le seguenti autorizzazioni e relazioni di fiducia. Durante la creazione del ruolo, scegliere S3 come Servizio AWS . Assegnare un nome al ruolo `OpenSearchSnapshotRole` in modo che sia facile da trovare.

Autorizzazioni

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
      "Action": [
        "s3:ListBucket"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket"
      ]
    },
    {
      "Action": [
        "s3:GetObject",
        "s3:PutObject",
        "s3:DeleteObject"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket/*"
      ]
    }
  ]
```

```
}
```

Relazione di attendibilità

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
      "Effect": "Allow",
      "Principal": {
        "Service": "es.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

Quindi assegna al tuo ruolo IAM personale le autorizzazioni per assumere `OpenSearchSnapshotRole`. Creare la policy seguente e collegarla alla propria identità.

Autorizzazioni

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
      "Effect": "Allow",
      "Action": "iam:PassRole",
      "Resource": "arn:aws:iam::123456789012:role/OpenSearchSnapshotRole"
    }
  ]
}
```

Mappa il ruolo dello snapshot nelle OpenSearch dashboard (se utilizzi un controllo granulare degli accessi)

Se è stato abilitato il controllo granulare degli accessi, anche se si utilizza l'autenticazione di base HTTP per tutti gli altri scopi, poter utilizzare gli snapshot sarà necessario mappare il ruolo `manage_snapshots` al ruolo IAM.

Come concedere all'identità le autorizzazioni per l'utilizzo degli snapshot

1. Accedi ai dashboard utilizzando le credenziali dell'utente principale che hai specificato quando hai creato il dominio del servizio. OpenSearch Puoi trovare l'URL delle dashboard nella console di servizio. OpenSearch Presenta il formato `https://domain-endpoint/_dashboards/`.
2. Dal menu principale scegliere Sicurezza, Ruoli e selezionare il ruolo manage_snapshots.
3. Scegliere Utenti mappati, Gestisci mappatura.
4. Aggiungere l'ARN del dominio del ruolo IAM personale nel campo appropriato. L'ARN deve avere uno dei seguenti formati:

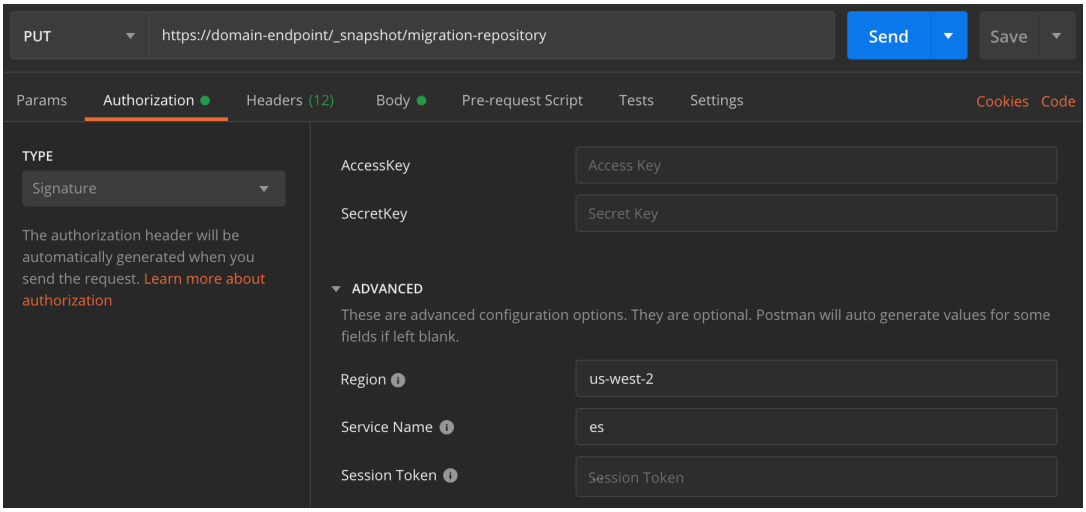
   ```
   arn:aws:iam::123456789123:user/user-name
   ```

   ```
   arn:aws:iam::123456789123:role/role-name
   ```
5. Seleziona Map (Mappa) e verifica che il ruolo sia visualizzato in Mapped users (Utenti mappati).

Ripristino dello snapshot

A questo punto, hai due modi per accedere al tuo dominio di OpenSearch servizio: l'autenticazione di base HTTP con le credenziali dell'utente principale o l' AWS autenticazione tramite le tue credenziali IAM. Poiché le istantanee utilizzano Amazon S3, che non ha il concetto di utente principale, devi utilizzare le tue credenziali IAM per registrare l'archivio di snapshot nel tuo dominio di servizio. OpenSearch

La maggior parte dei linguaggi di programmazione dispone di librerie che facilitano la firma delle richieste, ma l'approccio più semplice consiste nell'utilizzare uno strumento come Postman e inserire le credenziali IAM nella sezione Autorizzazione.



Ripristinare lo snapshot

1. Indipendentemente da come si sceglie di firmare le richieste, il primo passo è registrare il repository:

```
PUT _snapshot/my-snapshot-repo-name
{
  "type": "s3",
  "settings": {
    "bucket": "amzn-s3-demo-bucket",
    "region": "us-west-2",
    "role_arn": "arn:aws:iam::123456789012:role/OpenSearchSnapshotRole"
  }
}
```

2. Quindi elencare lo snapshot nel repository e trovare quello che si desidera ripristinare. A questo punto, continuare a usare Postman o passare a uno strumento come curl.

   Sintassi abbreviata

   ```
   GET _snapshot/my-snapshot-repo-name/_all
   ```

   curl

   ```
   curl -XGET -u 'master-user:master-user-password' https://domain-endpoint/
   _snapshot/my-snapshot-repo-name/_all
   ```

3. Ripristinare lo snapshot:

Sintassi abbreviata

```
POST _snapshot/my-snapshot-repo-name/my-snapshot-name/_restore
{
  "indices": "migration-index1,migration-index2,other-indices-*",
  "include_global_state": false
}
```

curl

```
curl -XPOST -u 'master-user:master-user-password' https://domain-endpoint/
_snapshot/my-snapshot-repo-name/my-snapshot-name/_restore \
  -H 'Content-Type: application/json' \
  -d '{"indices":"migration-index1,migration-index2,other-indices-
*","include_global_state":false}'
```

4. Infine, verifica che gli indici siano ripristinati come previsto.

Sintassi abbreviata

```
GET _cat/indices?v
```

curl

```
curl -XGET -u 'master-user:master-user-password' https://domain-endpoint/_cat/
indices?v
```

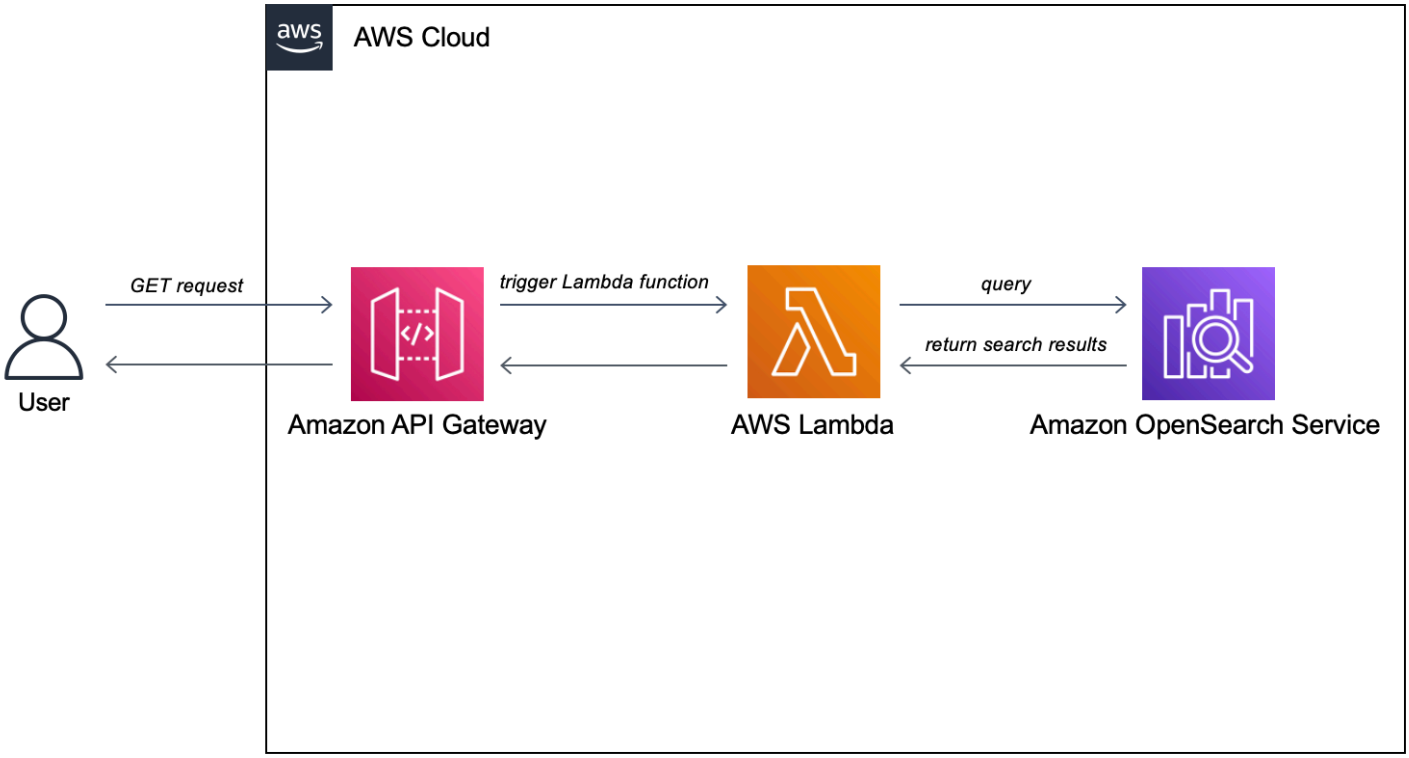
A questo punto, la migrazione è completa. Puoi configurare i tuoi client per utilizzare il nuovo endpoint OpenSearch Service, ridimensionare il dominio in base al tuo carico di lavoro, controllare il numero di shard per i tuoi indici, passare a un utente master IAM o iniziare a creare visualizzazioni nelle dashboard. OpenSearch

Tutorial: creazione di un'applicazione di ricerca con Amazon OpenSearch Service

Un modo comune per creare un'applicazione di ricerca con Amazon OpenSearch Service consiste nell'utilizzare moduli Web per inviare le richieste degli utenti a un server. Quindi puoi autorizzare

il server a OpenSearch APIs chiamarlo direttamente e fare in modo che il server invii richieste al OpenSearch servizio. Se desideri scrivere codice lato client che non si basa su un server, tuttavia, devi compensare per i rischi di sicurezza e prestazioni. Consentire l'accesso pubblico non firmato a OpenSearch APIs è sconsigliabile. Gli utenti possono accedere a endpoint non protetti o influire sulle prestazioni del cluster tramite query eccessivamente estese (o troppe query).

Questo capitolo presenta una soluzione: usa Amazon API Gateway per limitare gli utenti a un sottoinsieme di OpenSearch APIs e AWS Lambda firmare le richieste da API Gateway to OpenSearch Service.



> **Note**
>
> Si applicano i prezzi standard di API Gateway e Lambda, ma entro l'uso limitato di questo tutorial, i costi dovrebbero essere trascurabili.

Prerequisiti

Un prerequisito per questo tutorial è un dominio di OpenSearch servizio. Se non ne hai già uno, segui i passaggi in Creare un dominio di OpenSearch servizio per crearne uno.

Fase 1: Indicizzazione dei dati di esempio

Scaricare sample-movies.zip, decomprimerlo e utilizzare l'operazione dell'API _bulk per aggiungere i 5.000 documenti all'indice `movies`:

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/_bulk
{ "index": { "_index": "movies", "_id": "tt1979320" } }
{"directors":["Ron
 Howard"],"release_date":"2013-09-02T00:00:00Z","rating":8.3,"genres":
["Action","Biography","Drama","Sport"],"image_url":"http://ia.media-imdb.com/images/
M/MV5BMTQyMDE0MTY0OV5BMl5BanBnXkFtZTcwMjI2OTI0OQ@@._V1_SX400_.jpg","plot":"A re-
creation of the merciless 1970s rivalry between Formula One rivals James Hunt and
 Niki Lauda.","title":"Rush","rank":2,"running_time_secs":7380,"actors":["Daniel
 Brühl","Chris Hemsworth","Olivia Wilde"],"year":2013,"id":"tt1979320","type":"add"}
{ "index": { "_index": "movies", "_id": "tt1951264" } }
{"directors":["Francis Lawrence"],"release_date":"2013-11-11T00:00:00Z","genres":
["Action","Adventure","Sci-Fi","Thriller"],"image_url":"http://ia.media-imdb.com/
images/M/
MV5BMTAyMjQ3OTAxMzNeQTJeQWpwZ15BbWU4MDU0NzA1MzAx._V1_SX400_.jpg","plot":"Katniss
 Everdeen and Peeta Mellark become targets of the Capitol after
 their victory in the 74th Hunger Games sparks a rebellion in
 the Districts of Panem.","title":"The Hunger Games: Catching
 Fire","rank":4,"running_time_secs":8760,"actors":["Jennifer Lawrence","Josh
 Hutcherson","Liam Hemsworth"],"year":2013,"id":"tt1951264","type":"add"}
...
```

Nota che quanto sopra è un comando di esempio con un piccolo sottoinsieme dei dati disponibili. Per eseguire l'`_bulk`operazione, è necessario copiare e incollare l'intero contenuto del `sample-movies` file. Per ulteriori istruzioni, vederethe section called “Opzione 2: Caricamento di più documenti”.

Puoi anche usare il seguente comando curl per ottenere lo stesso risultato:

```
curl -XPOST -u 'master-user:master-user-password' 'domain-endpoint/_bulk' --data-binary
 @bulk_movies.json -H 'Content-Type: application/json'
```

Fase 2: Creare e distribuire la funzione Lambda

Prima di creare l'API in API Gateway, crea la funzione Lambda a cui passa le richieste.

Creazione della funzione Lambda

In questa soluzione, API Gateway passa le richieste a una funzione Lambda, che interroga il OpenSearch servizio e restituisce i risultati. Poiché questa funzione di esempio utilizza librerie esterne, è necessario creare un pacchetto di distribuzione e caricarlo su Lambda.

Per creare il pacchetto di implementazione

1. Apri un prompt dei comandi e crea una directory di progetto `my-opensearch-function`. Ad esempio, su macOS:

```
mkdir my-opensearch-function
```

2. Passa alla directory del progetto `my-sourcecode-function`.

```
cd my-opensearch-function
```

3. Copia il contenuto del seguente codice Python di esempio e salvalo in un nuovo file denominato. `opensearch-lambda.py` Aggiungi la tua regione e l'endpoint host al file.

```
import boto3
import json
import requests
from requests_aws4auth import AWS4Auth

region = '' # For example, us-west-1
service = 'es'
credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

host = '' # The OpenSearch domain endpoint with https:// and without a trailing
 slash
index = 'movies'
url = host + '/' + index + '/_search'

# Lambda execution starts here
def lambda_handler(event, context):

    # Put the user query into the query DSL for more accurate search results.
    # Note that certain fields are boosted (^).
    query = {
```

```
        "size": 25,
        "query": {
            "multi_match": {
                "query": event['queryStringParameters']['q'],
                "fields": ["title^4", "plot^2", "actors", "directors"]
            }
        }
    }

    # Elasticsearch 6.x requires an explicit Content-Type header
    headers = { "Content-Type": "application/json" }

    # Make the signed HTTP request
    r = requests.get(url, auth=awsauth, headers=headers, data=json.dumps(query))

    # Create the response and add some extra content to support CORS
    response = {
        "statusCode": 200,
        "headers": {
            "Access-Control-Allow-Origin": '*'
        },
        "isBase64Encoded": False
    }

    # Add the search results to the response
    response['body'] = r.text
    return response
```

4. Installa le librerie esterne in una nuova `package` directory.

```
pip3 install --target ./package boto3
pip3 install --target ./package requests
pip3 install --target ./package requests_aws4auth
```

5. Crea un pacchetto di implementazione con le librerie installate nella directory principale. Il comando seguente genera un `my-deployment-package.zip` file nella directory del progetto.

```
cd package
zip -r ../my-deployment-package.zip .
```

6. Aggiungi il file `opensearch-lambda.py` alla radice del file .zip.

```
cd ..
```

```
zip my-deployment-package.zip opensearch-lambda.py
```

Per ulteriori informazioni sulla creazione di funzioni Lambda e sui pacchetti di implementazione, consultare Implementa funzioni Lambda per Python con gli archivi di file .zip nella Guida per gli sviluppatori di AWS Lambda e the section called “Creazione il pacchetto di implementazione Lambda” in questa guida.

Per creare la tua funzione utilizzando la console Lambda

1. Accedi alla console Lambda a casahttps://console.aws.amazon.com/lambda/. Nel riquadro di navigazione a sinistra, scegli Funzioni.
2. Seleziona Crea funzione.
3. Configura i campi seguenti:

   - Nome della funzione: opensearch-function
   - Runtime: Python 3.9
   - Architettura: x86_64

   Mantieni tutte le altre opzioni predefinite e scegli Crea funzione.
4. Nella sezione Codice sorgente della pagina di riepilogo della funzione, scegli il menu a discesa Carica da e seleziona il file.zip. Individua il **`my-deployment-package.zip`** file che hai creato e scegli Salva.
5. Il gestore è il metodo nel codice della funzione che elabora gli eventi. In Impostazioni di runtime, scegli Modifica e modifica il nome del gestore in base al nome del file nel pacchetto di distribuzione in cui si trova la funzione Lambda. Poiché il file ha un nome`opensearch-lambda.py`, rinomina il gestore in. *`opensearch-lambda`*`.lambda_handler` Per ulteriori informazioni, consulta Gestore della funzione Lambda in Python.

Fase 3: Creare l'API in API Gateway

L'utilizzo di API Gateway consente di creare un'API più limitata e semplifica il processo di interazione con l' OpenSearch `_search`API. API Gateway consente inoltre di abilitare caratteristiche di sicurezza come l'autenticazione di Amazon Cognito e la limitazione delle richieste. Completare la procedura seguente per creare e implementare un'API:

Creazione e configurazione dell'API

Come creare l'API utilizzando la console API Gateway

1. Accedi alla console API Gateway da https://console.aws.amazon.com/apigateway/casa. Nel riquadro di navigazione a sinistra, scegli APIs.
2. Individuare REST API (non privato) e scegliere Crea.
3. Nella pagina seguente, individua la sezione Crea nuova API e assicurati che sia selezionata Nuova API.
4. Configura i campi seguenti:
   - Nome API: opensearch-api
   - Descrizione: API pubblica per la ricerca di un dominio Amazon OpenSearch Service
   - Tipo di endpoint: regionale
5. Seleziona Create API (Crea API).
6. Scegliere Operazioni quindi Crea metodo.
7. Selezionare GET nel menu a discesa e fare clic sul segno di spunta per confermare.
8. Configurare le impostazioni seguenti, quindi scegliere Salva:

Impostazione	Valore
Tipo di integrazione	Funzione Lambda
Utilizzo dellintegrazione proxy Lambda	Sì
Regione Lambda	*us-west-1*
funzione Lambda	`opensearch-lambda`
Utilizzo del timeout di default	Sì

Configurazione della richiesta del metodo

Scegliere Richiesta metodo e configurare le impostazioni seguenti:

Impostazione	Valore
Autorizzazione	NONE
Convalidatore di richieste	Convalida dei parametri e delle intestazioni delle stringhe di query
Chiave API richiesta	false

In Parametri della stringa di query URL, scegli Aggiungi stringa di query e configura il seguente parametro:

Impostazione	Valore
Nome	q
Richiesto	Sì

Implementazione dell'API e configurazione di una fase

La console API Gateway consente di distribuire un'API creando una distribuzione e associandola a una fase nuova o esistente.

1. Scegliere Operazioni e Implementa API.
2. Per Fase di implementazione, scegliere Nuova fase e nominare la fase `opensearch-api-test`.
3. Seleziona Implementa.
4. Configurare le seguenti impostazioni nell'editor della fase, quindi scegliere Salva modifiche:

Impostazione	Valore
Abilitazione della limitazione	Sì
Tariffa	1000
Burst	500

Queste impostazioni configurano un'API che dispone di un solo metodo: una richiesta `GET` alla root dell'endpoint (`https://some-id.execute-api.us-west-1.amazonaws.com/search-es-api-test`). La richiesta richiede un singolo parametro (`q`), la stringa di query da ricercare. Quando viene chiamato, il metodo passa la richiesta a Lambda, che esegue la funzione `opensearch-lambda`. Per ulteriori informazioni, consultare Creazione di un'API in Amazon API Gateway e Implementazione di una REST API in Amazon API Gateway.

Fase 4: (facoltativa) Modifica della policy di accesso al dominio

Il dominio del OpenSearch servizio deve consentire alla funzione Lambda di effettuare `GET` richieste all'`movies`indice. Se il dominio ha una policy di accesso aperto con il controllo granulare degli accessi abilitato, è possibile lasciarlo così com'è:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
      "Action": "es:*",
      "Resource": "arn:aws:es:us-west-1:123456789012:domain/domain-name/*"
    }
  ]
}
```

In alternativa, è possibile scegliere di rendere la policy di accesso al dominio più granulare. Per esempio, la policy minima seguente consente a `opensearch-lambda-role` (creato tramite Lambda) l'accesso in lettura all'indice `movies`: Per ottenere il nome esatto del ruolo creato automaticamente da Lambda, vai alla console AWS Identity and Access Management (IAM), scegli Ruoli e cerca «lambda».

JSON

```
{
```

```
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::123456789012:role/service-role/opensearch-lambda-
role-1abcdefg"
      },
      "Action": "es:ESHttpGet",
      "Resource": "arn:aws:es:us-west-1:123456789012:domain/domain-name/movies/
_search"
    }
  ]
}
```

> **Important**
>
> Se hai abilitato il controllo granulare degli accessi per il dominio, devi anche mappare il ruolo a un utente nelle OpenSearch dashboard, altrimenti vedrai errori di autorizzazione.

Configurare le autorizzazioni del ruolo di esecuzione Lambda

Oltre a configurare la policy di accesso al dominio, devi anche assicurarti che il ruolo di esecuzione Lambda disponga delle autorizzazioni IAM necessarie per accedere OpenSearch al tuo dominio di servizio. La funzione Lambda richiede autorizzazioni specifiche a seconda che si stia utilizzando un dominio gestito o una raccolta OpenSearch Service Serverless.

Per i domini di servizio gestiti OpenSearch :

Allega la seguente policy IAM al tuo ruolo di esecuzione Lambda per consentirgli di effettuare richieste al tuo dominio di OpenSearch servizio:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
```

```
      "Action": [
        "es:ESHttpGet",
        "es:ESHttpPost"
      ],
      "Resource": "arn:aws:es:us-west-1:123456789012:domain/domain-name/*"
    }
  ]
}
```

Per le raccolte OpenSearch Service Serverless:

Se utilizzi OpenSearch Service Serverless, collega la seguente policy IAM al tuo ruolo di esecuzione Lambda:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "aoss:*",
      "Resource": "arn:aws:aoss:us-west-1:123456789012:collection/collection-id"
    }
  ]
}
```

Per allegare queste policy al tuo ruolo di esecuzione Lambda:

1. Accedi alla console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Scegli Ruoli e cerca il tuo ruolo di esecuzione Lambda (chiamato `opensearch-lambda-role-`*`xxxxxxxx`* in genere).
3. Scegli Aggiungi autorizzazioni, quindi Crea politica in linea.
4. Scegli la scheda JSON e incolla la politica appropriata dall'alto, sostituendo i valori segnaposto con la tua risorsa effettiva. ARNs
5. Scegli Review policy, fornisci un nome simile e scegli **`OpenSearchAccess`** Crea policy.

> **Note**
>
> Senza queste autorizzazioni IAM, la funzione Lambda riceverà l'errore «Accesso negato» quando tenta di interrogare OpenSearch il dominio del servizio, anche se la policy di accesso al dominio consente le richieste.

Per ulteriori informazioni sulle policy di accesso a , consulta the section called “Configurazione delle policy di accesso”.

Mappatura del ruolo Lambda (se si utilizza il controllo granulare degli accessi)

Il controllo granulare degli accessi introduce un passaggio aggiuntivo prima che sia possibile testare l'applicazione. Anche se si utilizza l'autenticazione di base HTTP per tutti gli altri scopi, è necessario mappare il ruolo Lambda all'utente, altrimenti si riceveranno errori di autorizzazione.

1. Vai all'URL delle OpenSearch dashboard per il dominio.
2. Dal menu principale, scegli Sicurezza, Ruoli e seleziona il link al `all_access` ruolo a cui devi mappare il ruolo Lambda.
3. Scegliere Utenti mappati, Gestisci mappatura.
4. In Backend roles (Ruoli di backend), aggiungi il nome della risorsa Amazon (ARN) del ruolo Lambda. L'ARN dovrebbe assumere la forma di. `arn:aws:iam::123456789123:role/service-role/opensearch-lambda-role-1abcdefg`
5. Selezionare Mappa e confermare che l'utente o il ruolo venga visualizzato in Utenti mappati.

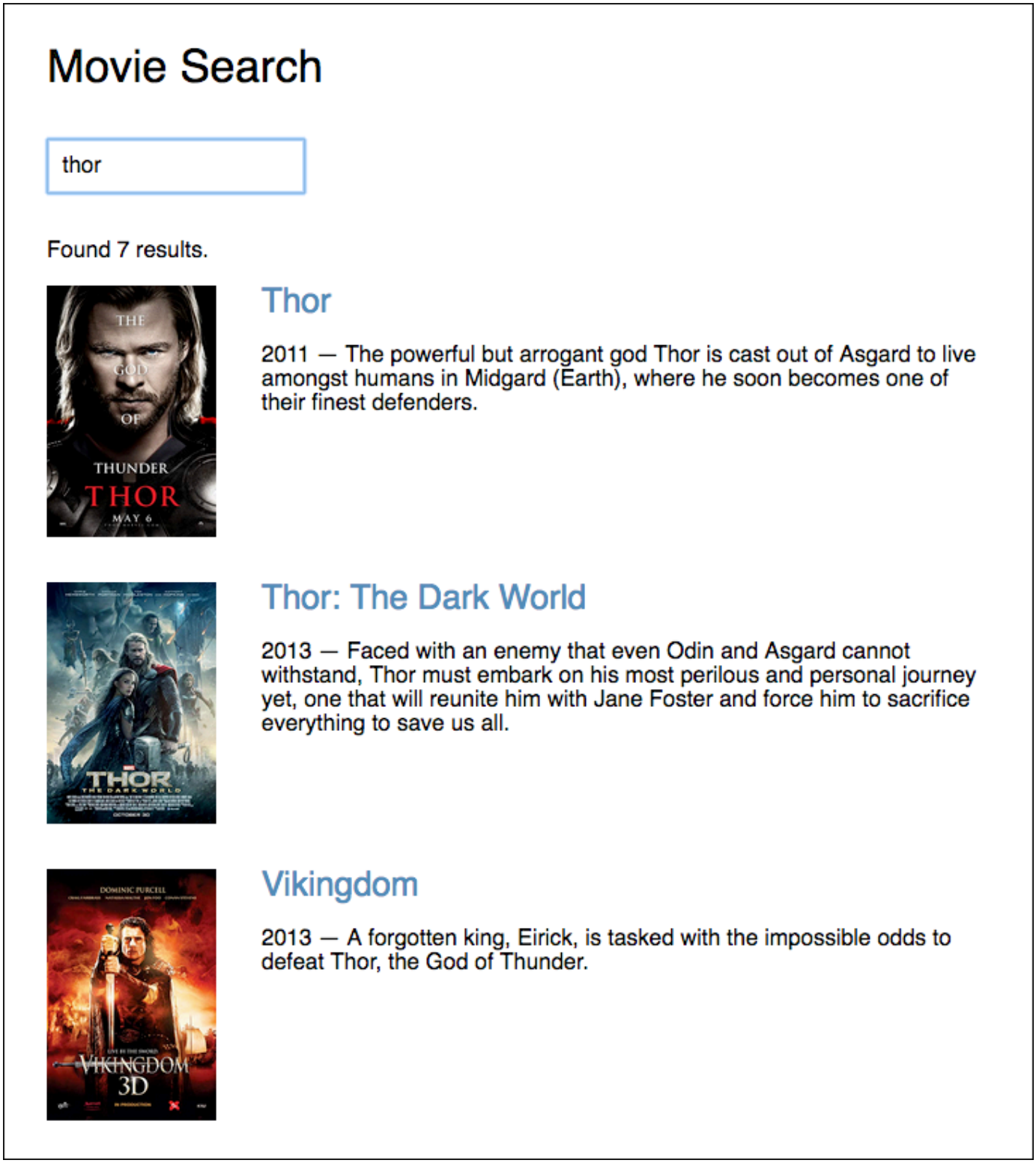
Fase 5: Test dell'applicazione Web

Per testare l'applicazione Web

1. Scaricare sample-site.zip, decomprimerlo e aprire `scripts/search.js` nell'editor di testo preferito.
2. Aggiorna la `apigatewayendpoint` variabile in modo che punti all'endpoint API Gateway e aggiungi una barra rovesciata alla fine del percorso specificato. Puoi trovare rapidamente l'endpoint in API Gateway scegliendo Stages (Fasi) e selezionando il nome dell'API. La

`apigatewayendpoint` variabile deve assumere la forma di /. `https://some-id.execute-api.us-west-1.amazonaws.com/opensearch-api-test`

3. Aprire `index.html` e provare a eseguire la ricerca di thor, house e qualche altro termine.



Risoluzione degli errori CORS

Anche se la funzione Lambda include contenuti nella risposta per il supporto di CORS, è comunque possibile che venga visualizzato il seguente errore:

```
Access to XMLHttpRequest at '<api-gateway-endpoint>' from origin 'null' has been
 blocked by CORS policy: No 'Access-Control-Allow-Origin' header is present in the
 requested resource.
```

In tal caso, prova quanto seguente:

1. Abilita CORS sulla risorsa GET. In Advanced (Avanzati), imposta Access-Control-Allow-Credentials a `'true'`.
2. Ridistribuisci la tua API in API Gateway (Actions (Operazioni),Deploy API (Distribuzione dell'API)).
3. Elimina e aggiungi nuovamente la tua attivazione della funzione Lambda. Aggiungila nuovamente, scegli Aggiungi trigger e crea l'endpoint HTTP che richiama la tua funzione. Il trigger deve avere la seguente configurazione:

Trigger	"Hello, World!"	Fase della distribuzione	Sicurezza
Gateway API	opensearch-api	opensearch-api-test	Aperta

Fasi successive

Questo capitolo è solo un punto di partenza per dimostrare un concetto. Potresti valutare se apportare le seguenti modifiche:

- Aggiungi i tuoi dati al dominio del servizio. OpenSearch
- Aggiungere metodi all'API.
- Nella funzione Lambda, modificare la query di ricerca o potenziare campi diversi.
- Utilizzare uno stile diverso per i risultati o modificare `search.js` per visualizzare campi diversi all'utente.

Tutorial: visualizzazione delle chiamate all'assistenza clienti con OpenSearch Service e dashboard OpenSearch

Questo capitolo è un'analisi completa della seguente situazione: un'azienda riceve un numero di chiamate al servizio di assistenza clienti e desidera analizzarle. Qual è l'argomento di ogni chiamata? Quante sono state positive? Quante sono state negative? In che modo i responsabili possono cercare o analizzare le trascrizioni di queste chiamate?

Un flusso di lavoro manuale potrebbe coinvolgere dipendenti che ascoltano le registrazioni, annotano l'oggetto di ciascuna chiamata e decidono se l'interazione con il cliente è stata positiva.

Tale processo sarebbe estremamente impegnativo. Considerando un tempo medio di 10 minuti per chiamata, ogni dipendente potrebbe ascoltare solo 48 chiamate al giorno. Ad eccezione di eventuali distorsioni umane, i dati generati sarebbero altamente accurati, ma la quantità di dati sarebbe minima: solo l'oggetto della chiamata e un Boolean se il cliente è soddisfatto oppure no. Qualcosa di più complesso, come una trascrizione completa, potrebbe richiedere una grande quantità di tempo.

Utilizzando Amazon S3, Amazon Transcribe, Amazon Comprehend e Amazon Service OpenSearch , puoi automatizzare un processo simile con pochissimo codice e ottenere molti più dati. Ad esempio, è possibile ottenere una trascrizione completa della chiamata, le parole chiave dalla trascrizione e un "sentiment" della chiamata (positivo, negativo, neutro o misto). Quindi puoi utilizzare OpenSearch e OpenSearch Dashboards per cercare e visualizzare i dati.

Sebbene sia possibile utilizzare questa procedura dettagliata così com'è, l'intento è quello di stimolare idee su come arricchire i documenti JSON prima di indicizzarli in Service. OpenSearch

Costi stimati

In generale, l'esecuzione della procedura indicata in questo scenario avrebbe un costo inferiore a $2. La procedura guidata utilizza le risorse seguenti:

- Bucket S3 con meno di 100 MB trasferiti e memorizzati

  Per ulteriori informazioni, consultare Amazon S3 Pricing.
- OpenSearch Dominio di servizio con un'`t2.medium`istanza e 10 GiB di storage EBS per diverse ore

  Per ulteriori informazioni, consulta la pagina dei prezzi OpenSearch di Amazon Service.
- Chiamate multiple ad Amazon Transcribe

Per ulteriori informazioni, consultare Prezzi di Amazon Transcribe.

- Diverse chiamate di elaborazione del linguaggio naturale ad Amazon Comprehend

  Per ulteriori informazioni, consultare Prezzi di Amazon Comprehend.

Argomenti

- Fase 1: Configurazione dei prerequisiti
- Fase 2: Copia del codice di esempio
- (Facoltativo) Fase 3: Indicizzazione dei dati di esempio
- Fase 4: Analisi e visualizzazione dei dati
- Fase 5: Pulizia delle risorse e fasi successive

Fase 1: Configurazione dei prerequisiti

Prima di procedere, devi disporre delle risorse indicate di seguito.

Prerequisito	Description
Bucket Amazon S3	Per ulteriori informazioni, consulta Creazione di un bucket nella Guida per l'utente di Amazon Simple Storage Service.
OpenSearch Dominio del servizio	Destinazione per i dati. Per ulteriori informazioni, consulta Creazione di domini OpenSearch di servizio.

Se non si dispone già di queste risorse, è possibile crearle utilizzando i comandi AWS CLI seguenti:

```
aws s3 mb s3://my-transcribe-test --region us-west-2
```

```
aws opensearch create-domain --domain-name my-transcribe-test --engine-version
 OpenSearch_1.0 --cluster-config  InstanceType=t2.medium.search,InstanceCount=1
 --ebs-options EBSEnabled=true,VolumeType=standard,VolumeSize=10 --access-policies
 '{"Version": "2012-10-17",       "Statement":[{"Effect":"Allow","Principal":
{"AWS":"arn:aws:iam::123456789012:root"},"Action":"es:*","Resource":"arn:aws:es:us-
west-2:123456789012:domain/my-transcribe-test/*"}]}' --region us-west-2
```

> **Note**
>
> Questi comandi utilizzano la regione `us-west-2`, ma è possibile usare qualsiasi regione supportata da Amazon Comprehend. Per ulteriori informazioni consulta Riferimenti generali di AWS.

Fase 2: Copia del codice di esempio

1. Copiare e incollare il seguente codice di esempio Python 3 in un nuovo file denominato `call-center.py`:

```
import boto3
import datetime
import json
import requests
from requests_aws4auth import AWS4Auth
import time
import urllib.request

# Variables to update
audio_file_name = '' # For example, 000001.mp3
bucket_name = '' # For example, my-transcribe-test
domain = '' # For example, https://search-my-transcribe-test-12345.us-
west-2.es.amazonaws.com
index = 'support-calls'
type = '_doc'
region = 'us-west-2'

# Upload audio file to S3.
s3_client = boto3.client('s3')

audio_file = open(audio_file_name, 'rb')

print('Uploading ' + audio_file_name + '...')
response = s3_client.put_object(
    Body=audio_file,
    Bucket=bucket_name,
    Key=audio_file_name
)

# # Build the URL to the audio file on S3.
```

```
# # Only for the us-east-1 region.
# mp3_uri = 'https://' + bucket_name + '.s3.amazonaws.com/' + audio_file_name

# Get the necessary details and build the URL to the audio file on S3.
# For all other regions.
response = s3_client.get_bucket_location(
    Bucket=bucket_name
)
bucket_region = response['LocationConstraint']
mp3_uri = 'https://' + bucket_name + '.s3-' + bucket_region + '.amazonaws.com/' +
 audio_file_name

# Start transcription job.
transcribe_client = boto3.client('transcribe')

print('Starting transcription job...')
response = transcribe_client.start_transcription_job(
    TranscriptionJobName=audio_file_name,
    LanguageCode='en-US',
    MediaFormat='mp3',
    Media={
        'MediaFileUri': mp3_uri
    },
    Settings={
        'ShowSpeakerLabels': True,
        'MaxSpeakerLabels': 2 # assumes two people on a phone call
    }
)

# Wait for the transcription job to finish.
print('Waiting for job to complete...')
while True:
    response =
 transcribe_client.get_transcription_job(TranscriptionJobName=audio_file_name)
    if response['TranscriptionJob']['TranscriptionJobStatus'] in ['COMPLETED',
 'FAILED']:
        break
    else:
        print('Still waiting...')
    time.sleep(10)

transcript_uri = response['TranscriptionJob']['Transcript']['TranscriptFileUri']

# Open the JSON file, read it, and get the transcript.
```

```
response = urllib.request.urlopen(transcript_uri)
raw_json = response.read()
loaded_json = json.loads(raw_json)
transcript = loaded_json['results']['transcripts'][0]['transcript']

# Send transcript to Comprehend for key phrases and sentiment.
comprehend_client = boto3.client('comprehend')

# If necessary, trim the transcript.
# If the transcript is more than 5 KB, the Comprehend calls fail.
if len(transcript) > 5000:
    trimmed_transcript = transcript[:5000]
else:
    trimmed_transcript = transcript

print('Detecting key phrases...')
response = comprehend_client.detect_key_phrases(
    Text=trimmed_transcript,
    LanguageCode='en'
)

keywords = []
for keyword in response['KeyPhrases']:
    keywords.append(keyword['Text'])

print('Detecting sentiment...')
response = comprehend_client.detect_sentiment(
    Text=trimmed_transcript,
    LanguageCode='en'
)

sentiment = response['Sentiment']

# Build the Amazon OpenSearch Service URL.
id = audio_file_name.strip('.mp3')
url = domain + '/' + index + '/' + type + '/' + id

# Create the JSON document.
json_document = {'transcript': transcript, 'keywords': keywords, 'sentiment':
 sentiment, 'timestamp': datetime.datetime.now().isoformat()}

# Provide all details necessary to sign the indexing request.
credentials = boto3.Session().get_credentials()
```

```
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region,
 'opensearchservice', session_token=credentials.token)

# Index the document.
print('Indexing document...')
response = requests.put(url, auth=awsauth, json=json_document, headers=headers)

print(response)
print(response.json())
```

2. Aggiornare le prime sei variabili.
3. Installare i pacchetti necessari utilizzando i comandi seguenti:

```
pip install boto3
pip install requests
pip install requests_aws4auth
```

4. Inserisci il tuo MP3 nella stessa directory `call-center.py` ed esegui lo script. Di seguito è riportato un output di esempio:

```
$ python call-center.py
Uploading 000001.mp3...
Starting transcription job...
Waiting for job to complete...
Still waiting...
Still waiting...
Still waiting...
Still waiting...
Still waiting...
Still waiting...
Still waiting...
Detecting key phrases...
Detecting sentiment...
Indexing document...
<Response [201]>
{u'_type': u'call', u'_seq_no': 0, u'_shards': {u'successful': 1, u'failed': 0,
 u'total': 2}, u'_index': u'support-calls4', u'_version': 1, u'_primary_term': 1,
 u'result': u'created', u'_id': u'000001'}
```

`call-center.py` esegue una serie di operazioni:

1. Lo script carica un file audio (in questo caso, un MP3, ma Amazon Transcribe supporta diversi formati) nel tuo bucket S3.
2. Invia l'URL del file audio ad Amazon Transcribe e attende il completamento del processo di trascrizione.

   Il tempo necessario per completare il processo di trascrizione dipende dalla durata del file audio. Supponiamo siano minuti, non secondi.

   > **Tip**
   >
   > Per migliorare la qualità della trascrizione, è possibile configurare un vocabolario personalizzato per Amazon Transcribe.

3. Una volta completato il processo di trascrizione, lo script estrae la trascrizione, la riduce a 5.000 caratteri e la invia ad Amazon Comprehend per l'analisi del sentiment e delle parole chiave.
4. Infine, lo script aggiunge la trascrizione completa, le parole chiave, il sentimento e il timestamp corrente a un documento JSON e lo indicizza in Service. OpenSearch

> **Tip**
>
> LibriVoxdispone di audiolibri di dominio pubblico che puoi usare per i test.

(Facoltativo) Fase 3: Indicizzazione dei dati di esempio

Se non si dispone di un gruppo di registrazioni di chiamate (chi le ha, effettivamente?) è possibile indicizzare i documenti campione in sample-calls.zip, con un risultato simile a quello prodotto da `call-center.py`.

1. Creare un file denominato `bulk-helper.py`:

```
import boto3
from opensearchpy import OpenSearch, RequestsHttpConnection
import json
from requests_aws4auth import AWS4Auth

host = '' # For example, my-test-domain.us-west-2.es.amazonaws.com
region = '' # For example, us-west-2
```

```
service = 'es'

bulk_file = open('sample-calls.bulk', 'r').read()

credentials = boto3.Session().get_credentials()
awsauth = AWS4Auth(credentials.access_key, credentials.secret_key, region, service,
 session_token=credentials.token)

search = OpenSearch(
    hosts = [{'host': host, 'port': 443}],
    http_auth = awsauth,
    use_ssl = True,
    verify_certs = True,
    connection_class = RequestsHttpConnection
)

response = search.bulk(bulk_file)
print(json.dumps(response, indent=2, sort_keys=True))
```

2. Aggiornare le prime due variabili per `host` e `region`.
3. Installare il pacchetto necessario utilizzando il seguente comando:

```
pip install opensearch-py
```

4. Scaricare e decomprimere sample-calls.zip.
5. Posizionare `sample-calls.bulk` nella stessa directory di `bulk-helper.py` ed eseguire lo script helper. Di seguito è riportato un output di esempio:

```
$ python bulk-helper.py
{
  "errors": false,
  "items": [
    {
      "index": {
        "_id": "1",
        "_index": "support-calls",
        "_primary_term": 1,
        "_seq_no": 42,
        "_shards": {
          "failed": 0,
          "successful": 1,
          "total": 2
```

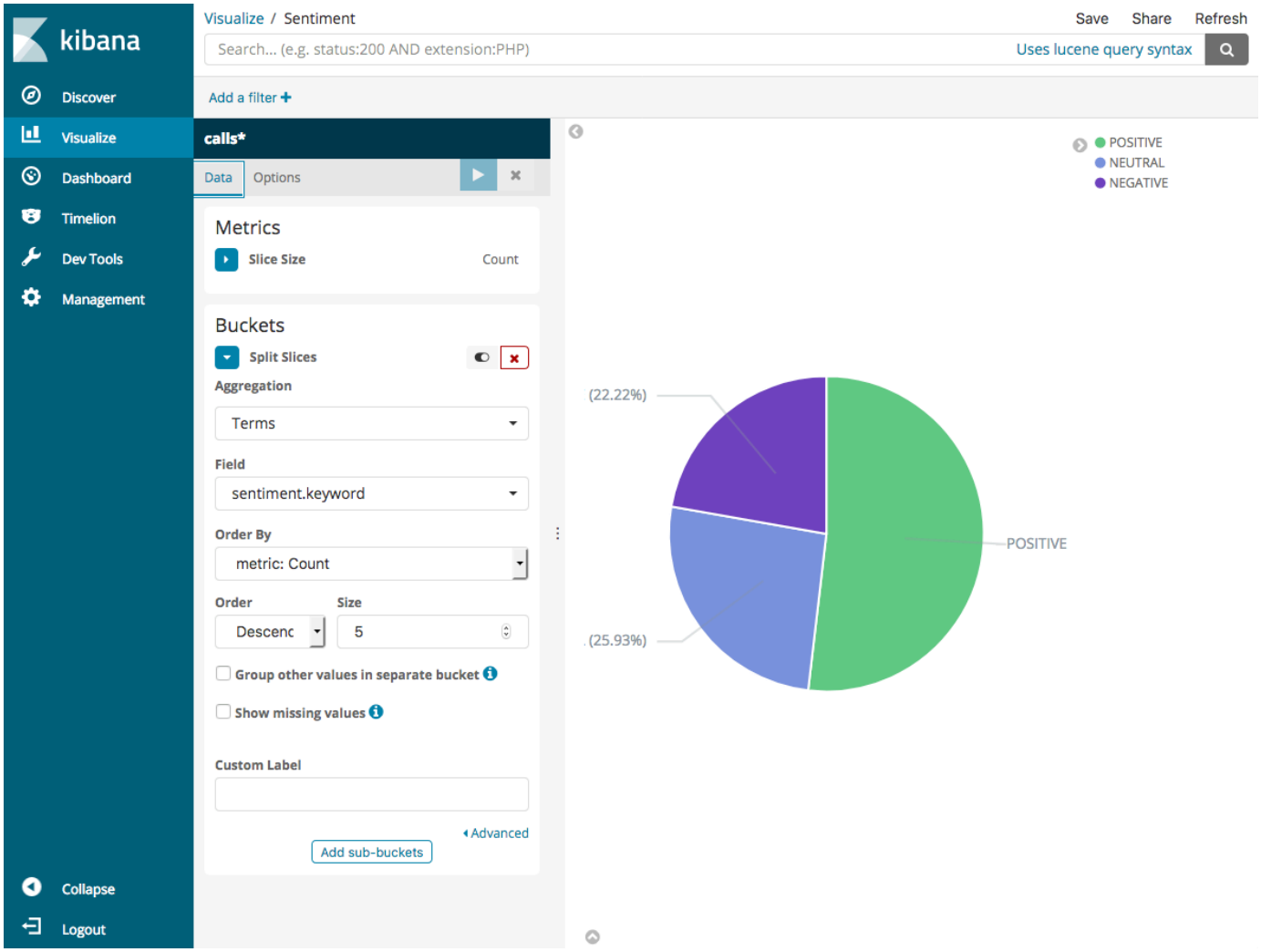
```
        },
        "_type": "_doc",
        "_version": 9,
        "result": "updated",
        "status": 200
      }
    },
    ...
  ],
  "took": 27
}
```

Fase 4: Analisi e visualizzazione dei dati

Ora che hai alcuni dati in OpenSearch Service, puoi visualizzarli utilizzando i dashboard. OpenSearch

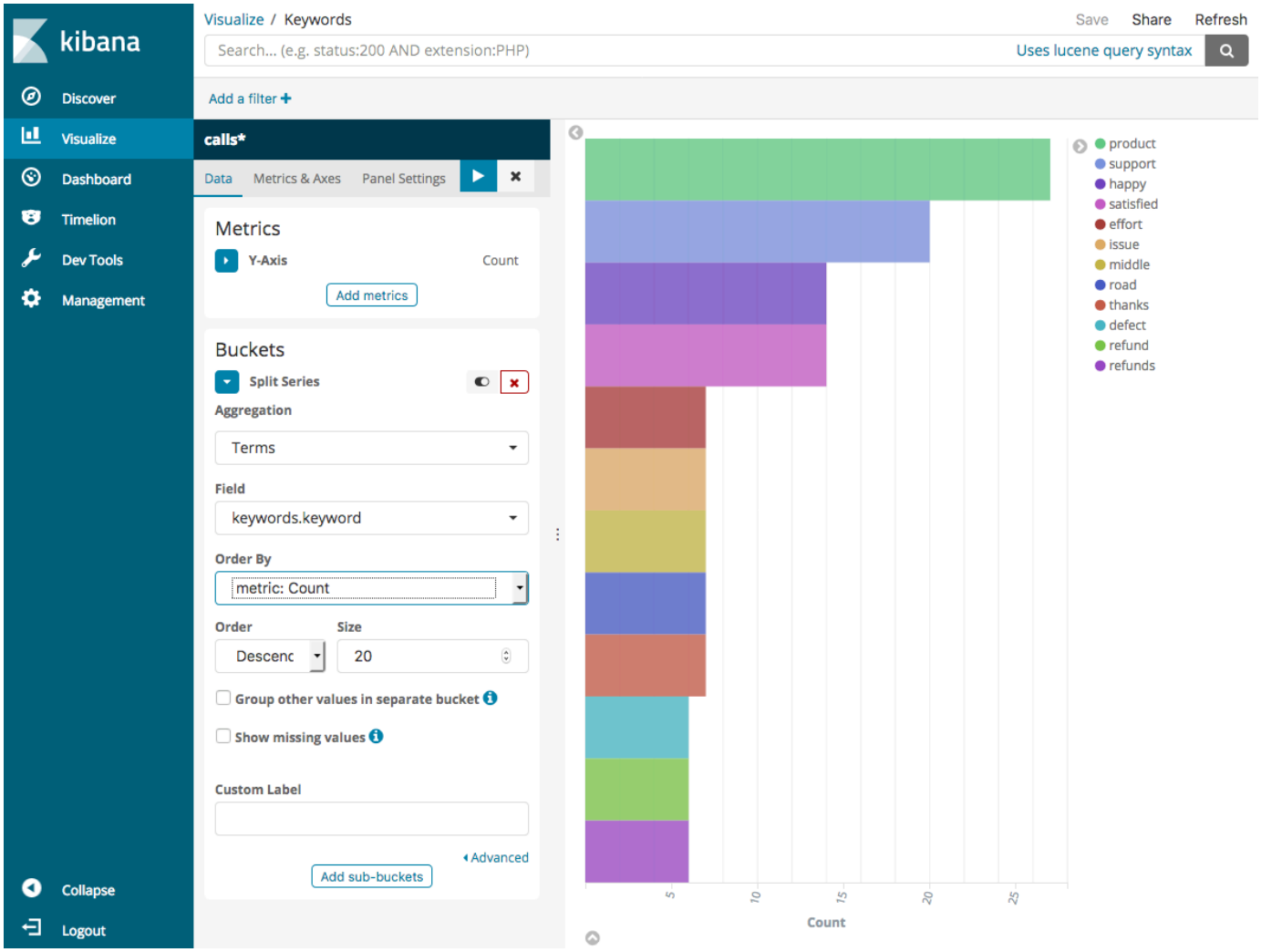
1. Accedi a `https://search-`*`domain`*`.`*`region`*`.es.amazonaws.com/_dashboards`.
2. Prima di poter utilizzare le OpenSearch dashboard, è necessario un modello di indice. Dashboards usa modelli di indice per limitare l'analisi a uno o più indici. Per abbinare l'indice `support-calls` creato da `call-center.py`, passare a Gestione degli stack, Modelli di indice e definire un modello di indice di `support*`, quindi scegliere Approfondimenti.
3. In Time Filter field name (Nome campo Filtro tempo), scegliere timestamp.
4. Ora è possibile iniziare a creare le visualizzazioni. Scegliere Visualize (Visualizza), quindi aggiungere una nuova visualizzazione.
5. Scegliere il grafico a torta e il modello dell'indice `support*`.
6. La visualizzazione predefinita è di base, quindi scegliere Split Slices (Dividi sezioni) per creare una visualizzazione più interessante.

   Per Aggregation (Aggregazione) scegliere Terms (Termini). In Field (Campo), scegliere sentiment.keyword. Quindi selezionare Apply changes (Applica modifiche) e poi Save (Salva).

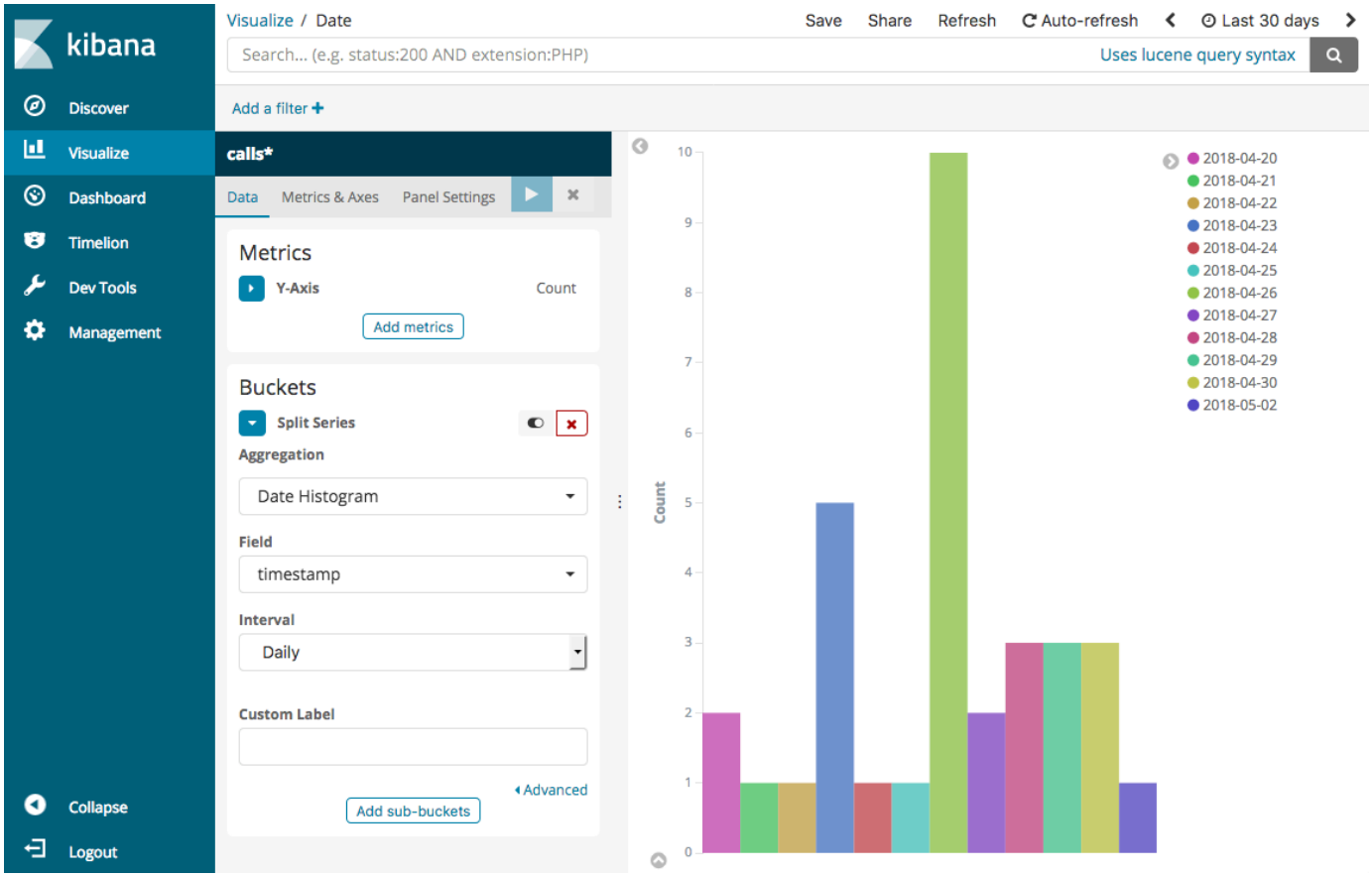


7. Tornare alla pagina Visualize (Visualizza) e aggiungere un'altra visualizzazione. Questa volta, scegliere il grafico a barre orizzontali.
8. Scegliere Split Series (Dividi serie).

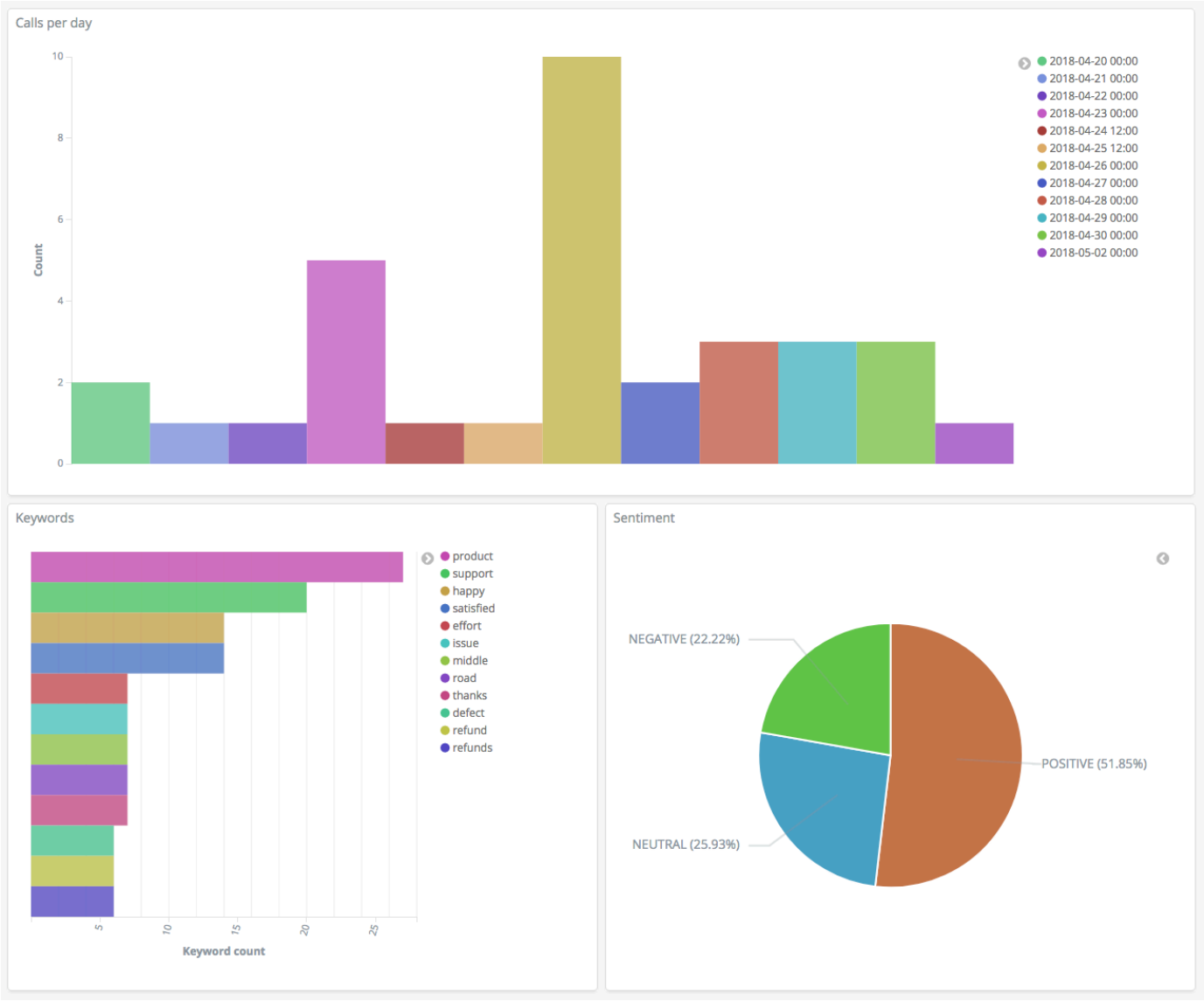
   Per Aggregation (Aggregazione) scegliere Terms (Termini). In Field (Campo), scegliere keywords.keyword e modificare Size (Dimensioni) in 20. Quindi selezionare Apply Changes (Applica modifiche) e poi Save (Salva).



9. Tornare alla pagina Visualize (Visualizza) e aggiungere una visualizzazione finale, un grafico a barre verticali.
10. Scegliere Split Series (Dividi serie). In Aggregation (Aggregazione), scegliere Date Histogram (Istogramma date). In Field (Campo), scegliere timestamp e modificare Interval (Intervallo) in Daily (Giornaliero).
11. Scegliere Metrics & Axes (Parametri e assi) e cambiare Mode (Modalità) in normal (normale).
12. Selezionare Apply Changes (Applica modifiche) e poi Save (Salva).



13. Ora che si dispone di tre visualizzazioni, è possibile aggiungerle a una visualizzazione di Dashboards. Scegliere Dashboard (Pannello di controllo), creare un pannello di controllo e aggiungere le visualizzazioni.



Fase 5: Pulizia delle risorse e fasi successive

Per evitare addebiti inutili, elimina il bucket S3 e OpenSearch il dominio di servizio. Per ulteriori informazioni, consulta Delete a Bucket nella Amazon Simple Storage Service User Guide e Delete an OpenSearch Service domain in questa guida.

Le trascrizioni richiedono molto meno spazio su disco rispetto ai file. MP3 Potresti riuscire ad abbreviare il MP3 periodo di conservazione, ad esempio da tre mesi di registrazioni delle chiamate a un mese, conservare anni di trascrizioni e comunque risparmiare sui costi di archiviazione.

Puoi anche automatizzare il processo di trascrizione utilizzando and AWS Step Functions Lambda, aggiungere metadati aggiuntivi prima dell'indicizzazione o creare visualizzazioni più complesse per adattarle al tuo caso d'uso esatto.

Amazon OpenSearch Service rename - Riepilogo delle modifiche

L'8 settembre 2021, la nostra suite di ricerca e analisi è stata rinominata Amazon OpenSearch Service. OpenSearch Il servizio supporta OpenSearch oltre al precedente sistema operativo Elasticsearch. Nelle sezioni seguenti vengono descritte le diverse parti del servizio modificate con la rinomina del servizio e le operazioni da eseguire per garantire che i domini continuino a funzionare correttamente.

Alcune di queste modifiche si applicano solo quando aggiorni i tuoi domini da Elasticsearch a. OpenSearch In altri casi, ad esempio nella console di gestione fatturazione e costi, l'esperienza cambia immediatamente.

Questo elenco non è completo. Mentre anche altre parti del prodotto sono cambiate, questi aggiornamenti sono i più rilevanti.

Nuova versione dell'API

La nuova versione dell'API di configurazione del OpenSearch servizio (2021-01-01) funziona anche con il precedente sistema operativo Elasticsearch. OpenSearch 21 operazioni API sono state sostituite con nomi più concisi e indipendenti dal motore (ad esempio, modificati in), ma Service continua a supportare entrambe le versioni dell'API. `CreateElasticsearchDomain` `CreateDomain` OpenSearch

Per creare e gestire domini da questo punto in poi, consigliamo di utilizzare le nuove operazioni API. Notare che quando si utilizzano le nuove operazioni API per creare un dominio, è necessario specificare il parametro `EngineVersion` nel formato `Elasticsearch_X.Y` o `OpenSearch_X.Y`, piuttosto che solo il numero di versione. Se non si specifica una versione, per impostazione predefinita viene utilizzata la versione più recente di. OpenSearch

Aggiorna il tuo AWS CLI alla versione 1.20.40 o successiva per utilizzarlo per creare e gestire `aws opensearch ...` i tuoi domini. Per il nuovo formato CLI, consultate il riferimento CLIOpenSearch .

Tipi di istanza rinominati

I tipi di istanza in Amazon OpenSearch Service ora sono nel formato, ad `<type>.<size>.search` esempio, `m6g.large.`*`search`* anzichém`6g.large.`*`elasticsearch`*. Non è necessario eseguire

nessuna operazione. I domini esistenti inizieranno automaticamente con riferimento ai nuovi tipi di istanza all'interno dell'API e nella console di gestione fatturazione e costi.

Se disponi di istanze riservate (RIs), il tuo contratto non sarà influenzato dalla modifica. La versione precedente dell'API di configurazione è ancora compatibile con il vecchio formato di denominazione, ma se si desidera utilizzare la nuova versione dell'API, sarà necessario utilizzare il nuovo formato.

Modifiche delle policy di accesso

Nelle sezioni seguenti vengono descritte le azioni da intraprendere per aggiornare le policy di accesso.

Policy IAM

Consigliamo di aggiornare le policy IAM in modo da utilizzare le operazioni API rinominate. Tuttavia, OpenSearch Service continuerà a rispettare le politiche esistenti replicando internamente le vecchie autorizzazioni API. Ad esempio, se al momento si dispone dell'autorizzazione per eseguire l'operazione `CreateElasticsearchDomain`, è ora possibile effettuare sia chiamate a `CreateElasticsearchDomain` (vecchia operazione API) che a `CreateDomain` (nuova operazione API). Lo stesso vale per i rifiuti espliciti. Per un elenco delle operazioni API aggiornate, consultare il riferimento all'elemento policy.

Policy SCP

Le policy di controllo dei servizi (SCPs) introducono un ulteriore livello di complessità rispetto allo standard IAM. Per evitare che le policy SCP non vengano rispettate, è necessario aggiungere sia le vecchie operazioni che le nuove per ogni policy SCP. Ad esempio, se un utente dispone attualmente delle autorizzazioni per `CreateElasticsearchDomain`, è necessario concedere loro anche le autorizzazioni per `CreateDomain` in modo che possano mantenere la possibilità di creare domini. Lo stesso vale per i rifiuti espliciti.

Ad esempio:

```
"Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "es:CreateElasticsearchDomain",
```

```
        "es:CreateDomain"
         ...
      ],
    },
      "Effect": "Deny",
      "Action:" [
        "es:DeleteElasticsearchDomain",
        "es:DeleteDomain"
         ...
```

Nuovi tipi di risorsa

OpenSearch Il servizio introduce i seguenti nuovi tipi di risorse:

Risorsa	Description
`AWS::OpenSearchService::Domain`	Rappresenta un dominio Amazon OpenSearch Service. Questa risorsa esiste a livello di servizio e non è specifica del software in esecuzione nel dominio. Si applica a servizi come CloudFormation e AWS Resource Groups, in cui è possibile creare e gestire le risorse per il servizio nel suo complesso. Per istruzioni su come aggiornare i domini definiti in Elasticsearch CloudFormation a OpenSearch, consulta la sezione Note nella Guida per l'utente. CloudFormation
`AWS::OpenSearch::Domain`	Rappresenta il OpenSearch/Elasticsearch software in esecuzione su un dominio. Questa risorsa si applica a servizi come AWS CloudTrailand AWS Config, che fanno riferimento al software in esecuzione sul dominio anziché al OpenSearch servizio nel suo insieme. Questi servizi ora contengono tipi di risorse separati per i domini che eseguono Elasticsearch (`AWS::Elasticsearch`

Risorsa	Description
	`::Domain`) rispetto ai domini che eseguono (). OpenSearch `AWS::OpenSearch::Domain`

> **Note**
>
> In AWS Config, continuerai a visualizzare i tuoi dati nel tipo di `AWS::Elasticsearch::Domain` risorsa esistente per diverse settimane, anche se esegui l'upgrade di uno o più domini a. OpenSearch

Kibana è stato rinominato in Dashboards OpenSearch

OpenSearch Dashboards, l' AWS alternativa a Kibana, è uno strumento di visualizzazione open source progettato per funzionare con. OpenSearch Dopo aver aggiornato un dominio da Elasticsearch a, l'endpoint cambia in. OpenSearch `/_plugin/kibana` `/_dashboards` OpenSearch Il servizio reindirizzerà tutte le richieste al nuovo endpoint, ma se utilizzi l'endpoint Kibana in una delle tue policy IAM, aggiorna tali policy per includere anche il nuovo endpoint. `/_dashboards`

Se utilizzithe section called “Autenticazione SAML per dashboard OpenSearch”, prima di aggiornare il tuo dominio a OpenSearch, devi modificare tutti i Kibana URLs configurati nel tuo provider di identità (IdP) da a. `/_plugin/kibana` `/_dashboards` I più comuni URLs sono Assertion Consumer Service (ACS) e Destinatario. URLs URLs

Il `kibana_read_only` ruolo predefinito per OpenSearch Dashboards è stato rinominato in e il `kibana_user` ruolo è stato rinominato in. `opensearch_dashboards_read_only` `opensearch_dashboards_user` La modifica si applica a tutti i 1 appena creati. OpenSearch domini x che eseguono il software di servizio R20211203 o versione successiva. Se si aggiorna un dominio esistente al software del servizio R20211203, i nomi dei ruoli rimangono invariati.

CloudWatch Metriche rinominate

Diverse CloudWatch metriche cambiano per i domini in esecuzione. OpenSearch Quando aggiorni un dominio a OpenSearch, le metriche cambiano automaticamente e gli CloudWatch allarmi attuali

si interrompono. Prima di aggiornare il cluster da una versione di Elasticsearch a una OpenSearch versione, assicurati di aggiornare gli CloudWatch allarmi per utilizzare le nuove metriche.

Sono stati modificati i seguenti parametri:

Nome parametro originale	Nuovo nome
KibanaHealthyNodes	OpenSearchDashboardsHealthyNodes
KibanaConcurrentConnections	OpenSearchDashboardsConcurrentConnections
KibanaHeapTotal	OpenSearchDashboardsHeapTotal
KibanaHeapUsed	OpenSearchDashboardsHeapUsed
KibanaHeapUtilization	OpenSearchDashboardsHeapUtilization
KibanaOS1MinuteLoad	OpenSearchDashboardsOS1MinuteLoad
KibanaRequestTotal	OpenSearchDashboardsRequestTotal
KibanaResponseTimesMaxInMillis	OpenSearchDashboardsResponseTimesMaxInMillis
ESReportingFailedRequestSysErrCount	KibanaReportingFailedRequestSysErrCount
ESReportingRequestCount	KibanaReportingRequestCount
ESReportingFailedRequestUserErrCount	KibanaReportingFailedRequestUserErrCount
ESReportingSuccessCount	KibanaReportingSuccessCount
ElasticsearchRequests	OpenSearchRequests

Per un elenco completo delle metriche inviate dal OpenSearch Servizio ad Amazon CloudWatch, consultathe section called “Monitoraggio dei parametri del cluster”.

Modifiche della console Gestione fatturazione e costi

I dati storici nella console di Billing and Cost Management e nei report sui costi e sull'utilizzo continueranno a utilizzare il vecchio nome di servizio, quindi devi iniziare a utilizzare i filtri sia per OpenSearch Amazon Service che per il nome precedente di Elasticsearch durante la ricerca dei dati. Se hai già dei report salvati, aggiorna i filtri per assicurarti che includano anche Service. OpenSearch Inizialmente potresti ricevere un avviso quando l'utilizzo di Elasticsearch diminuisce o aumenta OpenSearch, ma scompare nel giro di alcuni giorni.

Oltre al nome del servizio, i campi seguenti verranno modificati per tutti i report, le fatture e le operazioni API del listino prezzi:

Campo	Vecchio formato	Nuovo formato
Tipo di istanza	`m5.large.elasticsearch`	`m5.large.search`
Famiglia di prodotti	Istanza Elasticsearch Volume Elasticsearch	Istanza OpenSearch di servizio Amazon Volume OpenSearch di servizi Amazon
Descrizione dei prezzi	5.098 USD per ora di istanza c5.18xlarge.elasticsearch (o ora parziale) - EU	5.098 USD per ora di istanza c5.18xlarge.search (o ora parziale) - EU
Famiglia di istanze	`ultrawarm.elasticsearch`	`ultrawarm.search`

Nuovo formato evento

Il formato degli eventi che il OpenSearch Servizio invia ad Amazon EventBridge e Amazon CloudWatch è cambiato, in particolare il `detail-type` campo. Il campo di origine (`aws.es`) rimane

lo stesso. Per il formato completo per ogni tipo di evento, consultare the section called “Monitoraggio degli eventi”. Se si dispone di regole di evento esistenti che dipendono dal formato precedente, assicurarsi di aggiornarle in modo che siano conformi al nuovo formato.

Cosa rimane lo stesso?

Le seguenti caratteristiche e funzionalità, tra le altre non elencate, rimarranno invariate:

- Principale del servizio (`es.amazonaws.com`)
- Codice fornitore
- Dominio ARNs
- Endpoint di dominio

Inizia: aggiorna i tuoi domini alla OpenSearch versione 1.x

OpenSearch 1. x supporta gli aggiornamenti dalle versioni 6.8 e 7 di Elasticsearch. x. Per istruzioni su come aggiornare il dominio, consultare the section called “Aggiornamento di un dominio (console)”. Se utilizzi l'API AWS CLI o di configurazione per aggiornare il tuo dominio, devi specificare l'`TargetVersion`as`OpenSearch_1.x`.

OpenSearch 1. x introduce un'impostazione di dominio aggiuntiva denominata Abilita la modalità di compatibilità. Poiché alcuni client e plugin Elasticsearch OSS controllano la versione del cluster prima di connettersi, la modalità di compatibilità imposta in modo OpenSearch da riportare la versione 7.10 in modo che questi client continuino a funzionare.

Puoi abilitare la modalità di compatibilità quando crei OpenSearch domini per la prima volta o quando esegui l'aggiornamento da una versione di Elasticsearch. OpenSearch Se non è impostato, viene ripristinato il parametro predefinito, `false`, quando crei un dominio e `true` quando esegui l'aggiornamento di un dominio.

Per abilitare la modalità di compatibilità utilizzando l'API di configurazione, impostare `override_main_response_version` su `true`:

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/upgradeDomain
{
  "DomainName": "domain-name",
  "TargetVersion": "OpenSearch_1.0",
  "AdvancedOptions": {
```

```
    "override_main_response_version": "true"
   }
}
```

Per abilitare o disabilitare la modalità di compatibilità sui OpenSearch domini esistenti, devi utilizzare l'operazione API _cluster/settings: OpenSearch

```
PUT /_cluster/settings
{
  "persistent" : {
    "compatibility.override_main_response_version" : true
  }
}
```

Risoluzione dei problemi con Amazon OpenSearch Service

Questo argomento descrive come identificare e risolvere i problemi più comuni OpenSearch di Amazon Service. Consulta le informazioni contenute in questa sezione prima di contattare AWS Support.

Impossibile accedere alle OpenSearch dashboard

L'endpoint OpenSearch Dashboards non supporta le richieste firmate. Se la policy di controllo accessi per il dominio consente l'accesso solo a determinati ruoli IAM e non è stata configurata l'autenticazione Amazon Cognito, quando si prova ad accedere a Dashboards è possibile che venga ricevuto il seguente messaggio di errore:

```
"User: anonymous is not authorized to perform: es:ESHttpGet"
```

Se il tuo dominio di OpenSearch servizio utilizza l'accesso VPC, potresti non ricevere questo errore, ma la richiesta potrebbe scadere. Per ulteriori informazioni su come correggere questo problema e le varie opzioni di configurazione disponibili, consulta the section called “Controllo dell'accesso ai pannelli di controllo”, the section called “Informazioni sulle policy d'accesso nei domini VPC” e the section called “Identity and Access Management”.

Impossibile accedere al dominio VPC

Consulta the section called “Informazioni sulle policy d'accesso nei domini VPC” e the section called “Test dei domini VPC”.

Cluster in stato di sola lettura

Rispetto alle versioni precedenti di Elasticsearch ed Elasticsearch 7 OpenSearch . x utilizza un sistema diverso per il coordinamento dei cluster. In questo nuovo sistema, quando il cluster perde il quorum, il cluster non è disponibile finché non si esegue un'azione. La perdita del quorum può assumere due forme:

- Se il cluster utilizza nodi master dedicati, la perdita del quorum si verifica quando metà o più nodi non sono disponibili.

- Se il cluster non utilizza nodi master dedicati, la perdita del quorum si verifica quando metà o più nodi di dati non sono disponibili.

Se si verifica una perdita del quorum e il cluster ha più di un nodo, OpenSearch Service ripristina il quorum e imposta il cluster in uno stato di sola lettura. Sono disponibili due opzioni:

- Rimuovi lo stato di sola lettura e utilizza il cluster così com'è.
- Ripristina il cluster o i singoli indici da uno snapshot.

Se preferisci utilizzare il cluster così com'è, verifica che lo stato del cluster sia verde utilizzando la seguente richiesta:

```
GET _cat/health?v
```

Se lo stato del cluster è rosso, ti consigliamo di ripristinare il cluster da uno snapshot. Puoi anche consultare the section called “Cluster in stato rosso” per la risoluzione dei problemi. Se l'integrità del cluster è verde, controlla che tutti gli indici previsti siano presenti utilizzando la seguente richiesta:

```
GET _cat/indices?v
```

Quindi esegui alcune ricerche per verificare che i dati previsti siano presenti. In caso affermativo, puoi rimuovere lo stato di sola lettura utilizzando la seguente richiesta:

```
PUT _cluster/settings
{
  "persistent": {
    "cluster.blocks.read_only": false
  }
}
```

Se si verifica una perdita del quorum e il cluster ha un solo nodo, OpenSearch Service sostituisce il nodo e non pone il cluster in uno stato di sola lettura. In caso contrario, le opzioni sono le stesse: utilizza il cluster così com'è o ripristina da uno snapshot.

In entrambe le situazioni, OpenSearch Service invia due eventi al tuo. Health Dashboard Il primo segnala la perdita del quorum. Il secondo si verifica dopo che il OpenSearch Servizio ha ripristinato correttamente il quorum. Per ulteriori informazioni sull'utilizzo di Health Dashboard, consulta la Guida per l'AWS Health utente.

Cluster in stato rosso

Lo stato rosso del cluster indica che almeno uno shard primario e le relative repliche non sono allocati a un nodo. OpenSearch Il servizio continua a cercare di scattare istantanee automatiche di tutti gli indici indipendentemente dal loro stato, ma le istantanee falliscono finché persiste lo stato rosso del cluster.

Le cause più comuni di questa condizione sono nodi del cluster con errori e crash del processo OpenSearch causato da un carico di elaborazione costantemente elevato.

> **Note**
>
> OpenSearch Il servizio archivia le istantanee automatizzate per 14 giorni indipendentemente dallo stato del cluster. Pertanto, se lo stato rosso del cluster persiste per più di due settimane, l'ultimo snapshot automatico integro verrà eliminato e i dati del cluster potrebbero essere persi definitivamente. Se il dominio del OpenSearch servizio assume lo stato di cluster rosso, Supporto potresti contattarti per chiederti se desideri risolvere il problema da solo o se desideri ricevere assistenza dal team di supporto. Puoi impostare un CloudWatch allarme per avvisarti quando si verifica lo stato di un cluster rosso.

In definitiva, partizioni rosse causano cluster rossi e indici rossi causano partizioni rosse. Per identificare gli indici che causano lo stato rosso del cluster, OpenSearch ha alcune informazioni utili. APIs

- `GET /_cluster/allocation/explain` sceglie la prima partizione non assegnata che trova e spiega perché non può essere allocata a un nodo:

```
{
    "index": "test4",
    "shard": 0,
    "primary": true,
    "current_state": "unassigned",
    "can_allocate": "no",
    "allocate_explanation": "cannot allocate because allocation is not permitted to
 any of the nodes"
}
```

- `GET /_cat/indices?v` mostra lo stato di integrità, il numero di documenti e l'utilizzo del disco per ogni indice:

```
health status index           uuid                   pri rep docs.count docs.deleted
 store.size pri.store.size
green  open   test1           30h1EiMvS5uAFr2t5CEVoQ   5   0        820            0
       14mb           14mb
green  open   test2           sdIxs_WDT56afFGu5KPbFQ   1   0          0            0
       233b           233b
green  open   test3           GGRZp_TBRZuSaZpAGk2pmw   1   1          2            0
     14.7kb          7.3kb
red    open   test4           BJxfAErbTtu5HBjIXJV_7A   1   0
green  open   test5           _8C6MIXOSxCqVYicH3jsEA   1   0          7            0
     24.3kb         24.3kb
```

L'eliminazione degli indici rossi è il modo più rapido per risolvere uno stato rosso del cluster. A seconda del motivo dello stato del cluster rosso, è possibile ridimensionare il dominio di OpenSearch servizio in modo da utilizzare tipi di istanze più grandi, più istanze o più storage basato su EBS e provare a ricreare gli indici problematici.

Se non è possibile eliminare un indice problematico, puoi ripristinare una snapshot, eliminare documenti dall'indice, modificarne le impostazioni, ridurre il numero di repliche o eliminare altri indici per liberare spazio su disco. Il passaggio importante consiste nel risolvere lo stato rosso del cluster prima di riconfigurare il dominio di servizio. OpenSearch Riconfigurare un dominio con un cluster in stato rosso può peggiorare il problema e portare al blocco del dominio nello stato di configurazione Processing (Elaborazione) finché lo stato rosso non sarà risolto.

Correzione automatico di cluster rossi

Se lo stato del cluster è continuamente rosso per più di un'ora, il OpenSearch Servizio tenta di correggerlo automaticamente reindirizzando gli shard non allocati o ripristinando le istantanee precedenti.

Se non riesce a correggere uno o più indici rossi e lo stato del cluster rimane rosso per un totale di 14 giorni, il OpenSearch Servizio intraprende ulteriori azioni solo se il cluster soddisfa almeno uno dei seguenti criteri:

- Ha una sola zona di disponibilità
- Non ha nodi principali dedicati
- Contiene i tipi di istanze espandibili (T2 o T3)

Al momento, se il cluster soddisfa uno di questi criteri, il OpenSearch Servizio invia notifiche giornaliere nei prossimi 7 giorni per spiegare che se non correggi questi indici, tutti gli shard non assegnati verranno eliminati. Se lo stato del cluster è ancora rosso dopo 21 giorni, il OpenSearch Servizio elimina gli shard non assegnati (archiviazione e calcolo) su tutti gli indici rossi. Riceverai notifiche nel pannello Notifiche della console di OpenSearch servizio per ciascuno di questi eventi. Per ulteriori informazioni, consulta the section called “Eventi sull'integrità del cluster”.

Ripristino da un carico di elaborazione costantemente elevato

Per determinare se un cluster è in stato rosso a causa di un carico di elaborazione costantemente elevato su un nodo di dati, monitora i seguenti parametri del cluster.

Parametro pertinente	Description	Ripristino
JVMMemoryPressione	Specifica la percentuale massima dell'heap di Java utilizzata per tutti i nodi di dati in un cluster. Visualizza la statistica Massimo per questo parametro e cerca drop sempre minori nella pressione della memoria mano aman mano che il garbage collector Java riesce a recuperare memoria sufficiente. Questo modello è probabilmente dovuto a query complesse o campi di dati di grandi dimensioni. I tipi di istanza x86 utilizzano il garbage collector (GC) Concurrent Mark Sweep (CMS), che viene eseguito insieme ai thread dell'applicazione per tenere le pause brevi. Se CMS non è in grado di recuperare memoria sufficiente durante le raccolte normali, attiva una garbage collection (GC) completa, che può portare a lunghe pause dell'appl	Imposta interruttori di memoria per JVM. Per ulteriori informazioni, consulta the section called “JVM OutOfMemoryError”. Se il problema persiste, elimina indici non necessari, riduci il numero o la complessità delle richieste per il dominio, aggiungi istanze oppure utilizza tipi di istanza di dimensioni maggiori.

Parametro pertinente	Description	Ripristino
	icazione con un impatto sulla stabilità del cluster. I tipi di istanza Graviton basati su ARM utilizzano il garbage collector (GC) Garbage-First (G1), che è simile a CMS, ma utilizza ulteriori brevi pause e deframmentazione heap per ridurre ulteriormente la necessità di garbage collection complete. In entrambi i casi, se l'utilizzo della memoria continua a crescere oltre i limiti che il Garbage Collector può recuperare durante le Garbage Collection complete, si OpenSearch blocca con un errore di memoria esaurita. Su tutti i tipi di istanza, è preferibile mantenere l'utilizzo al di sotto dell'80%. L'API `_nodes/stats/jvm` offre un riepilogo delle statistiche JVM, dell'utilizzo dei pool di memoria e della garbage collection: `GET domain-endpoint /_nodes/stats/jvm?pretty`	
CPUUtilization	Specifica la percentuale di risorse della CPU utilizzate per i nodi di dati in un cluster. Visualizza la statistica Maximum (Massimo) per questo parametro e cerca un modello continuo di utilizzo elevato.	Aggiungi nodi di dati o passa a tipi di istanza più grandi per i nodi di dati esistenti.

Parametro pertinente	Description	Ripristino
Nodi	Specifica il numero di nodi in un cluster. Visualizza la statistica Minimum (Minimo) per questo parametro. Questo valore fluttua quando il servizio distribuisce un nuovo parco di istanze per un cluster.	Aggiungi nodi di dati.

Stato giallo del cluster

Un cluster in stato giallo indica che le partizioni principali per tutti gli indici sono allocate sui nodi in un cluster, ma che le partizioni di replica per almeno un indice non lo sono. I cluster a nodo singolo si inizializzano sempre con uno stato di cluster giallo perché non esiste nessun altro nodo a cui Service può assegnare una replica. OpenSearch Per portare il cluster in stato verde, aumenta il numero di nodi. Per ulteriori informazioni, consultare the section called “Dimensionamento dei domini”.

I cluster a più nodi potrebbero avere brevemente uno stato di cluster giallo dopo la creazione di un nuovo indice o dopo un errore di nodo. Questo stato si risolve automaticamente durante la replica dei dati nel cluster. OpenSearch Anche la mancanza di spazio su disco può provocare uno stato del cluster giallo; il cluster può distribuire partizioni di replica solo se i nodi dispongono dello spazio su disco per ospitarle.

ClusterBlockException

L'errore `ClusterBlockException` può presentarsi per i motivi indicati di seguito.

Mancanza di spazio di archiviazione disponibile

Se uno o più nodi del cluster hanno uno spazio di archiviazione inferiore al valore minimo di 1) 20% dello spazio di archiviazione disponibile o 2) 20 GiB di spazio di archiviazione, le operazioni di scrittura di base come l'aggiunta di documenti e la creazione di indici possono iniziare a fallire. the section called “Calcolo dei requisiti di archiviazione”fornisce un riepilogo di come il OpenSearch Servizio utilizza lo spazio su disco.

Per evitare problemi, monitora la `FreeStorageSpace` metrica nella console di OpenSearch servizio e crea CloudWatch allarmi da attivare quando `FreeStorageSpace` scende al di sotto di una certa soglia. `GET /_cat/allocation?v`fornisce anche un utile riepilogo dell'allocazione degli shard e dell'utilizzo del disco. Per risolvere i problemi associati alla mancanza di spazio di archiviazione, ridimensiona il dominio del OpenSearch servizio per utilizzare tipi di istanze più grandi, più istanze o più storage basato su EBS.

Pressione di memoria JVM elevata

Quando la metrica JVMMemoryPressure supera il 92% per 30 minuti, OpenSearch Service attiva un meccanismo di protezione e blocca tutte le operazioni di scrittura per impedire che il cluster raggiunga lo stato rosso. Quando la protezione è attiva, le operazioni di scrittura hanno esito negativo con errore `ClusterBlockException`, non è possibile creare nuovi indici e viene generato l'errore `IndexCreateBlockException`.

Quando la metrica JVMMemoryPressure torna all'88% o inferiore per cinque minuti, la protezione viene disabilitata e le operazioni di scrittura sul cluster vengono sbloccate.

Una pressione della memoria JVM elevata può essere causata da picchi nel numero di richieste al cluster, allocazioni di partizioni sbilanciate tra i nodi, troppe partizioni in un cluster, esplosioni di dati di campo o di mappatura degli indici o tipi di istanze che non sono in grado di gestire i carichi in ingresso. Può essere causata anche dall'utilizzo di aggregazioni, caratteri jolly o ampi intervalli temporali nelle query.

Per ridurre il traffico verso il cluster e risolvere problemi di pressione della memoria JVM elevata, prova una o più delle seguenti operazioni:

- Scala il dominio in modo che la dimensione massima dell'heap per nodo sia di 32 GB.
- Riduci il numero di partizioni eliminando gli indici vecchi o inutilizzati.
- Cancella la cache dei dati con l'operazione API `POST` *`index-name`*`/_cache/clear?fielddata=true`. Tieni presente che la cancellazione della cache può interrompere le query in corso.

In generale, per evitare una pressione della memoria JVM elevata in futuro, segui queste best practice:

- Evita l'aggregazione su campi di testo o modifica il tipo mappatura per i tuoi indici in `keyword`.
- Ottimizza le richieste di ricerca e indicizzazione scegliendo il numero corretto di partizioni.

- Imposta le policy ISM (Index State Management) per rimuovere gli indici inutilizzati con regolarità.

Errore durante la migrazione a Multi-AZ con Standby

I seguenti problemi potrebbero verificarsi quando si migra un dominio esistente su Multi-AZ in modalità standby.

Creazione di un indice, di un modello di indice o di una politica ISM durante la migrazione da domini senza standby a domini con standby

Se si crea un indice durante la migrazione di un dominio da Multi-AZ senza Standby a con Standby e il modello di indice o la politica ISM non seguono le linee guida consigliate sulla copia dei dati, ciò può causare un'incoerenza dei dati e la migrazione potrebbe non riuscire. Per evitare questa situazione, crea il nuovo indice con un numero di copie dei dati (compresi i nodi primari e le repliche) multiplo di tre. Puoi controllare l'avanzamento della migrazione utilizzando l'`DescribeDomainChangeProgress`API. Se si verifica un errore di conteggio delle repliche, correggilo e contatta l'AWS assistenza per riprovare la migrazione.

Numero errato di copie dei dati

Se non disponi del numero corretto di copie dei dati nel tuo dominio, la migrazione a Multi-AZ con Standby avrà esito negativo.

JVM OutOfMemoryError

Un errore `OutOfMemoryError` JVM in genere significa che è stato raggiunto uno dei seguenti interruttori JVM.

Interruttore	Description	Proprietà impostazione cluster
Interruttore padre	Percentuale totale di memoria heap JVM consentita per tutti gli interruttori. Il valore di default è 95%.	`indices.breaker.total.limit`
Interruttore campo dati	Percentuale di memoria heap JVM consentita per	`indices.breaker.fielddata.limit`

Interruttore	Description	Proprietà impostazione cluster
	caricare in memoria un singolo campo di dati. Il valore predefinito è 40%. Se vengono caricati dati con campi di grandi dimensioni, è possibile che sia necessario aumentare tale limite.	
Interruttore richieste	Percentuale di memoria heap JVM consentita per le strutture di dati utilizzat e nelle risposte a una richiesta di servizio. Il valore predefinito è 60%. Se le richieste di servizio prevedono il calcolo di aggregazioni, è consiglia bile aumentare tale limite.	`indices.breaker.request.limit`

Nodi cluster con errori

Le istanze Amazon EC2 potrebbero andare incontro a terminazioni e riavvii imprevisti. In genere, OpenSearch Service riavvia i nodi automaticamente. ma è possibile che uno o più nodi in un cluster OpenSearch rimangano in una condizione di errore.

Per verificare questa condizione, apri la dashboard del dominio nella console OpenSearch di servizio. Passa alla scheda Integrità del cluster e scegli il parametro Numero totale di nodi. Verifica se il numero di nodi indicati è inferiore al numero che hai configurato per il tuo cluster. Se dal parametro emerge che uno o più nodi restano non disponibili per più di un giorno, contatta AWS Support.

Puoi anche impostare un CloudWatch allarme per avvisarti quando si verifica questo problema.

> **Note**
>
> Il parametro Numero totale di nodi non è preciso durante le modifiche della configurazione del cluster e durante la manutenzione di routine per il servizio. Si tratta di un comportamento normale. Il parametro tornerà a indicare il numero corretto di nodi del cluster a breve. Per ulteriori informazioni, consulta the section called “Modifiche di configurazione”.

Per proteggere i cluster da terminazioni e riavvii imprevisti dei nodi, crea almeno una replica per ogni indice del dominio di servizio. OpenSearch

Limite massimo di partizioni superato

OpenSearch oltre a 7. le versioni x di Elasticsearch hanno un'impostazione predefinita di non più di 1.000 shard per nodo. OpenSearch/Elasticsearch genera un errore se una richiesta, ad esempio la creazione di un nuovo indice, comporta il superamento di questo limite. Se si verifica questo errore, sono disponibili diverse opzioni:

- Aggiungere più nodi di dati al cluster.
- Aumentare l'impostazione `_cluster/settings/cluster.max_shards_per_node`.
- Utilizzare l'API _shrink per ridurre il numero di partizioni sul nodo.

Dominio bloccato nello stato di elaborazione

Il dominio del OpenSearch servizio entra nello stato «Elaborazione» quando è nel bel mezzo di una modifica della configurazione. Quando si avvia una modifica alla configurazione, lo stato del dominio passa a «Elaborazione» mentre il OpenSearch Servizio crea un nuovo ambiente. Nel nuovo ambiente, OpenSearch Service lancia un nuovo set di nodi applicabili (come data, master o UltraWarm). Al termine della migrazione, i nodi più vecchi vengono terminati.

Il cluster può rimanere bloccato nello stato "Elaborazione" se si verifica una di queste situazioni:

- Un nuovo set di nodi di dati non può essere avviato.
- La migrazione della partizione al nuovo set di nodi di dati non riesce.
- Il controllo di convalida non è riuscito con errori.

Per i passaggi di risoluzione dettagliati in ciascuna di queste situazioni, consulta Perché il mio dominio Amazon OpenSearch Service è bloccato nello stato «Elaborazione»? .

Saldo di burst EBS basso

OpenSearch Il servizio ti invia una notifica sulla console quando il saldo di burst EBS su uno dei tuoi volumi General Purpose (SSD) è inferiore al 70% e una notifica di follow-up se il saldo scende al di sotto del 20%. Per risolvere questo problema, puoi aumentare le dimensioni del cluster o ridurre gli IOPS di lettura e scrittura in modo da poter accreditare il saldo di burst. Il saldo di espansione rimane a 0 per i domini con tipi di volumi gp3 e i domini con volumi gp2 con una dimensione del volume superiore a 1.000 GiB. Per ulteriori informazioni, consulta General Purpose SSD volumes (gp2) (Volumi SSD a scopo generico [gp2]). Puoi monitorare il burst balance di EBS con la metrica. `BurstBalance` CloudWatch

La metrica EBS aumenta durante il ridimensionamento del volume

Quando modifichi le dimensioni dei volumi di Amazon Elastic Block Store, potresti osservare aumenti temporanei di vari parametri EBS come`Write Throughput`,`Write Throughput Micro bursting`, `Disk Queue Depth` e. `IOPS` Si tratta di un comportamento previsto durante l'operazione di ridimensionamento e in genere dura alcuni secondi. La durata può tuttavia variare in base alla dimensione del volume da modificare.

Per evitare problemi di latenza e il rifiuto delle richieste, esegui il ridimensionamento dei volumi EBS solo quando il cluster è integro e il traffico di rete è basso.

Impossibile abilitare i log di verifica

È possibile che si verifichi il seguente errore quando si tenta di abilitare la pubblicazione dei registri di controllo utilizzando la console di servizio: OpenSearch

La politica di accesso alle risorse specificata per il gruppo di log CloudWatch Logs non concede autorizzazioni sufficienti ad Amazon OpenSearch Service per creare un flusso di log. Controllare la policy di accesso alle risorse.

Se si verifica questo errore, verificare che l'elemento `resource` della policy includa l'ARN del gruppo di log corretto. Se lo fa, procedere come indicato di seguito:

1. Attendere qualche minuto.

2. Aggiornare la pagina nel browser Web.
3. Scegli Seleziona gruppo esistente.
4. Per Gruppo di log esistente, scegliere il gruppo di log creato prima di ricevere il messaggio di errore.
5. Nella sezione Policy di accesso, scegli Seleziona policy esistente.
6. Per Policy esistente, scegliere la policy creata prima di ricevere il messaggio di errore.
7. Scegli Enable (Abilita).

Se l'errore persiste anche dopo aver ripetuto il processo più volte, contattare AWS Support.

Impossibile chiudere l'indice

OpenSearch Il servizio supporta l'`_close`API solo per le versioni 7.4 OpenSearch e successive di Elasticsearch. Se si utilizza una versione più vecchia e si sta ripristinando un indice da uno snapshot, è possibile eliminare l'indice esistente (prima o dopo la reindicizzazione).

Controlli delle licenze client

Le distribuzioni predefinite di Logstash e Beats includono un controllo della licenza proprietaria e non riescono a connettersi alla versione open source di. OpenSearch Assicurati di utilizzare le distribuzioni Apache 2.0 (OSS) di questi client con Service. OpenSearch

Limitazione delle richieste

Se si ricevono errori `403 Request throttled due to too many requests` o `429 Too Many Requests` persistenti, considerare il dimensionamento verticale. Amazon OpenSearch Service limita le richieste se il payload fa sì che l'utilizzo della memoria superi la dimensione massima dell'heap Java.

Impossibile eseguire SSH nel nodo

Non puoi usare SSH per accedere a nessuno dei nodi del tuo OpenSearch cluster e non puoi modificarlo direttamente. `opensearch.yml` Utilizza invece la console o SDKs per configurare il tuo dominio. AWS CLI Puoi specificare alcune impostazioni a livello di cluster anche utilizzando

OpenSearch REST APIs. Per ulteriori informazioni, consulta Amazon OpenSearch Service API Reference ethe section called “Operazioni supportate”.

Per ulteriori dettagli sulle prestazioni del cluster, puoi pubblicare log di errori e visualizzare log in CloudWatch.

Errore snapshot "Non valido per la classe di archiviazione dell'oggetto"

OpenSearch Gli snapshot dei servizi non supportano la classe di storage Amazon Glacier. Potresti riscontrare questo errore quando tenti di elencare le istantanee se il tuo bucket S3 include una regola del ciclo di vita che trasferisce gli oggetti alla classe di storage Amazon Glacier.

Se devi ripristinare uno snapshot dal bucket, ripristina gli oggetti da Amazon Glacier, copia gli oggetti in un nuovo bucket e registra il nuovo bucket come repository di snapshot.

Intestazione dell'host non valida

OpenSearch Il servizio richiede che i client lo specifichino nelle intestazioni della richiesta. `Host` A valore `Host` valido è l'endpoint del dominio senza `https://`, come:

```
Host: search-my-sample-domain-ih2lhn2ew2scurji.us-west-2.es.amazonaws.com
```

Se ricevi un `Invalid Host Header` errore durante la richiesta, verifica che il client o il proxy includa l'endpoint del dominio OpenSearch Service (e non, ad esempio, il relativo indirizzo IP) nell'`Host`intestazione.

Tipo di istanza M3 non valido

OpenSearch Il servizio non supporta l'aggiunta o la modifica di istanze M3 a domini esistenti che eseguono OpenSearch o eseguono versioni di Elasticsearch 6.7 e successive. È possibile continuare a utilizzare le istanze M3 con Elasticsearch 6.5 e versioni precedenti.

Si consiglia di scegliere un tipo di istanza più recente. Per i domini che eseguono Elasticsearch 6.7 OpenSearch o versioni successive, si applicano le seguenti restrizioni:

- Se il dominio esistente non utilizza istanze M3, non è più possibile passare a tali istanze.

- Se si modifica un dominio esistente da un tipo di istanza M3 a un altro tipo di istanza, non è possibile tornare indietro.

Le hot query smettono di funzionare dopo l'attivazione UltraWarm

Quando si abilita UltraWarm su un dominio, se non sono presenti sostituzioni preesistenti all'`search.max_bucketsimpostazione`, OpenSearch Service imposta automaticamente il valore per evitare che le query che richiedono molta memoria `10000` saturino i nodi caldi. Se le tue hot query utilizzano più di 10.000 bucket, potrebbero smettere di funzionare quando abiliti. UltraWarm

Poiché non puoi modificare questa impostazione a causa della natura gestita di Amazon OpenSearch Service, devi aprire una richiesta di assistenza per aumentare il limite. Gli aumenti di limite non richiedono un abbonamento Premium Support.

Impossibile eseguire il downgrade dopo l'aggiornamento

Gli aggiornamenti locali sono irreversibili, ma se si contatta il AWS Supporto, è possibile ripristinare lo snapshot automatico pre-aggiornamento in un nuovo dominio. Ad esempio, se aggiorni un dominio da Elasticsearch 5.6 a 6.4, AWS Support può aiutarti a ripristinare lo snapshot precedente all'aggiornamento su un nuovo dominio Elasticsearch 5.6. Se dal dominio originale è stato preso uno snapshot manuale, è possibile eseguire questa fase da soli.

Hai bisogno di un riepilogo dei domini per tutti Regioni AWS

Lo script seguente utilizza il AWS CLI comando Amazon EC2 describe-regions per creare un elenco di tutte le regioni in cui OpenSearch il servizio potrebbe essere disponibile. Quindi chiama per ogni regione list-domain-names:

```
for region in `aws ec2 describe-regions --output text | cut -f4`
do
    echo "\nListing domains in region '$region':"
    aws opensearch list-domain-names --region $region --query 'DomainNames'
done
```

Per ogni regione si riceve il seguente output:

```
Listing domains in region:'us-west-2'...
```

```
[
  {
    "DomainName": "sample-domain"
  }
]
```

Le regioni in cui il OpenSearch servizio non è disponibile restituiscono «Impossibile connettersi all'URL dell'endpoint».

Errore del browser durante l'utilizzo OpenSearch delle dashboard

Il browser inserisce i messaggi di errore del servizio negli oggetti di risposta HTTP quando si utilizzano le dashboard per visualizzare i dati nel dominio del servizio. OpenSearch Puoi usare gli strumenti di sviluppo comunemente disponibili nei Web browser, ad esempio la modalità di sviluppo in Chrome, per visualizzare gli errori del servizio sottostanti e semplificare le attività di debug.

Per visualizzare gli errori del servizio in Chrome

1. Dalla barra dei menu principale di Chrome, scegliere Visualizza, Sviluppatore, Strumenti per gli sviluppatori.
2. Scegliere la scheda Network (Rete).
3. Nella colonna Status (Stato) scegliere qualsiasi sessione HTTP con stato 500.

Per visualizzare gli errori del servizio in Firefox

1. Dal menu scegliere Tools (Strumenti), Web Developer (Sviluppatore Web), Network (Rete).
2. Scegliere qualsiasi sessione HTTP con stato 500.
3. Scegliere la scheda Response (Risposta) per visualizzare la risposta del servizio.

Asimmetria di partizioni e storage di nodi

L'asimmetria di partizioni di nodi si verifica quando uno o più nodi all'interno di un cluster presentano un numero significativamente maggiore di partizioni rispetto agli altri nodi. L'asimmetria di storage di nodi si verifica quando uno o più nodi all'interno di un cluster presentano una quantità di storage significativamente maggiore (`disk.indices`) rispetto agli altri nodi. Sebbene entrambe queste condizioni possano verificarsi temporaneamente, ad esempio quando un dominio ha sostituito un nodo e sta ancora allocando partizioni su quel nodo, se persistono, devono essere risolte.

Per identificare entrambi i tipi di asimmetria, esegui l'operazione API _cat/allocation e confronta le voci `shards` e `disk.indices` nella risposta:

```
shards    | disk.indices  | disk.used  | disk.avail   | disk.total   | disk.percent |
host     | ip       | node
   264    |      465.3mb  |   229.9mb  |      1.4tb   |      1.5tb   |            0 |
 x.x.x.x  | x.x.x.x  | node1
   115    |        7.9mb  |    83.7mb  |     49.1gb   |     49.2gb   |            0 |
x.x.x.x  | x.x.x.x  | node2
   264    |      465.3mb  |   235.3mb  |      1.4tb   |      1.5tb   |            0 |
 x.x.x.x  | x.x.x.x  | node3
   116    |        7.9mb  |    82.8mb  |     49.1gb   |     49.2gb   |            0 |
x.x.x.x  | x.x.x.x  | node4
   115    |        8.4mb  |      85mb  |     49.1gb   |     49.2gb   |            0 |
x.x.x.x  | x.x.x.x  | node5
```

Sebbene alcune asimmetrie di storage siano normali, qualsiasi valore superiore al 10% rispetto alla media è significativo. Quando la distribuzione delle partizioni è asimmetrica, anche l'utilizzo della CPU, della rete e della larghezza di banda del disco può essere asimmetrico. Poiché un maggior numero di dati significa in genere un maggior numero di operazioni di indicizzazione e ricerca, i nodi più pesanti tendono ad essere anche quelli con maggiori risorse, mentre i nodi più leggeri rappresentano una capacità sottoutilizzata.

Correzione: utilizza conteggi di partizioni che siano multipli del numero di nodi di dati per garantire che ogni indice sia distribuito uniformemente tra i nodi di dati.

Asimmetria di partizioni e storage di indici

L'asimmetria di partizioni di indici si verifica quando uno o più nodi contengono più partizioni di un indice rispetto agli altri nodi. L'asimmetria di storage di indici si verifica quando uno o più nodi contengono una quantità sproporzionatamente elevata di storage totale di un indice.

L'asimmetria di indici è più difficile da identificare rispetto all'asimmetria di nodi perché richiede una certa manipolazione dell'output dell'API _cat/shards. Esamina l'asimmetria di indici per verificare se esiste qualche indicazione di asimmetria nei parametri del cluster o del nodo. Di seguito sono riportate le indicazioni comuni dell'asimmetria di indici:

- Errori HTTP 429 che si verificano su un sottoinsieme di nodi di dati
- Accodamento non uniforme delle operazioni di indicizzazione o di ricerca tra i nodi di dati
- Utilizzo irregolare della CPU JVM heap tra i nodi di dati and/or

Correzione: utilizza conteggi di partizioni che siano multipli del numero di nodi di dati per garantire che ogni indice sia distribuito uniformemente tra i nodi di dati. Se la memorizzazione dell'indice o lo shard skew persistono, potrebbe essere necessario forzare la riallocazione degli shard, operazione che si verifica con ogni distribuzione blu/verde del dominio di servizio. OpenSearch

Operazione non autorizzata dopo la selezione dell'accesso VPC

Quando crei un nuovo dominio utilizzando la console di OpenSearch servizio, hai la possibilità di selezionare VPC o accesso pubblico. Se si seleziona Accesso VPC, il OpenSearch servizio richiede informazioni sul VPC e fallisce se non si dispone delle autorizzazioni appropriate:

```
You are not authorized to perform this operation. (Service: AmazonEC2; Status Code:
 403; Error Code: UnauthorizedOperation
```

Per permettere questa query, è necessario disporre di accesso alle operazioni `ec2:DescribeVpcs`, `ec2:DescribeSubnets` ed `ec2:DescribeSecurityGroups`. Questo requisito riguarda solo la console. Se utilizzi la AWS CLI per creare e configurare un dominio con un endpoint VPC, non è necessario accedere a tali operazioni.

Caricamento bloccato dopo la creazione di un dominio VPC

Dopo aver creato un nuovo dominio che usa l'accesso VPC, il valore di Configuration state (Stato di configurazione) del dominio potrebbe non andare mai oltre Loading (Caricamento). Se si verifica questo problema, probabilmente hai disabilitato AWS Security Token Service (AWS STS) per la tua regione.

Per aggiungere endpoint VPC al tuo VPC, OpenSearch Service deve assumere il ruolo. `AWSServiceRoleForAmazonOpenSearchService` Pertanto, AWS STS deve essere abilitato a creare nuovi domini che utilizzano l'accesso VPC in una determinata regione. Per ulteriori informazioni sull'attivazione e la disabilitazione AWS STS, consulta la IAM User Guide.

Richieste rifiutate all'API OpenSearch

Con l'introduzione del controllo degli accessi basato su tag per l' OpenSearch API, potresti iniziare a vedere errori di accesso negato dove prima non si verificavano. Ciò potrebbe essere dovuto al fatto che una o più policy di accesso contiene `Deny` che usa la condizione `ResourceTag` e tali condizioni ora sono soddisfatte.

Ad esempio, la policy seguente utilizzata per negare l'accesso solo all'azione `CreateDomain` dall'API di configurazione, se il dominio aveva il tag `environment=production`. Anche se l'elenco delle azioni include anche `ESHttpPut`, la dichiarazione di negazione non si applicava a tale o ad altre azioni `ESHttp*`.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Action": [
      "es:CreateDomain",
      "es:ESHttpPut"
    ],
    "Effect": "Deny",
    "Resource": "*",
    "Condition": {
      "ForAnyValue:StringEquals": {
        "aws:ResourceTag/environment": [
          "production"
        ]
      }
    }
  }]
}
```

Con il supporto aggiunto dei tag per i metodi OpenSearch HTTP, una policy basata sull'identità IAM come quella sopra riportata comporterà il rifiuto dell'accesso all'azione all'utente collegato. `ESHttpPut` In precedenza, in assenza di convalida di tag, l'utente collegato poteva comunque inviare richieste PUT.

Se inizi a scoprire errori di accesso negato dopo aver aggiornato i domini al software di servizio R20220323 o versioni successive, controlla le policy di accesso basate sull'identità per verificare se il problema è questo e aggiornali, se necessario, per consentire l'accesso.

Impossibile connettersi da Alpine Linux

Alpine Linux limita la dimensione della risposta DNS a 512 byte. Se tenti di connetterti al tuo dominio di OpenSearch servizio da Alpine Linux versione 3.18.0 o precedente, la risoluzione DNS può fallire

se il dominio si trova in un VPC e ha più di 20 nodi. Se utilizzi una versione di Alpine Linux successiva alla 3.18.0, dovresti essere in grado di risolvere più di 20 host. Per ulteriori informazioni, consulta le note di rilascio di Alpine Linux 3.18.0.

Se il proprio dominio è in un VPC, si consiglia di usare altre distribuzioni Linux, come Debian, Ubuntu, CentOS, Red Hat Enterprise Linux o Amazon Linux 2, per connettersi a esso.

Troppe richieste per Search Backpressure

Il controllo degli accessi basato sulla CPU è un meccanismo di controllo che limita in modo proattivo il numero di richieste a un nodo in base alla sua capacità attuale, sia in caso di aumenti organici che di picchi di traffico. Un numero eccessivo di richieste restituisce un codice di stato HTTP 429 «Troppe richieste» in caso di rifiuto. Questi errori indicano risorse di cluster insufficienti, richieste di ricerca che richiedono molte risorse o un picco involontario del carico di lavoro.

Search Backpressure fornisce il motivo del rifiuto, che può aiutare a ottimizzare le richieste di ricerca che richiedono molte risorse. In caso di picchi di traffico, consigliamo di ripetere i tentativi sul lato client con backoff e jitter esponenziali.

Errore di certificato quando si utilizza un SDK

Poiché AWS SDKs utilizzi i certificati CA del tuo computer, le modifiche ai certificati sui AWS server possono causare errori di connessione quando tenti di utilizzare un SDK. I messaggi di errore variano, ma in genere contengono il testo seguente:

```
Failed to query OpenSearch
...
SSL3_GET_SERVER_CERTIFICATE:certificate verify failed
```

Puoi prevenire questi errori conservando i certificati CA e il sistema operativo del tuo computer. up-to-date Se si riscontra questo problema in un ambiente aziendale e non si gestisce il computer in uso, potrebbe essere necessario richiedere all'amministratore assistenza con il processo di aggiornamento.

Nell'elenco seguente vengono riportati i requisiti minimi del sistema operativo e delle versioni di Java:

- Le versioni di Microsoft Windows in cui sono installati aggiornamenti di gennaio 2005 o versioni successive contengono almeno uno dei requisiti richiesti CAs nell'elenco di attendibilità.

- Mac OS X 10.4 con Java per Mac OS X 10.4 Release 5 (febbraio 2007), Mac OS X 10.5 (ottobre 2007) e le versioni successive contengono almeno uno dei requisiti richiesti CAs nell'elenco di attendibilità.
- Red Hat Enterprise Linux 5 (marzo 2007), 6 e 7 e CentOS 5, 6 e 7 contengono tutti almeno uno dei requisiti richiesti CAs nell'elenco di CA affidabili predefinito.
- Java 1.4.2_12 (maggio 2006), 5 Update 2 (marzo 2005) e tutte le versioni successive, tra cui Java 6 (dicembre 2006), 7 e 8, contengono almeno una delle CA richieste CAs nell'elenco di CA affidabili predefinito.

Le tre autorità di certificazione sono:

- Amazon CA root 1
- Autorità di certificazione root dei servizi Starfield - G2
- Autorità di certificazione Starfield classe 2

I certificati root delle prime due autorità sono disponibili presso Amazon Trust Services, ma mantenere il computer up-to-date è la soluzione più semplice. Per ulteriori informazioni sui certificati forniti da ACM, consulta. AWS Certificate Manager FAQs

> **Note**
>
> Attualmente, i domini di OpenSearch servizio nella regione us-east-1 utilizzano certificati di un'autorità diversa. Prevediamo di aggiornare la regione per l'uso di queste nuove autorità di certificazione in futuro.

L'installazione del plug-in personalizzato non riesce a causa della compatibilità della versione

Problema: l'installazione del plug-in non è riuscita a causa di una mancata corrispondenza tra la versione del plug-in e l' OpenSearch istanza in esecuzione. Il sistema restituisce il seguente errore:

```
PluginValidationFailureReason : The provided plugin could not be loaded.
```

Causa: il plugin è stato compilato per OpenSearch ${*MAJOR*}. ${*MINOR*}.{*PATCH*}, ma il tuo ambiente è in esecuzione OpenSearch ${*MAJOR*}. $ {*MINOR*} .0. OpenSearch richiede l'esatta

corrispondenza della versione tra i plugin e l' OpenSearchinstallazione principale per motivi di stabilità e sicurezza.

Possibile correzione: crea il plugin con la OpenSearch versione ${*MAJOR*}. ${*MINOR*} .0 in modo che corrisponda alla versione del cluster in uso.

Per verificare e aggiornare la versione di OpenSearch

1. Utilizza l'API o la dashboard del cluster per eseguire il comando seguente. Sostituisci *default placeholder values* con le informazioni appropriate.

   Richiesta API:

   ```
   curl -X GET your-opensearch-endpoint/
   ```

   Console Dev Tools nella dashboard:

   ```
   GET /
   ```

   Il comando restituisce le informazioni nel seguente formato.

   ```
   {
     "name": "node-id",
     "cluster_name": "account-id:domain-name",
     "cluster_uuid": "cluster-uuid",
     "version": {
       "distribution": "opensearch",
       "number": "2.17.0",
       "build_type": "tar",
       "build_hash": "unknown",
       "build_date": "2024-12-17T11:00:09.799828091Z",
       "build_snapshot": false,
       "lucene_version": "9.11.1",
       "minimum_wire_compatibility_version": "7.10.0",
       "minimum_index_compatibility_version": "7.0.0"
     },
     "tagline": "The OpenSearch Project: https://opensearch.org/"
   }
   ```

2. Se il numero di versione non è ${*MAJOR*}. ${*MINOR*} .0, ricostruisci il plugin effettuando le seguenti operazioni:

a. Aggiorna i plugin `descriptor.properties` per specificare la versione $. *{MAJOR}* $ *{MINOR}* .0.

b. Ricostruisci il plugin usando il comando corrispondente al tuo tipo di progetto.

c. Esegui il comando update-package usando il file appena creato. `.zip`

d. Esegui il comando associate-package per associare l'ultima versione del plugin creata quando hai eseguito il `update-package` comando nel passaggio precedente.

Cronologia dei documenti per Amazon OpenSearch Service

Questo argomento descrive importanti modifiche ad Amazon OpenSearch Service. Gli aggiornamenti software del servizio aggiungono il supporto per nuove funzionalità, patch di sicurezza, correzioni di bug e altri miglioramenti. Per utilizzare nuove funzionalità, potrebbe essere necessario aggiornare il software del servizio nel dominio. Per ulteriori informazioni, consulta the section called "Aggiornamenti del software del servizio".

Le funzionalità del servizio vengono implementate in modo incrementale a seconda del Regioni AWS luogo in cui il servizio è disponibile. Aggiorniamo questa documentazione solo per la prima versione. Non forniamo informazioni sulla disponibilità delle regioni e non annunciamo implementazioni successive delle regioni. Per informazioni sulla disponibilità regionale delle funzionalità del servizio e per iscriverti alle notifiche sugli aggiornamenti, vedi Cosa c'è di AWS nuovo?

Per ricevere le notifiche sugli aggiornamenti, è possibile sottoscrivere il feed RSS.

> **Note**
>
> Rilascio patch: le versioni del software di servizio che terminano con "-P" e un numero, ad esempio R20211203-P4, sono versioni di patch. È probabile che le patch includano miglioramenti delle prestazioni, correzioni di bug lievi e correzioni di sicurezza o miglioramenti della posizione. Poiché le patch non includono nuove funzionalità o modifiche importanti, generalmente non hanno un impatto diretto sull'utente o sulla documentazione, per cui le specifiche di ogni patch non sono incluse nella cronologia del documento.

Modifica	Descrizione	Data
Support per OpenSearch 3.5	Amazon OpenSearch Service supporta le OpenSearch versioni 3.4 e 3.5. Per ulteriori informazioni, consulta le note di rilascio OpenSearch 3.4 e le note di rilascio OpenSearch 3.5. Per informazioni sulle operazioni OpenSearch 3.4 e 3.5, consultate il riferimento	17 marzo 2026

	all'API OpenSearch REST o il riferimento all'API per il plug-in specifico. Per un elenco completo delle modifiche più importanti, consulta la OpenSearch documentazione qui.	
Nuove funzionalità: accelerazione GPU per indici vettoriali, ottimizzazione automatica e ingestione vettoriale	Amazon OpenSearch Service ora supporta l'accelerazione GPU per gli indici vettoriali, consentendo operazioni di ricerca vettoriale più rapide. Amazon OpenSearch Service introduce anche l'ottimizzazione automatica per ottimizzare automaticamente le prestazioni di raccolta e l'ingestione vettoriale per semplificare il caricamento dei dati. Per ulteriori informazioni, consulta Accelerazione GPU per indici vettoriali, Ottimizzazione automatica e Inserimento vettoriale.	2 dicembre 2025

Support per OpenSearch 3.3	Amazon OpenSearch Service supporta le OpenSearch versioni 3.2 e 3.3. Per ulteriori informazioni, consulta le note di rilascio OpenSearch 3.2 e le note di rilascio OpenSearch 3.3. Per informazioni sulle operazioni OpenSearch 3.2 e 3.3, consultate il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.	19 novembre 2025
Support per famiglie di istanze basate su Graviton4 (c8g, m8g, r8g r8gd)	Amazon OpenSearch Service ora supporta famiglie di istanze basate su Graviton4 (c8g, m8g, r8g, r8gd). Per ulteriori informazioni, consulta Tipi di istanze supportati in Amazon OpenSearch Service.	17 ottobre 2025

Nuova funzionalità: OpenSearch Ingestion, apprendimento automatico, inferenza in batch offline	Le pipeline Amazon OpenSearch Ingestion (OSI) supportano l'elaborazione di inferenze in batch offline di machine learning (ML) per arricchire in modo efficiente grandi volumi di dati a basso costo. Usa l'inferenza in batch offline ogni volta che hai set di dati di grandi dimensioni che possono essere elaborati in modo asincrono. L'inferenza batch offline funziona con Amazon Bedrock e SageMaker i modelli. Questa funzionalità è disponibile in tutti i domini Regioni AWS che supportano OpenSearch Ingestion with OpenSearch Service 2.17+. Per ulteriori informazioni, consulta Utilizzo di una pipeline di OpenSearch ingestione con inferenza batch offline di machine learning.	2 ottobre 2025

Nuova funzionalità: CloudFormation modelli per configurare l'inferenza remota per l'evidenziazione semantica	L'evidenziazione semantica è una funzionalità di ricerca avanzata che migliora la pertinenza dei risultati analizzando il significato e il contesto delle query anziché basarsi esclusivamente sulle corrispondenze esatte delle parole chiave. Questa funzionalità utilizza modelli di apprendimento automatico per valutare la somiglianza semantica tra le query di ricerca e il contenuto dei documenti, identificando ed evidenziando le frasi o i passaggi più rilevanti dal punto di vista contestuale all'interno dei documenti. Amazon OpenSearch Service ora include un CloudFormation modello che configura automaticamente i SageMaker modelli Amazon per l'evidenziazione semantica. Per ulteriori informazioni, consulta Inferenza remota per modelli di evidenziazione semantica.	26 settembre 2025
La politica AmazonOpenSearchIngestionServiceRolePolicy gestita è stata aggiornata	L'aggiornamento concede a OpenSearch Ingestion l'autorizzazione a modificare gli endpoint VPC creati da per condividere le pipeline OpenSearch tra loro. VPCs	19 settembre 2025

Nuova funzionalità: configura le pipeline di OpenSearch ingestione per l'ingestione tra account	Per sorgenti basate su push come HTTP e OTel Amazon OpenSearch Ingestion ti consente di condividere le pipeline Account AWS da un cloud privato virtuale (VPC) a un endpoint di pipeline in un VPC separato. Per ulteriori informazioni, consulta Configurazione delle pipeline di ingestione per l'ingestione tra account OpenSearch	19 settembre 2025
Amazon OpenSearch Serverless supporta la ricerca vettoriale basata su disco	La ricerca vettoriale basata su disco riduce in modo significativo i costi operativi per i carichi di lavoro vettoriali in ambienti con poca memoria. Per ulteriori informazioni, consulta Ricerca vettoriale basata su disco.	18 settembre 2025

Support per OpenSearch 3.1	Amazon OpenSearch Service supporta OpenSearch la versione 3.1. Per informazioni sulla compatibilità degli aggiornamenti, consulta Percorsi di aggiornamento supportati. Per ulteriori informazioni sulle modifiche nelle nuove versioni, consultate la documentazione OpenSearch 3.0 e 3.1. Per informazioni sulle operazioni OpenSearch 3.1, consultate il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.	9 settembre 2025
Amazon OpenSearch Serverless supporta gli standard federali di elaborazione delle informazioni (FIPS) 140-2	FIPS 140-2 è uno standard governativo statunitense e canadese che specifica i requisiti di sicurezza per i moduli crittografici che proteggono le informazioni sensibili. Gli endpoint FIPS Amazon OpenSearch Serverless sono disponibili Regioni AWS laddove è supportato FIPS. Questi endpoint utilizzano TLS 1.2 o versioni successive e algoritmi crittografici convalidati da FIPS per tutte le comunicazioni.	5 settembre 2025

Nuova funzionalità: controllo granulare degli accessi basato sugli attributi SAML	Amazon OpenSearch Service supporta il controllo granulare degli accessi con SAML per mappare dinamicam ente utenti e gruppi dal tuo provider di identità a utenti e ruoli di controllo degli accessi OpenSearch granulari . Puoi assegnare questi ruoli a OpenSearch domini specifici e raccolte serverles s e definire autorizzazioni a livello di indice e sicurezza a livello di documento. Per ulteriori informazioni, consulta Configurare il controllo granulare degli accessi basato sugli attributi SAML.	8 agosto 2025

Nuovo argomento: endpoint e quote dell'interfaccia utente OpenSearch	Amazon OpenSearch UI estende la sua esperienza di analisi operativa modernizzata a nuove Regioni AWS, tra cui Asia Pacifico (Hyderabad), Asia Pacifico (Osaka), Asia Pacifico (Seoul), Europa (Milano), Europa (Zurigo), Europa (Spagna) e Stati Uniti occidentali (California settentrionale), consentendoti di ottenere informazioni su dati che spaziano dai domini gestiti alle raccolte serverless da un unico endpoint. OpenSearch L'interfaccia utente può connettersi a OpenSearch domini precedenti alla versione 1.3 e a raccolte OpenSearch serverless. Ora è disponibile in 22. Regioni AWS	7 agosto 2025

Nuove funzionalità: arricchimento semantico automatico	L'arricchimento semantico automatico semplifica l'implementazione e le funzionalità della ricerca semantica in Amazon Service. OpenSearch La ricerca semantica restituisce risultati di query che incorpora no non solo la corrispon denza delle parole chiave, ma anche l'intento e il significato contestuale della ricerca dell'utente. Per ulteriori informazioni, vedere Informazioni sull'arricchimento semantico automatico. L'arricchimento semantico automatico offre i seguenti vantaggi: Implementazione semplificata Non hai bisogno di competenze di machine learning (ML) o integrazioni complesse. Processo automatizzato L'arricchimento semantico avviene automaticamente durante l'ingestione dei dati. Migliore rilevanza della ricerca L'arricchimento semantico migliora la qualità e	6 agosto 2025

l'accuratezza contestuale dei risultati di ricerca.

Scalabilità

L'arricchimento semantico applica la ricerca semantica a set di dati di grandi dimensioni senza intervento manuale.

Nuova funzionalità: Machine Learning su Amazon OpenSearch Serverless	OpenSearch Serverless ora supporta funzionalità di Machine Learning per flussi di lavoro di intelligenza artificiale come RAG (Retrieval-Augmented Generation) e ricerca semantica. Utilizza i connettori per accedere agli algoritmi ML e ai modelli remoti ospitati su piattaforme di terze parti oppure. Servizi AWS Per ulteriori informazioni, consulta Configurare Machine Learning su Amazon OpenSearch Serverless. Machine Learning on OpenSearch Serverless include le seguenti funzionalità: Ricerca neurale e ricerca ibrida Esegui operazioni di ricerca semantica che comprendono il contesto e il significato delle query per ottenere risultati più pertinenti. Combina la ricerca semantica e basata su parole chiave con la ricerca ibrida per migliorare la pertinenza. Per ulteriori informazioni, consulta Configurare la ricerca	6 agosto 2025

	neurale e la ricerca ibrida su Serverless. OpenSearch
Connettori e modelli	
	Connect a modelli ospitati su piattaforme ML di terze parti e Servizi AWS. Configura i connettori con i modelli e implementali per eseguire previsioni e flussi di lavoro di intellige nza artificiale.
Flussi di lavoro	
	Automatizza complesse attività di configurazione e pre-elaborazione ML concatenando più chiamate API. I flussi di lavoro semplificano la creazione di applicazi oni AI. OpenSearch Per ulteriori informazi oni, consulta Configura re i flussi di lavoro su OpenSearch Serverless.

Nuove funzionalità: istantanee serverless OpenSearch	Le istantanee sono point-in-time backup delle tue raccolte Amazon OpenSearch Serverless che forniscono funzionalità di disaster recovery. OpenSearch Serverless crea e gestisce automaticamente istantanee delle tue raccolte, garantendo la continuità aziendale e la protezione dei dati. Vantaggi principali • Backup orari automatici senza necessità di configurazione manuale • Zero spese di manutenzione • Nessun costo di archiviazione aggiuntivo • Ripristino rapido in caso di perdita accidentale di dati • Capacità di ripristinare indici specifici da un'istantanea Per ulteriori informazioni, consulta Backup delle raccolte mediante istantanee.	30 luglio 2025
OpenSearch Ingestion supporta due nuove fonti	OpenSearch Ingestion ora supporta le seguenti fonti: • Amazon Aurora • Amazon RDS	17 luglio 2025

Nuove funzionalità: integrazione dei OpenSearch servizi con Amazon S3 Vectors	Questa integrazione offre le seguenti funzionalità per migliorare notevolmente la ricerca vettoriale in Service. OpenSearch • Amazon S3 Vectors offre il primo archivio di oggetti nel cloud con supporto nativo per archiviare e interrogare vettori. S3 Vectors offre uno storage vettoriale economico, elastico e durevole che può essere interrogato in base al significato e alla somiglianza semantici. Offre tempi di risposta alle query inferiori al secondo e costi fino al 90% inferiori per il caricamento, l'archiviazione e l'interrogazione dei vettori. • Amazon OpenSearch Service offre la possibilità di utilizzare Amazon S3 come motore vettoriale per indici vettoriali. Questa funzionalità consente di scaricare dati vettoriali su Amazon S3 mantenendo funzionalità di ricerca vettoriale in meno di un secondo a costi contenuti. Per ulteriori informazioni, consulta (Anteprima) Integrazi	15 luglio 2025

	one OpenSearch del servizio con Amazon S3 Vectors.	
Nuovo argomento: integrazione di Atlassian con Ingestion OpenSearch	Ora puoi utilizzare le pipeline di Amazon OpenSearch Ingestion per elaborare i dati dai servizi Atlassian Jira e Confluence. Questa integrazione ti consente di creare una knowledge base unificata e ricercabile sincronizzando progetti Jira completi e spazi Confluence in. OpenSearch L'integrazione mantiene la rilevanza in tempo reale attraverso il monitoraggio continuo e la sincronizzazione automatica degli aggiornamenti, offrendo al contempo opzioni di filtro flessibili per progetti, tipi di problemi e contenuti specifici. Per ulteriori informazioni, consulta Utilizzo di una pipeline di OpenSearch ingestione con Atlassian Services.	5 giugno 2025

Gestione personalizzata dei plugin con AWS CLI	Ora puoi gestire i plugin personalizzati per Amazon OpenSearch Service utilizzando il AWS CLI. Questa funzionalità consente di installare, aggiornare, proteggere e disinstallare i plug-in personalizzati tramite un flusso di lavoro semplificato a riga di comando. Puoi anche proteggere i plug-in con AWS KMS chiavi e gestire gli aggiornamenti dei plug-in durante l'aggiornamento a versioni più recenti di. OpenSearch Per ulteriori informazioni, consulta Gestione dei plug-in personalizzati utilizzando l'integrazione AWS CLI di pacchetti AWS KMS personalizzati OpenSearch di Amazon Service	27 maggio 2025
Support per OpenSearch 2.19	Amazon OpenSearch Service supporta la OpenSearch versione 2.19. Per ulteriori informazioni, consulta le note di rilascio della versione OpenSearch 2.19. Per informazioni sulle operazioni della OpenSearch versione 2.19, consultate il riferimento all'API OpenSearch REST o il riferimento all'API per il plug-in specifico.	30 aprile 2025

Nuovo supporto per specificare i ruoli della pipeline in Ingestion OpenSearch	Quando crei o aggiorni una pipeline in Amazon OpenSearch Ingestion, non hai più un ruolo di pipeline specifico in una configurazione di pipeline in formato YAML. Invece, nella console, specifichi il ruolo in un nuovo campo del ruolo Pipeline. Se si utilizza il AWS CLI, si utilizza un nuovo `--pipeline-role-arn` parametro. Per ulteriori informazioni, consulta i seguenti argomenti: • Configurazione di ruoli e utenti in Amazon OpenSearch Ingestion • Creazione di pipeline Amazon OpenSearch Ingestion • Tutorial: importazione di dati in un dominio utilizzando Amazon Ingestion OpenSearch • Tutorial: Inserimento di dati in una raccolta con Amazon Ingestion OpenSearch • CreatePipeline • UpdatePipeline	22 aprile 2025

Nuovo argomento	L'argomento sulla risoluzione dei problemi L'installazione del plug-in personalizzato non riesce a causa della compatibilità della versione, illustra i passaggi da eseguire quando si verifica l'errore`PluginValidationFailureReason : The provided plugin could not be loaded` . Questo errore è in genere il risultato di una mancata corrispondenza di versione tra il plug-in e l' OpenSearch istanza in esecuzione.	21 aprile 2025
Nuova funzionalità: compatibilità OpenSearch dell'interfaccia utente con la ricerca tra cluster	OpenSearch L'interfaccia utente è ora compatibile con la ricerca tra cluster. In questo modo è possibile utilizzare l' OpenSearch interfaccia utente in un'unica interfaccia Regione AWS per accedere ai cluster in una regione diversa. Questo viene fatto configurandolo come un cluster remoto connesso a un cluster nella stessa regione. Per ulteriori informazioni, consulta Accesso ai dati tra aree geografiche e tra account con ricerca tra cluster.	16 aprile 2025

Contenuto rivisto: «Utilizzo OpenSearch dell'interfaccia utente in Amazon OpenSearch Service»	Abbiamo revisionato e ampliato il capitolo Utilizzo dell'interfaccia utente OpenSearch in Amazon OpenSearch Service, includendo procedure estese, dettagli di configurazione aggiuntivi e una nuova cronologia delle versioni del supporto di Amazon OpenSearch Service per l'interfaccia utente. OpenSearch	15 aprile 2025
Sviluppatore Amazon Q per Amazon OpenSearch Service	L'integrazione di Amazon Q con Amazon OpenSearch Service offre le seguenti funzionalità generative: • Genera visualizzazioni utilizzando il linguaggio naturale • Visualizza riepiloghi e approfondimenti sugli avvisi • Visualizza i riepiloghi dei risultati delle query generati da Amazon Q nella pagina Discover • Visualizza i rilevatori di anomalie consigliati • Accedi alla chat di Amazon Q per OpenSearch domande correlate	31 marzo 2025

La policy AmazonOpenSearchServiceRolePolicy gestita è stata aggiornata	La policy consente OpenSearch di aggiornare l'ambito di accesso di qualsiasi AWS IAM Identity Center applicazione gestita solo da OpenSearch.	28 marzo 2025
OR2 e OM2 istanze per Amazon Service OpenSearch	Amazon OpenSearch Service ora supporta OR2 e OM2 istanze. Nei benchmark interni, OR2 il throughput di indicizzazione è risultato superiore del 26% rispetto a R7 e del 70%. OR1 OM2 è risultato migliore del 15% e del 66% rispetto a M7g. OR1	25 marzo 2025
Interrogazione diretta per CloudWatch Logs e Security Lake	Amazon OpenSearch Service ora supporta le query dirette per interrogare i dati in CloudWatch Logs e Security Lake.	1 dicembre 2024
Indici K-nn	A partire dalla OpenSearch versione 2.17, è possibile spostare gli indici K-nn sui livelli e cold storage. UltraWarm	13 novembre 2024
OpenSearch supporto 2.17	Amazon OpenSearch Service ora supporta la OpenSearch versione 2.17. Questa versione include tutte le funzionalità che facevano parte delle versioni 2.16 e 2.17. Per ulteriori informazioni, consulta le note di rilascio 2.16 e 2.17.	13 novembre 2024

OpenSearch Supporto 2.15	Amazon OpenSearch Service ora supporta la OpenSearch versione 2.15. Questa versione include tutte le funzionalità che facevano parte delle versioni 2.14 e 2.15. Per ulteriori informazioni, consulta le note di rilascio 2.14 e 2.15.	11 ottobre 2024
Nuovi ruoli collegati al servizio	Amazon OpenSearch Service aggiunge un ruolo collegato al servizio chiamato`AWSServiceRoleForOpensearchIngestionSelfManagedVpce` , che consente ad Amazon OpenSearch Ingestion di inviare dati metrici a Amazon CloudWatch pipeline con endpoint VPC autogestiti.	12 giugno 2024
Interrogazioni dirette	Amazon OpenSearch Service ora supporta l'esecuzione di query direttamente sui dati archiviati in Amazon S3, senza dover inserire i dati in un indice. OpenSearch	22 maggio 2024

OpenSearch Supporto 2.13	Amazon OpenSearch Service ora supporta la OpenSearch versione 2.13. Questa versione include tutte le funzionalità che facevano parte delle versioni 2.12 e 2.13. Per ulteriori informazioni, consulta le note di rilascio 2.12 e 2.13.	21 maggio 2024
Supporto Amazon OpenSearch Ingestion per Data Prepper versione 2.7	Amazon OpenSearch Ingestion aggiunge il supporto per Data Prepper versione 2.7. Per ulteriori informazioni, consulta le note di rilascio 2.7.	4 aprile 2024
Servizio AWS accesso privato per le raccolte OpenSearch Serverless	Ora puoi concedere l'accesso specifico Servizi AWS, come Amazon Bedrock, alle tue raccolte OpenSearch Serverless all'interno di una politica di accesso alla rete.	28 marzo 2024
Aggiornamenti EBS in loco	Ora puoi apportare alcune modifiche EBS ai tuoi domini senza una distribuzione blu/verde in Amazon Service. OpenSearch	14 febbraio 2024
Visibilità delle modifiche alla configurazione	Ora puoi tenere traccia delle modifiche alla configurazione del dominio nella console OpenSearch di Amazon Service e utilizzando l'API di configurazione.	6 febbraio 2024

Disponibilità generale delle raccolte di ricerca vettoriale	Le raccolte di ricerca vettoriale e Amazon OpenSearch Serverless sono ora disponibili a livello generale. I seguenti miglioramenti importanti sono stati apportati durante la fase di anteprima: • Le raccolte di ricerca vettoriale ora supportano carichi di lavoro con miliardi di vettori, ciascuno con un massimo di 128 dimensioni. • OpenSearch Le dashboard ora supportano le raccolte di ricerca vettoriale.	29 novembre 2023
OR1 Istanze	Amazon OpenSearch Service ora supporta i tipi di OR1 istanze.	29 novembre 2023
Interrogazioni dirette con Amazon S3 (anteprima)	Le query dirette forniscono una soluzione completamente gestita per rendere disponibili i dati transazionali in Amazon OpenSearch Service entro pochi secondi dalla loro scrittura su un bucket Amazon S3.	29 novembre 2023

Capacità di 10 TiB per le raccolte di serie temporali	Amazon OpenSearch Serverless aggiunge il supporto per un massimo di 10 TiB di dati indicizzati per le raccolte di serie temporali . Questa versione supporta anche una capacità massima consentita di 200 OCUs per tutti i tipi di raccolte e la possibilità di disabilitare le repliche in standby quando si crea una raccolta.	29 novembre 2023
OpenSearch Supporto 2.11	Amazon OpenSearch Service ora supporta la OpenSearch versione 2.11. Questa versione include tutte le funzionalità che facevano parte delle versioni 2.10 e 2.11. Per ulteriori informazioni, consulta le note di rilascio 2.10 e 2.11.	17 novembre 2023
Supporto Amazon OpenSearch Ingestion per Data Prepper versione 2.6	Amazon OpenSearch Ingestion aggiunge il supporto per Data Prepper versione 2.6. Per ulteriori informazioni, consulta le note di rilascio 2.6. Inoltre, puoi specificare Amazon DynamoDB come origine della pipeline. Per ulteriori informazioni, consulta Usare una pipeline di OpenSearch ingestione con Amazon DynamoDB.	17 novembre 2023

Supporto Amazon OpenSearch Ingestion per Data Prepper versione 2.5	Amazon OpenSearch Ingestion aggiunge il supporto per Data Prepper versione 2.5. Per ulteriori informazioni, consulta le note di rilascio 2.5. Inoltre, ora puoi specificare un dominio di OpenSearch servizio o una raccolta OpenSearch Serverless come origine della pipeline. Per ulteriori informazioni, consulta il plug-in di OpenSearch origine nella documentazione di Data Prepper.	17 novembre 2023
CloudFormation modello per l'inferenza remota	Per facilitare la configurazione dell'inferenza remota per la ricerca semantica, Amazon OpenSearch Service fornisce un AWS CloudFormation modello nella console che automatizza il processo di provisioning del modello per te.	7 novembre 2023

Aggiornamento della politica relativa ai ruoli collegati ai servizi	Aggiunge le autorizzazioni necessarie alla politica dei ruoli collegati al servizio per `AmazonOpenSearchServiceRolePolicy` assegnare e annullare l'assegnazione degli indirizzi. IPv6 Anche la politica obsoleta di Elasticsearch è stata aggiornata per garantire la compatibilità con le versioni precedenti. `AmazonElasticsearchServiceRolePolicy`	26 ottobre 2023
Policy del ciclo di vita OpenSearch di Amazon Serverless	Amazon OpenSearch Serverless introduce le policy del ciclo di vita degli indici per semplificare la gestione della conservazione e dell'eliminazione dei dati. Ora puoi utilizzare un'interfaccia di configurazione nella console per impostare politiche di conservazione dei dati per le raccolte di serie temporali, eliminando la necessità di creare indici APIs o script giornalieri per eliminare i vecchi dati.	25 ottobre 2023

Supporto per istanze Im4gn	Amazon OpenSearch Service ora supporta i tipi di istanze Im4gn. Le istanze Im4gn sono ottimizzate per carichi di lavoro che gestiscono set di dati di grandi dimensioni e richiedon o un'elevata densità di storage per vCPU.	20 ottobre 2023
Opzioni amministrative	Amazon OpenSearch Service offre ora diverse opzioni amministrative che forniscon o un controllo granulare se devi risolvere problemi con il tuo dominio. Queste opzioni includono la possibili tà di riavviare il OpenSearch processo su un nodo di dati e la possibilità di riavviare un nodo di dati.	17 ottobre 2023
Plugin opzionali	Amazon OpenSearch Service aggiunge il supporto per quattro nuovi plugin di analisi linguistica: Nori (coreano) , Sudachi (giapponese), Pinyin (cinese) e STConvert Analysis (cinese), oltre al plug-in Amazon Personalize Search Ranking.	16 ottobre 2023

OpenSearch Supporto 2.9	Amazon OpenSearch Service ora supporta la OpenSearch versione 2.9. Questa versione include tutte le funzionalità che facevano parte delle versioni 2.8 e 2.9. Per ulteriori informazioni, consultate le note di rilascio 2.8 e 2.9.	2 ottobre 2023
Connettori ML	Amazon OpenSearch Service aggiunge il supporto per i connettori di machine learning (ML). I connettori facilitano l'accesso ai modelli ML ospitati su altre Servizi AWS piattaforme di machine learning (ML) o su piattaforme di apprendimento automatico (ML) di terze parti.	6 settembre 2023
Amazon OpenSearch Ingestion aggiunge il supporto per Data Prepper versione 2.4	Amazon OpenSearch Ingestion aggiunge il supporto per Data Prepper versione 2.4. Per ulteriori informazioni, consulta le note di rilascio 2.4. Inoltre, ora puoi specificare Amazon Streaming gestito per Apache Kafka (Amazon MSK) come origine della pipeline.	31 agosto 2023

Capacità di 6 TiB per le raccolte di serie temporali	Amazon OpenSearch Serverless aggiunge il supporto per un massimo di 6 TiB di dati indicizzati per le raccolte di serie temporali . Questa versione supporta anche una capacità massima consentita di 100 OCUs per le raccolte di ricerche e di serie temporali.	15 agosto 2023
Raccolte di ricerca vettoriale	Amazon OpenSearch Serverless aggiunge la possibilità di creare una raccolta di ricerche vettorial i, che puoi utilizzare per archiviare incorporamenti vettoriali per potenziare ricerche semantiche e di similarità.	26 luglio 2023
OpenSearch 2.7 supporto	Amazon OpenSearch Service ora supporta OpenSearch la versione 2.7. Questa versione include tutte le funzionalità che facevano parte delle versioni 2.6 e 2.7. Per ulteriori informazioni, consultate le note di rilascio 2.6 e 2.7.	10 luglio 2023

Supporto per Data Prepper 2.3	Amazon OpenSearch Ingestion aggiunge il supporto Data Prepper versione 2.3. Per ulteriori informazioni, consulta le note di rilascio 2.3. Inoltre, ora puoi specificare Amazon Security Lake come origine della pipeline.	26 giugno 2023
Multi-AZ con Standby	Amazon OpenSearch Service aggiunge la possibilità di distribuire un dominio in tre zone di disponibilità (AZs), con ogni AZ contenente una copia completa dei dati e con i nodi in una di queste AZs che fungono da standby. L'opzione di implementazione Multi-AZ con standby offre una disponibilità del 99,99% e prestazioni costanti in caso di guasto dell'infrastruttura.	3 maggio 2023
Nuovi ruoli collegati al servizio	Amazon OpenSearch Service aggiunge un ruolo collegato al servizio chiamato`AWSServiceRoleForAmazonOpenSearchIngestionService` , che consente ad Amazon OpenSearch Ingestion di inviare dati metrici a. Amazon CloudWatch	26 aprile 2023

OpenSearch Ingestione di Amazon	Amazon OpenSearch Ingestion è un raccoglitore di dati completamente gestito che fornisce dati di log and trace in tempo reale a domini di OpenSearch servizio e raccolte Serverless. OpenSearch OpenSearch Ingestion elimina la necessità di utilizzare soluzioni di terze parti come Logstash o Jaeger per inserire dati nei tuoi domini e nelle tue raccolte.	26 aprile 2023
OpenSearch 2.5 supporto	Amazon OpenSearch Service ora supporta OpenSearch la versione 2.5. Questa versione include tutte le funzionalità che facevano parte delle versioni 2.4 e 2.5. Per ulteriori informazioni, consultate le note di rilascio 2.4 e 2.5.	13 marzo 2023

Finestre di manutenzione non di punta	Amazon OpenSearch Service aggiunge finestre non di punta, ovvero blocchi di tempo giornalieri di 10 ore a basso traffico durante i quali può pianificare gli aggiornamenti del software di servizio e le ottimizzazioni Auto-Tune che richiedono un'implementazione . blue/green Gli aggiornamenti non di punta aiutano a ridurre al minimo il carico sui nodi master dedicati di un cluster durante i periodi di traffico più intenso. Per i nuovi domini creati dopo il 16 febbraio, la finestra non di punta viene configurata automaticamente tra le 22:00 e le 8:00 ora locale. Per i domini esistenti, devi abilitare esplicitamente la finestra.	16 febbraio 2023
Configurazione dell'autenticazione SAML durante la creazione del dominio	Amazon OpenSearch Service ora supporta la configurazione dell'autenticazione SAML durante la creazione del dominio. In precedenza, era necessario configurare le opzioni SAML dopo che il dominio era già stato creato.	1 febbraio 2023

Reindicizzazione remota per i domini VPC	Amazon OpenSearch Service aggiunge l'opzione per una connessione endpoint VPC tra due domini. È ora possibile utilizzare la reindicizzazione remota per copiare gli indici da un dominio VPC a un altro senza un proxy inverso. Per utilizzare questa funzionalità, i domini VPC devono eseguire il software di servizio R20221114 o versione successiva.	31 gennaio 2023

Disponibilità generale di Amazon OpenSearch Serverless	Amazon OpenSearch Serverless è ora disponibile a livello generale. I seguenti miglioramenti importanti sono stati apportati durante la fase di anteprima: • La capacità può ora essere ridotta al minimo configurato in OCUs caso di diminuzione del traffico sull'endpoint di raccolta. • Il valore massimo consentito sia OCUs per l'indicizzazione che per la ricerca è stato aumentato da 20 a 50. Ogni OCU include una memoria temporanea a caldo sufficiente per 120 GiB di dati di indice. • È ora possibile configurare le impostazioni di accesso ai dati durante la creazione delle raccolte anziché in un flusso di lavoro separato.	25 gennaio 2023
Esecuzione asincrona	Amazon OpenSearch Service ora supporta l'async dry run, che ti consente di eseguire un controllo di convalida prima di apportare una modifica alla configurazione e ti avvisa se le modifiche causeranno una distribuzione. blue/green	19 gennaio 2023

Nuovi ruoli collegati al servizio	Amazon OpenSearch Service aggiunge un ruolo collegato al servizio chiamato`AWSServiceRoleForAmazonOpenSearchServerless` , che consente a OpenSearch Serverless di inviare dati metrici a. Amazon CloudWatch	29 novembre 2022
Anteprima di Amazon OpenSearch Serverless	Amazon OpenSearch Serverless è una configurazione serverless su richiesta, con scalabilità automatica per Amazon Service. OpenSearch Serverless rimuove le complessità operative legate al provisioning, alla configurazione e all'ottimizzazione dei cluster. OpenSearch	29 novembre 2022
OpenSearch 2.3 supporto	Amazon OpenSearch Service ora supporta OpenSearch la versione 2.3. Questa versione include tutte le funzionalità che facevano parte delle versioni 2.0, 2.1 e 2.2. Per ulteriori informazioni, consulta le note di rilascio 2.0, 2.1, 2.2 e 2.3. La versione 2.3 contiene una modifica importante. Per ulteriori informazioni, consulta la sezione Percorsi di upgrade supportati.	15 novembre 2022

Supporto per il plug-in Notifications	Amazon OpenSearch Service ora supporta il plug-in Notifications, che offre una posizione centrale per tutte le notifiche provenienti dai OpenSearch plug-in. A partire dalla versione 2.0, le destinazioni degli avvisi sono diventate obsolete e sono state sostituite con canali di notifica.	15 novembre 2022
Supporto per Kibana 7.1.1	I domini Amazon OpenSearch Service che eseguono Elasticsearch 7.1 ora supportano l'ultima versione di patch per Kibana 7.1.1, che aggiunge correzioni di bug e migliora la sicurezza. Quando aggiorni i domini 7.1 al software di servizio R20221114, Service li aggiornerà automaticamente a questa versione di patch. OpenSearch	15 novembre 2022

Supporto per Kibana 6.8.13	I domini Amazon OpenSearch Service che eseguono Elasticsearch 6.8 ora supportano l'ultima versione di patch per Kibana 6.8.13, che aggiunge correzioni di bug e migliora la sicurezza. Quando aggiorni i domini 6.8 al software di servizio R20221114, Service li aggiornerà automaticamente a questa versione di patch. OpenSearch	15 novembre 2022
Supporto per Kibana 6.3.2	I domini Amazon OpenSearch Service che eseguono Elasticsearch 6.3 ora supportano l'ultima versione di patch per Kibana 6.3.2, che aggiunge correzioni di bug e migliora la sicurezza. Quando aggiorni i domini 6.3 al software di servizio R20221114, Service li aggiornerà automaticamente a questa versione di patch. OpenSearch	15 novembre 2022

AWS PrivateLink	Con gli endpoint OpenSearch VPC gestiti da Amazon Service, puoi connetterti direttamente ai domini Service OpenSearch VPC utilizzando un endpoint VPC di interfaccia anziché collegarti tramite Internet. Un endpoint OpenSearch VPC gestito dai servizi è accessibile solo all'interno del VPC in cui viene fornito l'endpoint o da qualsiasi dispositivo peerizzato VPCs con il VPC in cui viene fornito l'endpoint, come consentito dalle tabelle di routing e dai gruppi di sicurezza. Il dominio VPC deve eseguire il software di servizio R20220928 o versione successiva per connettersi a un endpoint VPC di interfaccia.	7 novembre 2022
Correzioni di bug e miglioramenti delle prestazioni	Il software del servizio R20220928 include correzioni di bug e miglioramenti delle prestazioni, tra cui una migliore registrazione SAML. L'aggiornamento modifica anche il tenant predefinito in `Global` anziché `Private`.	3 ottobre 2022

Riferimento alle API migliorato	Amazon OpenSearch Service offre un riferimento API di configurazione migliorato e completo. Questo nuovo riferimento contiene tutte le operazioni e i tipi di dati disponibili, esempi di sintassi di richieste e risposte e collegamenti ai riferimenti SDK corrispondenti per tutti i linguaggi supportati.	13 settembre 2022
Convalida blu/verde	Amazon OpenSearch Service ora esegue un controllo di convalida prima delle blue/green distribuzioni e rileva errori di convalida se il tuo dominio non è idoneo per un aggiornamento.	16 agosto 2022
OpenSearch 1.3 supporto	Amazon OpenSearch Service ora supporta OpenSearch la versione 1.3. Per ulteriori informazioni, consulta le Note di rilascio 1.3.	27 luglio 2022
Supporto per il plugin ML Commons	Amazon OpenSearch Service aggiunge il supporto per il plug-in ML Commons, che fornisce una serie di algoritmi di apprendimento automatic o comuni tramite il trasporto e le chiamate API REST. Puoi anche interagire con il plugin ML Commons tramite i comandi PPL.	27 luglio 2022

Supporto del volume gp3	Amazon OpenSearch Service aggiunge il supporto per il tipo di volume SSD gp3 EBS General Purpose. Puoi specificare capacità di IOPS allocata e velocità di trasmissione effettiva aggiuntive quando crei o modifichi il dominio.	26 luglio 2022
Documentazione delle best practice avanzate	La documentazione OpenSearch di Amazon Service fornisce best practice operative migliorate e consigli generali per la creazione e la gestione dei domini OpenSearch di servizio.	6 luglio 2022
Integrazione con Service Quotas	Ora puoi visualizzare le quote per Amazon OpenSearch Service e richiedere aumenti delle quote dalla console Service Quotas.	29 giugno 2022
Controllo degli accessi basato su tag per l'API OpenSearch	Ora puoi utilizzare i tag per controllare l'accesso a. OpenSearch APIs In precedenza, i tag potevano essere utilizzati solo per controllare l'accesso all'API di configurazione.	16 giugno 2022

Ricerca tra cluster tra regioni	La ricerca tra cluster è ora supportata Regioni AWS purché entrambi i domini eseguano Elasticsearch versione 7.10 o successiv a o qualsiasi versione di. OpenSearch	14 giugno 2022
Supporto per Kibana 5.6 singolo	Amazon OpenSearch Service aggiunge il supporto per Kibana 5.6.16 singolo. Con Kibana 5.6.16 singolo, puoi utilizzare Kibana 5.6 come front-end durante la connessione alle versioni di Elasticsearch 5.1, 5.3, 5.5 e 5.6. Per utilizzare Kibana 5.6 singolo, è necessari o utilizzare il software di servizio R20220323 o versioni successive.	4 aprile 2022
R20220323-P1	Amazon OpenSearch Service ha recentemente rilasciato l'aggiornamento del software di servizio R20220323, ma l'aggiornamento è stato successivamente annullato a causa di un problema. Si consiglia di aggiornare i domini alla versione di patch R20220323-P1 o versione successiva, il che risolve il problema.	4 aprile 2022

OpenSearch 1.2 supporto	Amazon OpenSearch Service ora supporta OpenSearch la versione 1.2. Per ulteriori informazioni, consultare le Note di rilascio 1.2.	4 aprile 2022
Osservabilità	L'installazione predefinita di OpenSearch Dashboards for Amazon OpenSearch Service include il plug-in Observability, che puoi utilizzare per visualizzare eventi basati sui dati utilizzando Piped Processing Language (PPL) per esplorare e interrogare i tuoi dati. Il plug-in richiede OpenSearch 1.2 o versione successiva e il software di servizio R20220323 o successivo.	4 aprile 2022
Supporto per Kibana 7.7.1	I domini Amazon OpenSearch Service che eseguono Elasticsearch 7.7 ora supportano l'ultima versione di patch per Kibana 7.7, che aggiunge correzioni di bug e migliora la sicurezza. Quando aggiorni i domini 7.7 al software di servizio R20220323 o versione successiva, Service li aggiornerà automaticamente a questa versione di patch. OpenSearch	4 aprile 2022

Modifica dei parametri della pressione della memoria JVM	Amazon OpenSearch Service ha modificato la logica delle `JVMMemoryPressure` CloudWatch metriche per riflettere in modo più accurato l'utilizzo della memoria. In precedenza i parametri consideravano solo il pool di memoria di vecchia generazione dell'heap della JVM. Con questa modifica il parametro considera anche il pool di memoria di nuova generazione. Dopo l'aggiornamento del dominio al software di servizio R20220323, è possibile che si verifichi un aumento della `JVMMemoryPressure` , `MasterJVMMemoryPressure` , e/o parametri `WarmJVMMemoryPressure` .	4 aprile 2022
Dizionari personalizzati con il plug-in IK (Chinese) Analysis	Amazon OpenSearch Service ora supporta l'utilizzo di dizionari personalizzati con il plug-in IK (cinese) Analysis.	4 aprile 2022

Replica tra cluster su domini esistenti	Amazon OpenSearch Service ha rimosso la limitazione secondo cui è possibile implementare la ricerca e la replica tra cluster solo su domini creati a partire dal 3 giugno 2020. Ora puoi abilitare queste caratteristiche su tutti i domini indipendentemente dal momento in cui sono state create. Per entrambi i domini è necessario il software di servizio R20220323 o versioni successive.	4 aprile 2022
Visibilità blue/verde dell'implementazione	Amazon OpenSearch Service ora offre una maggiore visibilità sullo stato di avanzamento delle implementazioni blu/green. Puoi monitorare questi dettagli nella console o utilizzando l'API di configurazione.	27 gennaio 2022

Controllo granulare degli accessi sui domini esistenti	Ora puoi abilitare il controllo granulare degli accessi sui domini esistenti. Puoi abilitare un periodo di migrazione temporaneo per le policy di accesso open/basate su per garantire che gli utenti possano continuar e ad accedere al dominio durante la creazione e la mappatura dei ruoli. L'abilita zione del controllo granulare degli accessi sui domini esistenti richiede il software di servizio R20211203 o versioni successive.	6 gennaio 2022
OpenSearch Ruoli Dashboard s rinominati	Con il software di servizio R20211203, il ruolo `kibana_user` è stato rinominato `opensearc h_dashboards_user` , e `kibana_read_only` è stato rinominato `opensearc h_dashboards_read_ only` . Questa modifica si applica a tutti i 1 appena OpenSearch creati. x domini. Per i OpenSearch domini esistenti che vengono aggiornati al software di servizio R20211203, i ruoli rimangono gli stessi.	4 gennaio 2022

OpenSearch 1.1 supporto	Amazon OpenSearch Service ora supporta OpenSearch la versione 1.1. Per ulteriori informazioni, consultare le Note di rilascio 1.1.	4 gennaio 2022
Editor visivo ISM	L'installazione predefinita di OpenSearch Dashboard s for Amazon OpenSearch Service ora supporta l'editor visivo per le politiche ISM. Questa funzionalità richiede la OpenSearch versione 1.1 o successiva.	4 gennaio 2022
Aggiornamento della prevenzione del "confused deputy" tra servizi	Amazon OpenSearch Service supporta l'utilizzo delle chiavi di contesto `aws:SourceArn` e delle condizioni `aws:SourceAccount` globali nelle politiche delle risorse IAM per prevenire il confuso problema del vice. Per utilizzare queste chiavi di condizione, è necessario utilizzare il software di servizio R20211203 o versioni successive.	4 gennaio 2022

Patch Log4j	Il software di servizio R20211203-P2 aggiorna la versione di Log4j utilizzata in OpenSearch Service come consigliato dagli avvisi contenuti in CVE-2021-44228 e CVE-2021-45046. La patch si applica ai domini che eseguono tutte le versioni di Elasticsearch. OpenSearch OpenSearch Il servizio continuerà ad aggiornare internamente varie versioni di Log4j e non saranno necessariamente limitate all'ultima versione di Log4j. La versione di Log4j sul tuo dominio dipende dalla versione del software su cui il dominio è in esecuzione. Tuttavia, indipendentemente dalla versione di Log4j, se utilizzi R20211203-P2 o versioni successive, i domini contengono l'aggiornamento di Log4j necessario per risolvere i problemi di CVE-2021-44228 e CVE-2021-45046.	15 dicembre 2021

Replica tra cluster	La replica tra cluster consente di replicare indici, mappature e metadati da un dominio di servizio a un altro. OpenSearch La replica tra cluster richiede un dominio che esegue Elasticsearch 7.10 o 1.1 o versione successiva. OpenSearch	5 ottobre 2021
AWS Nuove politiche gestite	Il lancio di Amazon OpenSearch Service include nuove politiche AWS gestite e l'eliminazione delle vecchie politiche.	08 settembre 2021
Supporto per Kibana 6.4.3	I domini Amazon OpenSearch Service che eseguono la versione precedente di Elasticsearch 6.4 ora supportano l'ultima versione di patch per Kibana 6.4, che aggiunge correzioni di bug e migliora la sicurezza . OpenSearch Il servizio aggiornerà automaticamente i domini a questa versione di patch.	08 settembre 2021

Flussi di dati	Amazon OpenSearch Service aggiunge il supporto per i flussi di dati, che semplific ano il processo di gestione dei dati delle serie temporali . Il tuo dominio deve avere la OpenSearch versione 1.0 o successiva per utilizzare i flussi di dati.	08 settembre 2021
OpenSearch Servizio Amazon	AWS rinomina Amazon OpenSearch Service per rimuovere il marchio precedent e «Elasticsearch». Amazon OpenSearch Service supporta un sistema operativo OpenSearch Elasticsearch legacy. Quando crei un cluster, puoi scegliere quale motore di ricerca utilizzar e. OpenSearch Il servizio offre un'ampia compatibilità con Elasticsearch OSS 7.10, l'ultima versione open source del software.	08 settembre 2021

Archiviazione a freddo	L'archiviazione a freddo è un nuovo livello di archiviazione per i dati storici o a cui si accede raramente. Gli indici a freddo occupano solo l'archiviazione S3 e non hanno alcun calcolo collegato. L'archiviazione a freddo richiede un dominio che esegue Elasticsearch 7.9 o versione successiva e il software di servizio R20210426 o versione successiva.	13 maggio 2021
Istanze Graviton basate su ARM	Amazon OpenSearch Service ora supporta i tipi di istanze Graviton basati su ARM (M6G, C6G, R6G e R6GD). I tipi di istanza Graviton sono disponibili su domini nuovi ed esistenti che eseguono Elasticsearch 7.9 o versioni successive e software di servizio R20210331 o versioni successive.	4 maggio 2021

Modelli ISM	Amazon OpenSearch Service aggiunge il supporto per i modelli ISM, che consenton o di allegare automatic amente una policy ISM a un indice se l'indice corrispon de a uno schema definito nella policy. I modelli ISM richiedono il software di servizio R20210426 o versioni successive. Questo aggiornam ento rende obsoleta anche l'impostazione `policy_id` , il che significa che non è più possibile utilizzare modelli di indice per applicare policy ISM agli indici appena creati. L'aggiornamento introduce una modifica sostanziale per i CloudFormation modelli esistenti che utilizzano questa impostazione.	27 aprile 2021
Supporto per Elasticsearch 7.10	Amazon OpenSearch Service ora supporta la versione 7.10 di Elasticsearch. Per ulteriori informazioni, consultare Note di rilascio 7.10.	21 aprile 2021

Ricerca asincrona	Amazon OpenSearch Service ora supporta la ricerca asincrona, che consente di eseguire richieste di ricerca in background. La ricerca asincrona richiede un dominio che esegue Elasticsearch 7.10 o versione successiv a e il software di servizio R20210331 o versione successiva.	21 aprile 2021
Controllo degli accessi basato su tag per API di configura zione	Ora puoi utilizzare i AWS tag per controllare l'accesso all'API di configurazione di Amazon ES.	2 marzo 2021
Regolazione automatica	Amazon OpenSearch Service aggiunge Auto-Tune, che utilizza i parametri di prestazio ni e utilizzo del cluster per suggerire modifiche alle impostazioni JVM sui nodi. La regolazione automatic a richiede un dominio che esegue Elasticsearch 6.7 o versione successiva e il software di servizio R20201117 o versione successiva.	24 febbraio 2021

Trace Analytics	L'installazione predefinita di Kibana per Amazon OpenSearch Service ora include il plug-in di analisi delle tracce, che consente di monitorare i dati di traccia dalle applicazioni distribuite. Il plug-in richiede un dominio che esegue Elasticsearch 7.9 o versione successiva e il software di servizio R20210201 o versione successiva.	17 febbraio 2021
Parametri delle partizioni	Amazon OpenSearch Service aggiunge le seguenti CloudWatch metriche per tracciare lo stato degli shard:`Shards.active` ,,, `Shards.unassigned` `Shards.delayedUnassigned` `Shards.activePrimary` ,`Shards.initializing` . `Shards.relocating` I parametri sono disponibili nei domini con software di servizio R20210201 o versioni successive.	17 febbraio 2021

Report di Kibana	L'installazione predefinita di Kibana per Amazon OpenSearch Service ora supporta report su richiesta per le pagine Discover, Visualize e Dashboard. Questa funzionalità richiede Elasticsearch 7.9 o versione successiva e il software di servizio R20210201 o versione successiva.	17 febbraio 2021
Supporto per Kibana 5.6.16	I domini Amazon OpenSearch Service che eseguono Elasticsearch 5.6 ora supportano l'ultima versione di patch per Kibana 5.6, che aggiunge correzioni di bug e migliora la sicurezza. Amazon ES aggiornerà automaticamente i domini a questa versione di patch.	17 febbraio 2021
Crittografia per domini esistenti	Amazon OpenSearch Service ora supporta l'abilitazione della crittografia dei dati inattivi e della node-to-node crittografia su domini esistenti che eseguono Elasticsearch 6.7 o versione successiva. Dopo aver abilitato queste impostazioni, non è possibile disabilitarle.	27 gennaio 2021

Reindicizzazione remota	Amazon OpenSearch Service ora supporta la reindiciz zazione remota, che consente di migrare gli indici da domini remoti. Questa funzional ità richiede il software del servizio R20201117 o versioni successive.	24 novembre 2020
Piped Processing Language (PPL)	Amazon OpenSearch Service ora supporta Piped Processing Language (PPL), un linguaggi o di query che consente di utilizzare la sintassi pipe (|) per interrogare i dati archiviat i in Elasticsearch. Questa funzionalità richiede il software del servizio R20201117 o versioni successive. Per ulteriori informazioni, consultar e.	24 novembre 2020
Notebook Kibana	Amazon OpenSearch Service aggiunge il supporto per i notebook Kibana, che consente di combinare visualizzazioni dal vivo e testo narrativo in un'unica interfaccia. Questa funzional ità richiede il software del servizio R20201117 o versioni successive.	24 novembre 2020

Grafici Gantt	L'installazione predefinita di Kibana per Amazon OpenSearch Service ora supporta un nuovo tipo di visualizzazione, i diagrammi di Gantt. Questa funzionalità richiede il software del servizio R20201117 o versioni successive.	24 novembre 2020
Supporto per Elasticsearch 7.9	Amazon OpenSearch Service ora supporta la versione 7.9 di Elasticsearch. Per ulteriori informazioni, consultare Note di rilascio 7.9.	24 novembre 2020
Aggiornamenti al rilevamento delle anomalie	Il rilevamento delle anomalie per Amazon OpenSearch Service aggiunge il supporto per la cardinalità elevata, che consente di classificare le anomalie con una dimensione come indirizzo IP, ID prodotto, codice paese e così via. Questa funzionalità richiede il software del servizio R20201117 o versioni successive.	24 novembre 2020

Aggiornamenti del dizionario dinamico	Amazon OpenSearch Service ora ti consente di aggiornar e i tuoi analizzatori di ricerca senza reindicizzarli. È possibile aggiornare i file del dizionario su alcuni o tutti i tuoi domini e Amazon ES tiene traccia delle versioni dei pacchetti nel tempo, in modo da avere una cronologi a di ciò che è cambiato e quando. Questa funzional ità richiede il software del servizio R20201019 o versioni successive.	17 novembre 2020
Endpoint personalizzati	Amazon OpenSearc h Service ora supporta endpoint personalizzati, che ti consentono di assegnare al tuo dominio Amazon ES un nuovo URL. Se si scambiano domini, è possibile mantenere lo stesso URL. Questa funzionalità richiede il software del servizio R20201019 o versioni successive.	05 novembre 2020

Nuovi plug-in del linguaggio	Amazon OpenSearch Service ora supporta i plug-in IK (Chinese) Analysis, Vietnames e Analysis e Thai Analysis su domini che eseguono Elasticsearch 7.7 o versione successiva con il software di servizio R20201019 o successivo.	28 ottobre 2020
Supporto per Elasticsearch 7.8	Amazon OpenSearch Service ora supporta la versione 7.8 di Elasticsearch. Per ulteriori informazioni, consultare Note di rilascio 7.8.	28 ottobre 2020
Autenticazione SAML per Kibana	Amazon OpenSearch Service ora supporta l'autenticazione SAML per Kibana, che consente di utilizzare provider di identità di terze parti per accedere a Kibana, gestire il controllo granulare degli accessi, cercare dati e creare visualizzazioni. Questa funzionalità richiede il software del servizio R20201019 o versioni successive.	27 ottobre 2020
Istanze T3	Amazon OpenSearch Service ora supporta i tipi di `t3.medium` istanze `t3.small` e.	23 settembre 2020

Log di verifica	Amazon OpenSearch Service ora supporta i log di controllo per i tuoi dati, che ti consenton o di tenere traccia dei tentativi di accesso falliti, dell'accesso degli utenti a indici, documenti e campi e molto altro. Questa funzionalità richiede il software del servizio R20200910 o versioni successive.	16 settembre 2020
UltraWarm aggiornamenti	UltraWarm per Amazon OpenSearch Service aggiunge nuove metriche, nuove impostazioni, una coda di migrazione più ampia e un'API di cancellazione. Questi aggiornamenti richiedon o il software di servizio R20200910 o versione successiva. Per ulteriori informazioni, consultare .	14 settembre 2020
Learning to Rank	Amazon OpenSearch Service ora supporta il plug-in open source Learning to Rank, che consente di utilizzare tecnologi e di apprendimento automatic o per migliorare la pertinenz a delle ricerche. Questa funzionalità richiede il software del servizio R20200721 o versioni successive.	27 luglio 2020

Similitudine del coseno k-NN	k-Nearest Neighbor (k-NN) ora ti permette di cercare i "vicini più vicini" per similitudine del coseno oltre alla distanza euclidea. Questa funzionalità richiede il software del servizio R20200721 o versioni successive.	23 luglio 2020
Compressione gzip	Amazon OpenSearch Service ora supporta la compressione gzip per la maggior parte delle richieste e risposte HTTP, che può ridurre la latenza e conservare la larghezza di banda. Questa funzionalità richiede il software del servizio R20200721 o versioni successive.	23 luglio 2020
Supporto per Elasticsearch 7.7	Amazon OpenSearch Service ora supporta la versione 7.7 di Elasticsearch. Per ulteriori informazioni, consultare Note di rilascio 7.7.	23 luglio 2020
Servizio di mappa di Kibana	L'installazione predefinita di Kibana per Amazon OpenSearch Service ora include un server di mappe WMS, ad eccezione dei domini nelle regioni di India e Cina.	18 giugno 2020

Miglioramenti per SQL	Il supporto SQL per Amazon OpenSearch Service ora supporta molte nuove operazioni, un'interfaccia utente Kibana dedicata per l'esplorazione dei dati e una CLI interattiva.	3 giugno 2020
Funzionalità di ricerca tra cluster	Amazon OpenSearch Service ti consente di eseguire query e aggregazioni tra cluster su più domini connessi.	3 giugno 2020
Rilevamento di anomalie	Amazon OpenSearch Service ti consente di rilevare automaticamente le anomalie quasi in tempo reale.	3 giugno 2020
UltraWarm	UltraWarm lo storage per Amazon OpenSearch Service non è più disponibile in anteprima pubblica ed è ora disponibile a livello generale. La funzionalità ora supporta una gamma più ampia di versioni e Regioni AWS. Per ulteriori informazioni, consulta .	05 maggio 2020

Dizionari personalizzati	Amazon OpenSearch Service ti consente di caricare file di dizionario personalizzati da utilizzare con il tuo cluster. Questi file migliorano i risultati della ricerca dicendo a Elasticsearch di ignorare determinate parole ad alta frequenza o di trattare termini come equivalenti.	21 aprile 2020
Supporto per Elasticsearch 7.4	Amazon OpenSearch Service ora supporta la versione 7.4 di Elasticsearch. Per ulteriori informazioni, consultare Versioni supportate.	12 marzo 2020
k-NN	Amazon OpenSearch Service aggiunge il supporto per la ricerca K-Nearest Neighbor (k-NN). k-NN richiede il software di servizio R20200302 o successivo.	3 marzo 2020
Index State Management	Amazon OpenSearch Service aggiunge Index State Management (ISM), che consente di automatizzare le attività di routine, come l'eliminazione degli indici quando raggiungono una certa età. Questa funzionalità richiede il software del servizio R20200302 o versioni successive.	3 marzo 2020

Supporto per Elasticsearch 5.6.16	Amazon OpenSearch Service ora supporta l'ultima versione di patch per la versione 5.6, che aggiunge correzioni di bug e migliora la sicurezza . Amazon ES aggiornerà automaticamente i domini versione 5.6 esistenti a questa versione. Questa versione di Elasticsearch riporta erroneamente la versione come 5.6.17.	2 marzo 2020
Controllo granulare degli accessi	Amazon OpenSearch Service ora supporta il controllo granulare degli accessi, che offre sicurezza a livello di indice, documento e campo, la multi-tenancy Kibana e l'autenticazione di base HTTP opzionale per il cluster.	11 febbraio 2020
UltraWarm archiviazione (anteprima)	Amazon OpenSearch Service aggiunge UltraWarm un nuovo livello di storage caldo che utilizza Amazon S3 e una sofisticata soluzione di caching per migliorare le prestazio ni. Per gli indici su cui non si scrive attivamente e che vengono interrogati con minore frequenza, UltraWarm lo storage offre costi per GiB notevolmente inferiori.	3 dicembre 2019

Funzionalità di crittografia per le regioni Cina	La crittografia dei dati archiviati e node-to-node la crittografia sono ora disponibili nella regione `cn-north-1` Cina (Pechino) e nella regione `cn-northwest-1` Cina (Ningxia).	20 novembre 2019
Richiedi HTTPS	Adesso è possibile richiedere che tutto il traffico verso i domini Amazon ES arrivi tramite HTTPS. Durante la configurazione del dominio, seleziona la casella Require HTTPS (Richiedi HTTPS). Questa funzionalità richiede il software del servizio R20190808 o versioni successive.	3 ottobre 2019
Supporto per Elasticsearch 7.1 e 6.8	Amazon OpenSearch Service ora supporta le versioni 7.1 e 6.8 di Elasticsearch. Per ulteriori informazioni, consultare Versioni supportate.	13 agosto 2019
Snapshot orari	Invece di istantanee giornaliere, Amazon OpenSearch Service ora acquisisce istantanee orarie dei domini che eseguono Elasticsearch 5.3 e versioni successive, in modo da avere backup più frequenti da cui ripristinare i dati.	8 luglio 2019

Supporto per Elasticsearch 6.7	Amazon OpenSearch Service ora supporta la versione 6.7 di Elasticsearch. Per ulteriori informazioni, consultare Versioni supportate.	29 maggio 2019
Supporto per SQL	Amazon OpenSearch Service ora ti consente di interroga re i tuoi dati utilizzando SQL. Il supporto SQL richiede il software del servizio R20190418 o versioni successive.	15 maggio 2019
Tipi di istanza 5-series	Amazon OpenSearch Service ora supporta i tipi di istanze M5, C5 e R5. Rispetto ai tipi di istanze di vecchia generazio ne, questi nuovi tipi offrono prestazioni migliori a prezzi inferiori. Per ulteriori informazi oni, consulta Limiti.	24 aprile 2019
Supporto per Elasticsearch 6.5	Amazon OpenSearch Service ora supporta la versione 6.5 di Elasticsearch.	8 Aprile 2019
Avviso	Gli avvisi per Amazon OpenSearch Service ti avvisano quando i dati di uno o più indici Amazon ES soddisfano determinate condizioni. La funzionalità di avviso richiede il software del servizio R20190221 o versioni successive.	25 marzo 2019

Supporto per tre zone di disponibilità	Amazon OpenSearch Service ora supporta tre zone di disponibilità in molte regioni. Questa versione include anche un'esperienza della console ottimizzata. Questo multi-AZ richiede il software del servizio R20181023 o versioni successive.	7 febbraio 2019
Supporto per Elasticsearch 6.4	Amazon OpenSearch Service ora supporta la versione 6.4 di Elasticsearch.	23 gennaio 2019
Cluster di 200 nodi	Amazon ES ora consente di creare cluster con fino a 200 nodi di dati per un totale di 3 PB di archiviazione.	22 gennaio 2019
Aggiornamenti del software del servizio	Amazon ES ora consente di aggiornare manualmente il software di servizio per il dominio al fine di trarre vantaggio dalle nuove funzionalità più rapidamente o aggiornarlo in un momento con poco traffico. Per ulteriori informazioni, consultare.	20 novembre 2018
CloudWatch Nuove metriche	Amazon ES ora offre i parametri a livello di nodo e le nuove schede Integrità del cluster e Integrità dell'istanza nella console Amazon ES.	20 novembre 2018

Supporto per Cina (Pechino)	Amazon OpenSearch Service è ora disponibile nella regione cn-north-1, dove supporta i tipi di istanze M4, C4 e R4.	17 ottobre 2018
Node-to-node crittografia	Amazon OpenSearch Service ora supporta node-to-node la crittografia, che mantiene i dati crittografati mentre Amazon ES li distribuisce in tutto il cluster.	18 settembre 2018
Aggiornamenti locali della versione	Amazon OpenSearch Service ora supporta gli upgrade di versione in loco.	14 agosto 2018
Supporto per Elasticsearch 6.3 e 5.6	Amazon OpenSearch Service ora supporta le versioni 6.3 e 5.6 di Elasticsearch.	14 agosto 2018
Log di errore	Amazon ES ora consente di pubblicare i log degli errori di Elasticsearch su Amazon. CloudWatch	31 luglio 2018
Istanze riservate della regione Cina (Ningxia)	Amazon ES ora offre istanze riservate nella regione Cina (Ningxia).	29 maggio 2018
Istanze riservate	Amazon ES ora offre il supporto per le istanze riservate.	7 maggio 2018

Aggiornamenti precedenti

Nella tabella seguente sono riportate le modifiche importanti apportate ad Amazon ES prima di maggio 2018.

Modifica	Descrizione	Data
Autenticazione di Amazon Cognito per Kibana	Amazon ES offre ora protezione della pagina di accesso per Kibana. Per ulteriori informazioni, consultare the section called “Autenticazione Amazon Cognito per dashboard OpenSearch ”.	2 aprile 2018
Supporto per Elasticsearch 6.2	Amazon OpenSearch Service ora supporta la versione 6.2 di Elasticsearch.	14 marzo 2018
Plug-in di analisi coreano	Amazon ES supporta ora una versione ottimizzata per la memoria del plug-in di analisi coreano Seunjeon.	13 marzo 2018
Aggiornamenti istantanei per il controllo degli accessi	Le modifiche apportate alle policy di controllo degli accessi nei domini Amazon ES ora hanno effetto immediatamente.	07 marzo 2018
Scala in petabyte	Amazon ES supporta ora i tipi di istanza I3 e dimension i di archiviazione totali del dominio fino a 1,5 PB. Per ulteriori informazioni, consultare the section called “Scala in petabyte”.	19 dicembre 2017
Crittografia dei dati a riposo	Amazon ES supporta ora la crittografia dei dati a riposo. Per ulteriori informazioni, consultare the section called “Crittografia dei dati a riposo”.	7 dicembre 2017
Supporto per Elasticsearch 6.0	Amazon ES ora supporta Elasticsearch versione 6.0. Per le considerazioni e le istruzioni relative alla migrazione, consulta the section called “Aggiornamento dei domini”.	6 dicembre 2017
Supporto per VPC	Amazon ES consente ora di avviare domini all'interno di Amazon Virtual Private Cloud. il supporto per VPC offre un ulteriore livello di sicurezza e semplifica le comunicaz ioni tra Amazon ES e altri servizi all'interno di un VPC. Per ulteriori informazioni, consultare the section called “Supporto per VPC”.	17 ottobre 2017

Modifica	Descrizione	Data
Pubblicazione di log di query lente	Amazon ES ora supporta la pubblicazione di slow log su Logs. CloudWatch Per ulteriori informazioni, consultare the section called “Monitoraggio dei log”.	16 ottobre 2017
Supporto per Elasticsearch 5.5	Amazon ES ora supporta Elasticsearch versione 5.5. Adesso è possibile ripristinare snapshot automatici senza contattare Supporto e archiviare gli script tramite l'API `_scripts`.	07 settembre 2017
Supporto per Elasticsearch 5.3	Amazon ES ha aggiunto il supporto per Elasticsearch versione 5.3.	1 giugno 2017
Più istanze e capacità di EBS per ogni cluster	Amazon ES supporta ora fino a 100 nodi e 150 TB di capacità di EBS per ogni cluster.	5 Aprile 2017
Supporto per le regioni Canada (Centrale) e UE (Londra)	Amazon ES ha aggiunto il supporto per le seguenti regioni: Canada (Centrale), ca-central-1, ed Europa (Londra), eu-west-2.	20 marzo 2017
Più istanze e volumi EBS di dimensioni maggiori	Amazon ES ha aggiunto il supporto per più istanze e volumi EBS di dimensioni maggiori.	21 febbraio 2017
Supporto per Elasticsearch 5.1	Amazon ES ha aggiunto il supporto per Elasticsearch versione 5.1.	30 gennaio 2017
Supporto per il plug-in di analisi fonetica	Amazon ES fornisce ora l'integrazione predefinita con il plug-in di analisi fonetica, che permette di eseguire query basate sulla fonetica sui dati.	22 dicembre 2016
Supporto per la regione Stati Uniti orientali (Ohio)	Amazon ES ha aggiunto il supporto per la regione Stati Uniti orientali (Ohio), us-east-2.	17 ottobre 2016

Modifica	Descrizione	Data
Nuovo parametro per le prestazioni	Amazon ES ha aggiunto un parametro per le prestazioni, `ClusterUsedSpace` .	29 luglio 2016
Supporto per Elasticsearch 2.3	Amazon ES ha aggiunto il supporto per Elasticsearch versione 2.3.	27 luglio 2016
Supporto per la regione Asia Pacifico (Mumbai)	Amazon ES ha aggiunto il supporto per la seguente regione: Asia Pacifico (Mumbai), ap-south-1.	27 giugno 2016
Più istanze per cluster	Amazon ES ha aumentato il numero massimo di istanze (numero di istanze) per cluster da 10 a 20.	18 maggio 2016
Supporto per la regione Asia Pacifico (Seoul)	Amazon ES ha aggiunto il supporto per la seguente regione: Asia Pacifico (Seoul), ap-northeast-2.	28 gennaio 2016
Amazon ES	Versione iniziale.	1 Ottobre 2015

AWS Glossario

Per la AWS terminologia più recente, consultate il AWS glossario nella sezione Reference.Glossario AWS

Le traduzioni sono generate tramite traduzione automatica. In caso di conflitto tra il contenuto di una traduzione e la versione originale in Inglese, quest'ultima prevarrà.