aws

Guía para desarrolladores

OpenSearch Servicio Amazon



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

OpenSearch Servicio Amazon: Guía para desarrolladores

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

Las marcas comerciales y la imagen comercial de Amazon no se pueden utilizar en relación con ningún producto o servicio que no sea de Amazon de ninguna manera que pueda causar confusión entre los clientes y que menosprecie o desacredite a Amazon. Todas las demás marcas registradas que no son propiedad de Amazon son propiedad de sus respectivos propietarios, que pueden o no estar afiliados, conectados o patrocinados por Amazon.

Table of Contents

¿Qué es Amazon OpenSearch Service? 1
Características de Amazon OpenSearch Service 2
Cuándo se debe usar 3
Versiones compatibles de Elasticsearch y OpenSearch 4
Soporte extendido y estándar 4
Versiones estándar compatibles y ampliadas 7
Cálculo de los cargos por soporte extendido 7
Precios 9
Servicios relacionados 9
Configuración 12
Concesión de permisos 12
Concesión de acceso programático 12
Configure el AWS CLI 14
Abra la consola de 15
Introducción 16
Crear un dominio 17
Cargar datos para indexar 18
Opción 1: cargar un solo documento 19
Opción 2: cargar varios documentos 19
Buscar documentos 20
Para buscar documentos desde la línea de comandos 20
Busque documentos mediante paneles OpenSearch 21
Eliminar un dominio 22
OpenSearch Ingestión de Amazon 23
Conceptos clave 23
Ventajas 25
Limitaciones 26
Versiones de Data Prepper admitidas 26
Escala de canalizaciones 27
Precios 29
Soportado Regiones de AWS 29
Configuración de roles y usuarios 29
Rol de canalización 31
Rol de ingesta 34

Concesión de acceso a dominios para canalizaciones 36
Concesión de acceso a colecciones para canalizaciones 40
Cómo empezar con Ingestion OpenSearch 46
Tutorial: incorporar datos a un dominio 46
Tutorial: incorporar datos en una colección 55
Características de la canalización 62
Almacenamiento en búfer persistente 62
División 65
Encadenar 66
Colas de mensajes fallidos 67
Administración de índices 68
End-to-end reconocimiento 72
Contrapresión del origen 73
Creación de canalizaciones 74
Requisitos previos y rol de IAM requerido 74
Permisos de IAM necesarios 75
Especificar la versión de la canalización 77
Especificación de la ruta de ingesta 78
Creación de canalizaciones 78
Seguimiento del estado de creación de la canalización 83
Uso de esquemas 85
Visualización de canalizaciones 86
Actualización de las canalizaciones 89
Consideraciones 89
Permisos necesarios 90
Actualización de las canalizaciones 91
Implementaciones azul/verde para actualizaciones de canalización 92
Administración de los costos de las canalizaciones 93
Detención de una canalización 94
Inicio de una canalización 95
Eliminación de canalizaciones 96
Complementos y opciones compatibles 96
Complementos compatibles 97
Procesadores sin estado frente a procesadores con estado 99
Requisitos y restricciones de configuración 100
Integración de canalizaciones 105

Creación del punto de conexión de ingesta 105
Creación de un rol de ingesta 106
Servicios de Atlassian 108
Amazon Aurora 126
Amazon DynamoDB 156
Amazon DocumentDB 172
Confluent Cloud Kafka 191
Amazon MSK 201
Amazon RDS 209
Amazon S3 238
Amazon Security Lake 248
Fluent Bit 253
Fluentd 255
Inferencia en lotes de ML sin conexión 257
OpenTelemetry Colector 274
Servicio administrado por Amazon para Prometheus 280
Kafka autogestionado 289
Clústeres autogestionados OpenSearch 297
Amazon Kinesis Data Streams 305
Siguientes pasos 317
AWS Lambda 318
Migración de datos entre dominios y colecciones 321
Limitaciones 322
OpenSearch El servicio como fuente 322
Especificar varios sumideros OpenSearch de dominio de servicio 325
Migración de datos a una colección de OpenSearch VPC sin servidor 325
Administrar canalizaciones con AWS SDKs 326
Python 326
Seguridad en la ingestión OpenSearch 330
Configuración del acceso a la VPC para canalizaciones 331
Configuración de canalizaciones para la ingesta entre cuentas 336
Gestión de identidad y acceso 344
Supervisión con CloudTrail 352
Puntos de conexión de VPC (AWS PrivateLink) 356
Etiquetado de canalizaciones 359
Permisos necesarios 360

Uso de etiquetas (consola) ... 360
Uso de etiquetas (AWS CLI) ... 361
Registro y supervisión ... 361
Monitoreo de registros de canalización ... 361
Monitoreo de métricas de canalización ... 364
Prácticas recomendadas ... 396
Prácticas recomendadas generales ... 397
CloudWatch Alarmas recomendadas ... 397
Amazon OpenSearch Serverless ... 404
Ventajas ... 404
¿Qué es Amazon OpenSearch Serverless? ... 405
Casos de uso de Serverless OpenSearch ... 406
Funcionamiento ... 406
Elección de un tipo de colección ... 408
Precios ... 409
Compatible Regiones de AWS ... 410
Limitaciones ... 410
Comparación entre OpenSearch servicio y sin servidor OpenSearch ... 411
Tutorial: Cómo empezar a usar Serverless OpenSearch ... 415
Paso 1: configurar permisos ... 416
Paso 2: crear una colección ... 417
Paso 3: cargar y buscar datos ... 418
Gestión de errores ... 419
Paso 4: eliminar la colección ... 420
Siguientes pasos ... 420
Colecciones ... 421
Administración de colecciones ... 421
Trabajo con colecciones de búsqueda vectorial ... 445
Uso de políticas de ciclo de vida de los datos ... 453
Administrar las colecciones con AWS SDKs ... 462
Crear colecciones con CloudFormation ... 474
Hacer copias de seguridad de las colecciones mediante instantáneas ... 476
Soporte de códecs Zstandard en Amazon Serverless OpenSearch ... 483
Ahorre almacenamiento mediante el uso de una fuente derivada ... 485
Grupos de colecciones ... 487
Conceptos clave ... 488

Límites de capacidad de los grupos de recolección 488
Claves de cifrado y KMS en grupos de recopilación 492
Crea grupos de colecciones 493
Administre los grupos de recopilación 496
Administración de los límites de capacidad 499
Como establecer los parámetros de capacidad 501
Límites de la capacidad máxima 502
Monitoreo del uso de la capacidad 503
Ingesta de datos en las colecciones 503
Permisos mínimos necesarios 504
OpenSearch Ingestión 504
Fluent Bit 505
Amazon Data Firehose 506
Go 506
Java 508
JavaScript 510
Logstash 512
Python 515
Ruby 516
Otros clientes 517
Configurar Machine Learning 519
Machine Learning 519
Connectors 520
Modelos 520
Configurar permisos para Machine Learning 521
Funciones y características no compatibles APIs 522
Configure la búsqueda neuronal y la búsqueda híbrida 523
Búsqueda neuronal 523
Búsqueda híbrida 524
Consultas neuronales e híbridas 524
Configuración de permisos 526
Configure los flujos de trabajo 528
Flujos de trabajo 528
Configuración de permisos 528
Seguridad en Serverless OpenSearch 530
Políticas de cifrado 532

Políticas de red ... 532
Políticas de acceso a datos ... 533
Autenticación SAML e IAM ... 534
Seguridad de la infraestructura ... 535
Introducción a la seguridad ... 535
Gestión de identidad y acceso ... 550
Cifrado ... 568
Acceso a la red ... 580
Conformidad con FIPS ... 591
Control de acceso a los datos ... 597
Punto final de VPC del plano de datos ... 608
Punto final de VPC del plano de control ... 620
Autenticación SAML ... 623
Validación de conformidad ... 633
Etiquetado de colecciones ... 633
Permisos necesarios ... 634
Etiquetado de colecciones (consola) ... 634
Etiquetado de colecciones (AWS CLI) ... 635
Operaciones y complementos compatibles ... 635
Operaciones y permisos de OpenSearch API compatibles ... 635
OpenSearch Plugins compatibles ... 645
Supervisión sin servidor OpenSearch ... 646
Monitorización con CloudWatch ... 647
Monitorización con CloudTrail ... 654
Monitorización con EventBridge ... 659
Creación y administración de dominios ... 663
Creación de dominios OpenSearch de servicio ... 663
Creación OpenSearch de dominios de servicio (consola) ... 663
Crear dominios de OpenSearch servicio ()AWS CLI ... 670
Creación OpenSearch de dominios de servicio (AWS SDKs) ... 672
Creación OpenSearch de dominios de servicio (AWS CloudFormation) ... 672
Configurar políticas de acceso ... 673
Migración al OpenSearch servicio con Migration Assistant ... 673
Descripción general de ... 673
Escenarios de migración ... 674
Introducción del Asistente de Migración ... 675

Arquitectura de soluciones 676
Configuración avanzada de clústeres 676
Enriquecimiento semántico automático 677
Introducción 677
Detalles del modelo y punto de referencia de rendimiento 678
Idiomas compatibles 679
Configure un índice de enriquecimiento semántico automático para los dominios 679
Ingestión y búsqueda de datos 681
Configuración de permisos para el enriquecimiento semántico automático 682
Reescrituras de consultas 685
Limitaciones del enriquecimiento semántico automático 686
Precios 686
Cambios de configuración 687
Cambios que suelen provocar blue/green despliegues 687
Cambios que normalmente no provocan blue/green despliegues 688
Opciones de despliegue azul/verde 689
Determinar si un cambio provocará un blue/green despliegue 692
Seguimiento de un cambio de configuración 696
Etapas de un cambio de configuración 698
Impacto de las blue/green implementaciones en el rendimiento 701
Cargos por los cambios de configuración 702
Solución de errores de validación 702
Actualizaciones del software del servicio 708
Actualizaciones opcionales frente a actualizaciones obligatorias 709
Actualizaciones de revisión 710
Consideraciones 710
Iniciar una actualización 711
Intervalos de menor demanda 714
Supervisar actualizaciones 715
Cuando los dominios no son aptos para una actualización 716
Intervalos de menor demanda 717
Actualizaciones de software de servicio de menor demanda 718
Optimizaciones de ajuste automático durante horas de menor demanda 719
Habilitación del intervalo de menor demanda 719
Configuración de un intervalo de menor demanda personalizado 720
Visualización de acciones programadas 721

Reprogramación de acciones ... 723
Migración desde los intervalos de mantenimiento de ajuste automático ... 724
Notificaciones ... 725
Introducción a las notificaciones ... 726
Severidad de las notificaciones ... 727
Ejemplo de evento EventBridge ... 728
Configuración de un dominio multi-AZ ... 728
Multi-AZ con modo de espera ... 729
Multi-AZ sin modo de espera ... 730
Interrupción de las zonas de disponibilidad ... 735
Compatibilidad con VPC ... 737
VPC frente a dominios públicos ... 737
Limitaciones ... 738
Arquitectura ... 738
Crear instantáneas de índice ... 746
Requisitos previos ... 747
Registrar un repositorio de instantáneas manuales ... 751
Tomar instantáneas manuales ... 755
Restaurar instantáneas ... 757
Eliminar instantáneas manuales ... 760
Automatizar instantáneas con la administración de instantáneas ... 760
Automatizar instantáneas con la administración de estado de índice ... 763
Utilizar Curator para instantáneas ... 763
Actualización de dominios ... 764
Rutas de actualización admitidas ... 764
Actualización de un dominio (consola) ... 769
Actualización de un dominio (CLI) ... 769
Actualización de un dominio (SDK) ... 770
Solución de errores de validación ... 771
Solución de problemas de una actualización ... 772
Uso de una instantánea para migrar datos ... 775
Creación de un punto de conexión personalizado ... 783
Puntos de conexión personalizados para nuevos dominios ... 783
Puntos de conexión personalizados para dominios existentes ... 784
Asignación CNAME ... 784
Ajuste automático ... 785

Tipos de cambios ... 785
Habilitar o deshabilitar la función de ajuste automático ... 787
Programación de mejoras de ajuste automático ... 788
Supervisión de cambios de ajuste automático ... 789
Etiquetado de dominios ... 789
Ejemplos de etiquetas ... 790
Etiquetado de dominios (consola) ... 791
Etiquetado de dominios (AWS CLI) ... 791
Etiquetado de dominios (AWS SDKs) ... 793
Ejecución de acciones administrativas ... 794
Reiniciar el OpenSearch proceso en un nodo de datos ... 795
Reinicio de un nodo de datos ... 795
Reiniciar el proceso de Dashboards ... 796
Limitaciones ... 797
Uso de consultas directas ... 798
Precios ... 798
Limitaciones ... 799
Limitaciones generales ... 799
Limitaciones de Amazon S3 ... 799
Limitaciones de Amazon CloudWatch Logs ... 800
Limitaciones de Amazon Security Lake ... 800
Recomendaciones ... 801
Recomendaciones generales ... 801
Recomendaciones para Amazon S3 ... 802
CloudWatch Registra recomendaciones ... 802
Recomendaciones para Security Lake ... 803
Cuotas ... 804
Cuotas para Amazon S3 ... 804
Cuotas de CloudWatch registros ... 805
Cuotas para Security Lake ... 806
Compatible Regiones de AWS ... 807
Disponible Regiones de AWS para Amazon S3 ... 808
Disponible Regiones de AWS para CloudWatch registros ... 808
Disponible Regiones de AWS para Security Lake ... 809
Consultas directas en S3 ... 809
Creación de un origen de datos de S3 ... 810

Configuración de un origen de datos de S3 ... 818
Consultas directas en Logs CloudWatch ... 821
Crear una fuente de datos de CloudWatch Logs ... 821
Configuración de una fuente de datos de CloudWatch Logs ... 828
Consultas directas en Security Lake ... 830
Creación de un origen de datos de Security Lake ... 831
Configuración de un origen de datos de Security Lake ... 837
Administración de un origen de datos ... 840
Supervisión con CloudWatch métricas y fuentes de datos ... 840
Activación y desactivación de orígenes de datos ... 844
Supervisión con AWS presupuesto ... 844
Eliminación de un origen de datos ... 845
Optimización del rendimiento de las consultas ... 846
Índices de omisión ... 846
Vistas materializadas ... 847
Índices de cobertura ... 847
Comandos SQL y PPL compatibles ... 847
Comandos SQL compatibles ... 849
Comandos compatibles PPL ... 1057
Supervisión de dominios ... 1242
Monitoreo de métricas del clúster ... 1243
Visualización de las métricas en CloudWatch ... 1244
Interpretación de las historias clínicas en OpenSearch Service ... 1245
Métricas de clúster ... 1245
Métricas de nodo maestro dedicado ... 1254
Métricas de nodos coordinadores dedicados ... 1255
Métricas de volumen de EBS ... 1256
Métricas de la instancia ... 1259
Métricas cálidas ... 1273
Métricas de almacenamiento en frío ... 1280
OpenSearch Métricas de instancias optimizadas (OR1) ... 1282
Métricas de alertas ... 1282
Métricas de detección de anomalías ... 1284
Métricas de búsqueda asíncrona ... 1286
Métricas de ajuste automático ... 1288
Métricas de Multi-AZ con modo de espera ... 1288

Métricas de un momento dado ... 1291
Métricas de SQL ... 1292
Métricas k-NN ... 1293
Métricas de búsqueda entre clústeres ... 1296
Métricas de replicación entre clústeres ... 1297
Aprender a clasificar métricas ... 1299
Métricas del lenguaje de procesamiento de canalizaciones ... 1300
Monitoreo de registros ... 1300
Habilitación de la publicación de registros (consola) ... 1302
Habilitación de la publicación de registros (AWS CLI) ... 1304
Habilitación de la publicación de registros (AWS SDKs) ... 1306
Habilitación de la publicación de registros (CloudFormation) ... 1307
Establecimiento de umbrales de registros lentos de solicitudes de búsqueda ... 1309
Establecimiento de umbrales de registros lentos de particiones ... 1310
Prueba de registros lentos ... 1310
Visualización de registros ... 1311
Monitoreo de registros de auditoría ... 1311
Limitaciones ... 1312
Habilitación de los registros de auditoría ... 1312
Habilite el registro de auditorías mediante el AWS CLI ... 1314
Habilite el registro de auditoría con la API de configuración ... 1315
Categorías y capas de registro de auditoría ... 1315
Configuración de registros de auditoría ... 1318
Ejemplo de registro de auditoría ... 1321
Configuración de registros de auditoría mediante la API REST ... 1324
Supervisión de eventos ... 1325
Eventos de actualización del software de servicio ... 1326
Eventos de ajuste automático ... 1333
Eventos del estado del clúster ... 1338
Eventos del punto de conexión de VPC ... 1352
Eventos de retirada de nodos ... 1354
Eventos de retirada de nodos degradados ... 1356
Eventos de error de dominio ... 1358
Tutorial: Cómo escuchar los eventos OpenSearch de Service ... 1361
Tutorial: Sending SNS alerts for available updates ... 1363
Monitorear con CloudTrail ... 1364

Información OpenSearch de Amazon Service en CloudTrail 654
Descripción de las entradas de los archivos de registro de Amazon OpenSearch Service ... 658
Información sobre clústeres 1369
Ventajas 1369
Acceda a Cluster Insights a través de la consola 1370
Acceda a las estadísticas y métricas detalladas del clúster a través OpenSearch de la interfaz de usuario 1370
Cree y configure una aplicación para ver Cluster Insights 1371
Comprensión de las perspectivas de los 1373
Panel de información general 1374
Estado actual del clúster 1374
Tendencia de Insights 1374
Perspectivas abiertas actuales 1375
OpenSearch Clústeres de servicios 1375
Principales datos desglosados por gravedad 1375
Detalles de la información 1375
Detalles de clúster 1375
Descripción general de 1375
Prácticas recomendadas de configuración 1376
Información 1376
Sección de métricas 1377
Vistas de nodos, índices y fragmentos 1377
Vista de consulta 1378
Acceda a Insights a través de EventBridge los eventos de Amazon 1380
Catálogo de información 1380
Seguridad 1384
Protección de datos 1385
Cifrado en reposo 1386
Node-to-node cifrado 1390
Gestión de identidad y acceso 1391
Tipos de políticas 1391
Realizar y firmar solicitudes OpenSearch de servicio 1400
Cuando las políticas chocan 1402
Referencia de los elementos de las políticas 1403
Opciones avanzadas y consideraciones de la API 1409
Configurar políticas de acceso 1413

Políticas de muestra adicionales 1413
Referencia de permisos de la API 1413
AWS políticas gestionadas 1414
Prevención de la sustitución confusa entre servicios 1424
Control de acceso detallado 1426
Un panorama más amplio: control de acceso detallado y seguridad del servicio OpenSearch 1427
Conceptos clave 1430
Acerca del usuario maestro 1431
Habilitar el control de acceso detallado 1432
Acceder a los OpenSearch paneles de control como usuario maestro 1436
Administrar permisos 1438
Configuraciones recomendadas 1444
Limitaciones 1447
Modificar el usuario maestro 1449
Usuarios maestros adicionales 1449
Instantáneas manuales 1451
Integraciones 1451
Diferencias en la API de REST 1452
Tutorial: Control de acceso detallado con la autenticación de Cognito 1454
Tutorial: Base de datos de usuarios interna con autenticación básica 1459
Validación de conformidad 1462
Resiliencia 1463
JSON Web Tokens 1463
Consideraciones 1463
Modificación de la política de acceso al dominio 1464
Configuración de la autenticación y la autorización de JWT 1464
Uso de un JWT para enviar una solicitud de prueba 1465
Seguridad de infraestructuras 1467
Trabajar con puntos finales de OpenSearch VPC gestionados por el servicio 1468
Autenticación SAML para paneles OpenSearch 1472
Información general de la configuración de SAML 1473
Consideraciones 1473
Autenticación SAML para dominios de VPC 1474
Modificación de la política de acceso al dominio 1474

Configuración de la autenticación iniciada por proveedor de servicios o por proveedor de identidades ... 1476
Configuración de la autenticación iniciada por proveedor de servicios y por proveedor de identidades ... 1483
Configuración de la autenticación SAML (AWS CLI) ... 1483
Configuración de la autenticación SAML (API de configuración) ... 1484
Solución de problemas de SAML ... 1484
Deshabilitar la autenticación SAML ... 1488
Soporte de IAM Identity Center para OpenSearch ... 1489
Consideraciones ... 1489
Modificación de la política de acceso al dominio ... 1490
Configuración de la autenticación y autorización de Centro de identidades de IAM (consola) ... 1490
Configuración del control de acceso detallado ... 1491
Configuración de la autenticación y autorización del Centro de identidades de IAM (CLI) ... 1491
Deshabilitación de la autenticación del Centro de identidades de IAM en el dominio ... 1492
Autenticación de Amazon Cognito para paneles OpenSearch ... 1492
Requisitos previos ... 1493
Configuración de un dominio para utilizar la autenticación de Amazon Cognito ... 1497
Permitir el rol autenticado ... 1500
Configuración de proveedores de identidades ... 1502
(Opcional) Configuración de acceso pormenorizado ... 1502
(Opcional) Personalización de la página de inicio de sesión ... 1503
(Opcional) Configuración de seguridad avanzada ... 1504
Testeo ... 1504
Cuotas ... 1504
Problemas habituales de configuración ... 1505
Desactivación de la autenticación de Amazon Cognito para paneles OpenSearch ... 1510
Eliminar dominios que utilizan la autenticación de Amazon Cognito para los paneles OpenSearch ... 1510
Cómo utilizar roles vinculados a servicios ... 1510
Rol de creación de dominios de VPC y orígenes de datos ... 1511
Rol de creación de colecciones ... 1514
Rol de creación de canalizaciones ... 1517
Código de muestra ... 1521
Compatibilidad con clientes de Elasticsearch ... 1521

Compresión de solicitudes HTTP 1522
Habilitar la compresión gzip 1522
Encabezados obligatorios 1523
Código de muestra (Python 3) 1523
Usando el AWS SDKs 1525
Java 1525
Python 1536
Nodo 1539
Indexación de datos 1542
Restricciones de nomenclatura de los índices 1542
Reducción del tamaño de la respuesta 1543
Códecs de índice 1545
Cargando datos de streaming en el OpenSearch servicio 1545
Cargando datos de streaming desde Ingestion OpenSearch 1546
Carga de datos de streaming desde Amazon S3 1546
Cargar datos de streaming desde Amazon Kinesis Data Streams 1552
Carga de datos de streaming desde Amazon DynamoDB 1556
Carga de datos de streaming desde Amazon Data Firehose 1560
Carga de datos de streaming desde Amazon CloudWatch 1561
Cargando datos de streaming desde AWS IoT 1561
Carga de datos con Logstash 1561
Configuración 1561
Buscar datos 1565
Búsquedas de URI 1566
Búsquedas de cuerpo de la solicitud 1567
Potenciar campos 1569
Resaltado de resultados de búsqueda 1569
API de recuento 1571
Paginar los resultados de búsqueda 1572
Punto en el tiempo 1572
Los parámetros `from` y `size` 1572
Lenguaje de consulta de paneles 1573
Paquetes 1575
Permisos necesarios 1576
Carga de paquetes en Amazon S3 1576
Importación y asociación de paquetes 1577

Uso de paquetes con OpenSearch ... 1578
Actualización de paquetes ... 1583
Actualización manual de los índices con un nuevo diccionario ... 1586
Disociación y eliminación de paquetes ... 1588
Plugins personalizados ... 1589
Complementos de terceros ... 1607
Compatibilidad con SQL ... 1611
Ejemplo de llamada ... 1613
Notas y diferencias ... 1613
SQL Workbench ... 1614
CLI SQL ... 1466
Controlador JDBC ... 1614
Controlador ODBC ... 1615
Búsqueda en clústeres ... 1615
Limitaciones ... 1616
Requisitos previos de búsqueda entre clústeres ... 1617
Precio de búsqueda entre clústeres ... 1617
Configuración de una conexión ... 1617
Eliminación de una conexión ... 1618
Configuración de seguridad y explicación de ejemplo ... 1619
OpenSearch Cuadros de mando ... 1625
Learning to Rank ... 1625
Introducción a Aprender a clasificar ... 1626
API de Aprender a clasificar ... 1648
Búsqueda asíncrona ... 1654
Ejemplo de llamada de búsqueda ... 1654
Permisos de búsqueda asíncrona ... 1656
Configuración de búsqueda asincrónica ... 1657
Búsqueda en clústeres ... 1657
UltraWarm ... 1659
Punto en el tiempo ... 1659
Consideraciones ... 1660
Creación de un PIT ... 1660
Permisos de puntos en el tiempo ... 1662
Configuración de PIT ... 1663
Búsqueda en clústeres ... 1663

UltraWarm 1664
Búsqueda de agentes 1664
Búsqueda semántica 1664
Búsqueda de segmentos simultánea 1665
Generación de consultas en lenguaje natural 1666
Requisitos previos 1666
Introducción 1666
Configuración de permisos 1667
Automatización de la configuración 1667
Búsqueda vectorial 1668
Importación de vectores de Amazon S3 a OpenSearch Serverless 1669
Arquitecturas de integración 1670
Permisos de IAM necesarios 1672
Configuración de la integración de Amazon S3 Vectors con OpenSearch 1675
Capacidades de búsqueda avanzada con un motor vectorial Amazon S3 1677
Activación de S3 Vectors 1678
Creación de índices con el motor vectorial de S3 1678
Limitaciones de funcionalidad 1681
Indexación de documentos 1682
Buscar documentos 1683
Parámetros de asignación compatibles 1684
Facturación y medición 1685
Deshabilitación del motor s3vector 1685
Búsqueda k-NN 1686
Introducción a k-NN 1686
Diferencias, ajuste y limitaciones de k-NN 1689
Ingestión vectorial 1689
Funcionamiento 1690
OpenSearch Precios de ingestión vectorial 1690
Requisitos previos 1690
Crear una base de datos vectorial 1690
Funciones relacionadas 1692
Exporte el índice vectorial de Amazon S3 al motor vectorial de OpenSearch servicios 1692
Importación del espacio de nombres vectorial de Amazon S3 al motor vectorial OpenSearch de servicios 1693
Ver los trabajos de ingestión de vectores y el historial de importación 1695

Optimizar automáticamente ... 1696
Ventajas ... 1697
Funcionamiento ... 1697
Requisitos previos ... 1698
Casos de uso para la optimización automática ... 1699
Limitaciones ... 1700
Facturación y costos ... 1701
Formatos de datos admitidos ... 1702
Uso de la optimización automática en la consola ... 1704
Características relacionadas ... 1711
Aceleración por GPU para la indexación vectorial ... 1711
Requisitos previos ... 1711
Funcionamiento ... 1711
Precios de aceleración de GPU ... 1712
Operaciones de escritura y aceleración de la GPU ... 1713
Configuraciones de índice compatibles ... 1713
Compatible Regiones de AWS ... 1713
Habilitar la aceleración por GPU ... 1714
Cree índices vectoriales acelerados por GPU ... 1717
Indexación y fusión forzada ... 1719
Prácticas recomendadas ... 1719
OpenSearch Cuadros de mando ... 1721
Control del acceso a Dashboards ... 1722
Uso de un proxy para acceder al servicio desde los paneles OpenSearch ... 1722
Configuración de OpenSearch Dashboards para utilizar un servidor de mapas de WMS ... 1726
Conexión de un servidor de paneles de control local al servicio OpenSearch ... 1727
Administración de índices en OpenSearch Dashboards ... 1729
Características adicionales ... 1729
OpenSearch INTERFAZ DE USUARIO ... 1730
Historial de versiones ... 1731
Introducción ... 1734
Permisos necesarios para crear aplicaciones OpenSearch de Amazon Service ... 1734
Creación de una aplicación de ... 1738
Administrar los administradores de aplicaciones ... 1746
IA de agencia ... 1750
Capacidades clave ... 1750

Límites de precio y uso 1751
Seguridad 1751
Activación y desactivación de las funciones de IA de los agentes 1752
Disponibilidad por región 1753
Chat de agencia 1753
Agente de investigación 1758
Memoria de agente 1760
Cifrado y clave gestionada por el cliente 1761
Requisitos previos para usar claves administradas por el cliente 1762
Creación de una aplicación con cifrado de claves gestionado por el cliente mediante la consola 1764
Crear una aplicación con cifrado de clave gestionado por el cliente mediante AWS CLI 1764
Supervisión del uso de claves gestionadas por el cliente 1765
Actualización de la configuración de cifrado 1765
Habilitar la federación SAML con IAM 1766
Paso 1: Configurar la aplicación de proveedor de identidades (Okta) 1767
Paso 2: Configura la AWS configuración de Okta 1771
Paso 3: Crear la política de acceso a Amazon OpenSearch Service en IAM 1773
Paso 4: Compruebe la experiencia de inicio de sesión único iniciada por el proveedor de identidad con SAML 1776
Paso 5: Configuración del control de acceso detallado basado en atributos de SAML 1780
Administración de las asociaciones de orígenes de datos y los permisos de acceso a la VPC 1784
Asociar una fuente de datos a una aplicación de interfaz de OpenSearch usuario 1784
Administrar el acceso a los dominios en una VPC 1785
Configuración del acceso a las colecciones OpenSearch sin servidor en una VPC 1787
Uso de los espacios de trabajo OpenSearch de Amazon Service 1790
Creación de espacios de trabajo de aplicaciones de OpenSearch interfaz de usuario 1790
La privacidad del espacio de trabajo y los colaboradores 1791
Tipos de WorkSpace 1792
Acceso a datos entre regiones y cuentas 1793
Acceso a los datos de varias cuentas a los OpenSearch dominios 1794
Búsqueda en clústeres 1817
Administrar el acceso a la OpenSearch interfaz de usuario desde un punto final de VPC 1828
Crear una conexión privada entre una VPC y una interfaz de usuario OpenSearch 1828
Actualización de la política de puntos finales de la VPC para permitir el acceso a la aplicación de interfaz de usuario OpenSearch 1830

Revocar el acceso a la OpenSearch interfaz de usuario en una política de puntos finales de VPC ... 1830
Migre desde el panel OpenSearch ... 1831
Información general sobre la migración ... 1832
Requisitos previos ... 1832
Paso 1: Exporte los objetos guardados desde los OpenSearch paneles ... 1833
Paso 2: Importa los objetos guardados a la interfaz de usuario OpenSearch ... 1836
Cuotas y puntos de conexión ... 1838
OpenSearch Puntos finales de interfaz de usuario ... 1838
OpenSearch Cuotas de servicio de interfaz ... 1841
Administrar índices ... 1843
Instancias optimizadas para OpenSearch ... 1843
Limitaciones ... 1844
Ajustes para mejorar el rendimiento de la ingesta ... 1845
En qué se diferencian las instancias OpenSearch optimizadas de otras instancias ... 1845
En qué se diferencian las instancias OpenSearch optimizadas de UltraWarm las instancias ... 1846
Aprovisionamiento de un dominio con instancias OpenSearch optimizadas ... 1846
Almacenamiento de varios niveles ... 1848
Arquitectura de almacenamiento de varios niveles ... 1849
UltraWarm almacenamiento ... 1855
Requisitos previos ... 1856
UltraWarm requisitos de almacenamiento y consideraciones de rendimiento ... 1858
UltraWarm precios ... 1859
Habilitando UltraWarm ... 1859
Migración de índices al almacenamiento UltraWarm ... 1862
Automatizar migraciones ... 1865
Ajuste de migración ... 1865
Cancelación de migraciones ... 1866
Listado de índices calientes y templados ... 1866
Devolución de índices templados al almacenamiento caliente ... 1866
Restauración de índices templados a partir de instantáneas ... 1867
Instantáneas manuales de índices templados ... 1868
Migración de índices templados al almacenamiento frío ... 1869
Prácticas recomendadas para índices KNN ... 1869
Inhabilitando UltraWarm ... 1870

Almacenamiento en frío 1871
Requisitos previos 1872
Requisitos de almacenamiento en frío y consideraciones de rendimiento 1873
Precio del almacenamiento en frío 1873
Habilitación del almacenamiento en frío 1874
Gestión de índices inactivos en los paneles OpenSearch 1876
Migración de índices al almacenamiento frío 1876
Automatización de las migraciones al almacenamiento en frío 1878
Cancelación de migraciones al almacenamiento en frío 1878
Listado de índices almacenados en frío 1879
Migración de índices fríos al almacenamiento templado 1883
Restauración de índices fríos a partir de instantáneas 1884
Cancelación de migraciones del almacenamiento en frío al templado 1884
Actualización de metadatos de índices almacenados en frío 1885
Eliminación de índices almacenados en frío 1885
Deshabilitar el almacenamiento en frío 1886
Index State Management 1886
Crear una política de ISM 1887
Ejemplos de política 1888
Plantillas de ISM 1892
Diferencias 1892
Tutorial: Automatización de procesos de ISM 1894
Acumulaciones de índices 1899
Crear un trabajo acumulativo de índices 1899
Transformaciones de índices 1901
Creación de un trabajo de transformación de índice 1901
Replicación entre clústeres 1903
Limitaciones 1904
Requisitos previos 1904
Requisitos de los permisos 1905
Configuración de una conexión entre clústeres 1906
Inicio de la replicación 1907
Confirmación de replicación 1908
Pausa y reanudación de la replicación 1909
Detención de la replicación 1910
Seguimiento automático 1910

Actualización de los dominios conectados ... 1912
Reindexación remota ... 1912
Requisitos previos ... 1913
Reindexe los datos entre los dominios de Internet OpenSearch del Servicio ... 1914
Reindexación de datos cuando el dominio remoto esté en una VPC ... 1915
Vuelva a indexar los datos entre dominios que no son de servicio OpenSearch ... 1920
Reindexar conjuntos de datos grandes ... 1921
Configuración remota de reindexación ... 1922
Data Streams ... 1923
Introducción a Data Streams ... 1923
Monitoreo de datos ... 1927
Alertas ... 1927
Permisos de alertas ... 1928
Introducción a las alertas ... 1928
Notificaciones ... 1929
Diferencias ... 1930
Detección de anomalías ... 1931
Requisitos previos ... 1932
Introducción a la detección de anomalías ... 1932
Tutorial: Detectar un uso elevado de la CPU con la detección de anomalías ... 1935
Asistente de IA ... 1939
Compatible Regiones de AWS ... 1939
Configuración de AI Assistant para el OpenSearch servicio ... 1940
Genere visualizaciones utilizando un lenguaje natural ... 1940
Ver resúmenes e información sobre alertas ... 1941
Antes de empezar ... 1944
Visualización de resúmenes de alertas y conocimientos ... 1945
Consulta los resúmenes de los resultados de las consultas generados por AI Assistant en la página Discover ... 1946
Vea los detectores de anomalías recomendados ... 1948
Accede al chat de AI Assistant si tienes preguntas sobre el OpenSearch servicio ... 1949
Machine learning ... 1952
Conectores para Servicios de AWS ... 1952
Requisitos previos ... 1953
Crea un conector de OpenSearch servicio ... 1956
Conectores para plataformas externas ... 1958

Requisitos previos ... 1959
Crea un conector de OpenSearch servicio ... 1962
CloudFormation integraciones de plantillas ... 1964
Plantillas disponibles CloudFormation ... 1965
Requisitos previos ... 1966
Plantillas de Amazon Bedrock ... 1968
Configuración de la búsqueda de agentes con Bedrock Claude ... 1969
Plantillas de integración de servidores MCP ... 1970
SageMaker Plantillas de Amazon ... 1973
Inferencia remota para plantillas de resaltado semántico ... 1974
Configuración de ML Commons no compatible ... 1975
Complemento de Flow Framework ... 1976
Creación de conectores ML en Service OpenSearch ... 1977
Configuración de permisos ... 1984
Security Analytics ... 1986
Componentes y conceptos de Security Analytics ... 1986
Tipos de registro ... 1987
Detectores ... 1987
Reglas ... 1987
Hallazgos ... 1988
Alertas ... 1988
Exploración de Security Analytics ... 1988
Configuración de permisos ... 1990
Resolución de problemas ... 1992
No existe tal error de índice ... 1992
Observabilidad ... 1993
El flujo de trabajo de observabilidad en Service OpenSearch ... 1993
OpenSearch Interfaz de usuario y paneles OpenSearch ... 1995
Registros ... 1996
OpenSearch Espacio de trabajo de interfaz de usuario y observabilidad ... 1997
Consulta tus registros mediante PPL ... 1998
Consulta tus registros mediante IA ... 1999
Consulta tus registros mediante SQL ... 1999
Consulta de los registros mediante DQL ... 2000
Cuadros de mando y alertas para los registros ... 2000
Rastros ... 2001

Procesamiento e ingestión de datos de rastreo 2001
OpenSearch Espacio de trabajo de interfaz de usuario y observabilidad 1997
Características principales de 2002
Prácticas recomendadas 2006
Monitorización y alertas 2006
Configure CloudWatch las alarmas 2006
Habilitación de la publicación de registros 2007
Estrategia de particiones 2007
Determinar los recuentos de particiones y nodos de datos 2008
Evite el sesgo de almacenamiento 2009
Stability 2009
Manténgase al día con OpenSearch 2009
Cómo mejorar el rendimiento de las instantáneas 2010
Habilite los nodos maestros dedicados 2011
Implemente en varias zonas de disponibilidad 2011
Controle el flujo de incorporación y el almacenamiento en búfer 2011
Cree asignaciones para cargas de trabajo de búsqueda 2012
Utilice plantillas de índice 2013
Administre los índices con la administración de estado de índice 2014
Eliminar los índices que no se utilizan 2014
Utilice varios dominios para disfrutar de una alta disponibilidad 2014
Desempeño 2015
Optimice el tamaño y la compresión de solicitudes 2015
Reduzca el tamaño de las respuestas de solicitudes masivas 2015
Ajuste intervalos de actualización 2016
Habilite el ajuste automático 2016
Seguridad 2016
Cómo habilitar el control de acceso detallado 2016
Implemente dominios dentro de una VPC 2017
Aplique una política de acceso restrictivo 2017
Habilite el cifrado en reposo 2017
Habilita el cifrado node-to-node 2018
Supervisa con AWS Security Hub CSPM 2018
Optimización de costos 2018
Use los tipos de instancia de última generación 2018
Use los volúmenes gp3 de Amazon EBS más recientes 2019

Utilice UltraWarm y almacene en frío los datos de registro de series temporales 2019
Revise las recomendaciones para instancias reservadas 2020
CloudWatch Alarmas recomendadas 2020
Otras alarmas para tener en cuenta 2025
Determinación del tamaño de dominios 2029
Cálculo de requisitos de almacenamiento 2029
Selección del número de particiones 2031
Selección de tipos de instancias y pruebas 2033
Escala de petabytes 2035
Nodos coordinadores dedicados 2036
Cuándo utilizar nodos coordinadores dedicados 2037
Arquitectura y comportamiento 2037
Requisitos y limitaciones 2038
Aprovisionamiento de nodos coordinadores dedicados 2038
Prácticas recomendadas 2039
Nodos maestros dedicados 2041
Elección del número de nodos maestros dedicados 2042
Elección de tipos de instancias para nodos principales dedicados 2043
Referencia general 2045
Tipos de instancias admitidas 2045
Tipos de instancias de generación actual 2046
Tipos de instancias de generación anterior 2062
Características por versión de motor 2073
Complementos por versión de motor 2080
Complementos opcionales 2084
Operaciones admitidas 2084
Diferencias de API destacadas 2084
Cuotas 2141
Cuotas de almacenamiento en nodos cálidos 2141
Número de nodos de datos por AZ 2043
Límite total de nodos por familia de instancias 2143
Cuotas de tamaño del volumen de EBS 2144
Cuotas de red 2156
Cuotas de tamaño de la partición 2166
Cuotas de recuento de la partición 2167
Cuota de procesamiento de Java 2168

Cuota de políticas de dominio 2168
instancias reservadas 2168
Comprar instancias reservadas (consola) 2169
Adquisición de instancias reservadas (AWS CLI) 2170
Compra de instancias reservadas (AWS SDKs) 2172
Examinar los costos 2174
Optimización de costos 2174
Optimización de costes para los clústeres gestionados por Amazon OpenSearch Service . 2175
Optimización de costes para Amazon OpenSearch Service Serverless 2180
Otros recursos admitidos 2185
Tutoriales 2187
Creación y búsqueda de documentos 2187
Requisitos previos 2187
Adición de un documento a un índice 2188
La creación se genera automáticamente IDs 2189
Actualización de un documento con un comando POST 2190
Ejecución de acciones por lotes 2191
Búsqueda de documentos 2192
Recursos relacionados 2194
Migración al servicio OpenSearch 2194
Tomar y cargar la instantánea 2194
Crear un dominio 2196
Conceda permisos al bucket de S3. 2197
Restaurar la instantánea 2199
Creación de una aplicación de búsqueda 2202
Requisitos previos 2203
Paso 1: Indexe los datos de muestra 2203
Paso 2: crear e implementar una función de Lambda 2204
Paso 3: crear la API en API Gateway 2207
Paso 4: (opcional) modificar la política de acceso a dominios 2209
Asigne el rol de Lambda (si utiliza un control de acceso detallado) 2212
Paso 5: Pruebe la aplicación web 2213
Siguientes pasos 2215
Visualización de las llamadas del servicio de atención 2216
Paso 1: configure los requisitos previos 2217
Paso 2: copie el código de muestra 2218

(Opcional) Paso 3: indexe los datos de ejemplo ... 2222
Paso 4: analice y visualice sus datos ... 2224
Paso 5: elimine recursos y pasos siguientes ... 2228
Cambio de nombre OpenSearch de Amazon Service ... 2230
Nueva versión de la API ... 2230
Tipos de instancia con cambio de nombre ... 2231
Cambios en las políticas de acceso ... 2231
Políticas de IAM ... 2231
Políticas de SCP ... 2231
Nuevos tipos de recursos ... 2232
Kibana ha cambiado su nombre a Dashboards OpenSearch ... 2233
CloudWatch Métricas renombradas ... 2234
Cambios en la consola de Billing and Cost Management ... 2235
Nuevo formato de evento ... 2236
¿Qué permanece igual? ... 2236
Comience: actualice sus dominios a OpenSearch 1.x ... 2236
Resolución de problemas ... 2238
¿No puedes acceder a los OpenSearch paneles ... 2238
No se puede obtener acceso al dominio de VPC ... 2238
Clúster en estado de solo lectura ... 2238
Estado rojo del clúster ... 2240
Corrección automática de clústeres en rojo ... 2241
Recuperación de una carga de procesamiento elevada continua ... 2242
Estado amarillo del clúster ... 2244
ClusterBlockException ... 2244
Falta de espacio de almacenamiento disponible ... 2245
Presión alta de memoria de JVM ... 2245
Error al migrar a multi-AZ con modo de espera ... 2246
Creación de un índice, una plantilla de índice o una política de ISM durante la migración de dominios sin modo de espera a dominios con modo de espera ... 1992
Número incorrecto de copias de datos ... 2246
JVM OutOfMemoryError ... 2247
Nodos de clúster defectuosos ... 2248
Límite máximo de fragmentos superado ... 2248
Dominio atascado en estado de procesamiento ... 2249
Bajo balance de ráfaga EBS ... 2249

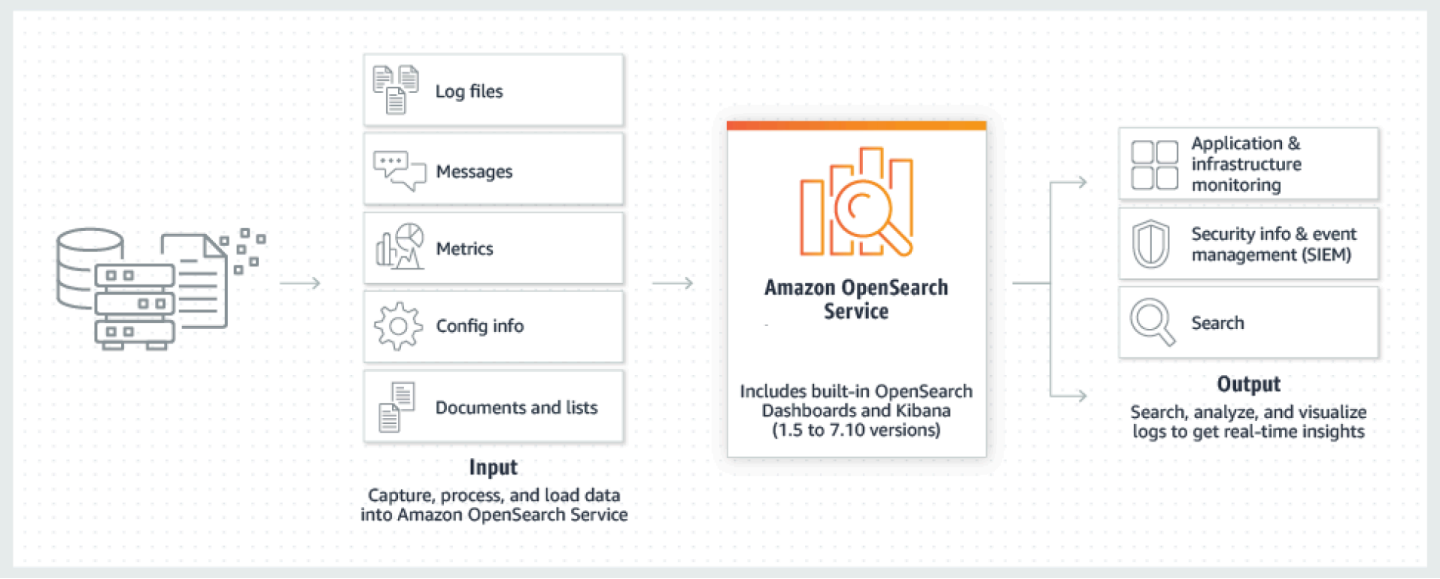
La métrica de EBS aumenta durante el redimensionamiento del volumen 2249
No se pueden habilitar los registros de auditoría 2250
No se puede cerrar el índice 2250
Verificaciones de licencias del cliente 2251
Limitación controlada de solicitudes 2251
No se puede usar SSH en nodo 2251
Error de instantánea "No válido para la clase de almacenamiento del objeto" 2251
Encabezado de host no válido 2252
Tipo de instancia M3 no válido 2252
Las consultas activas dejan de funcionar después de habilitarlas UltraWarm 2252
No se puede cambiar a una versión anterior después de una actualización 2253
¿Necesitas un resumen de los dominios de todos Regiones de AWS 2253
Error del navegador al usar los OpenSearch paneles 2253
Partición de nodos y sesgo de almacenamiento 2254
Partición del índice y el sesgo de almacenamiento 2255
Operación no autorizada tras seleccionar acceso a la VPC 2256
Bloqueo al cargar después de crear el dominio de la VPC 2256
Solicitudes denegadas a la API OpenSearch 2256
No se puede conectar desde Alpine Linux 2257
Hay demasiadas solicitudes de Search Backpressure 2258
Error de certificado cuando se utiliza un SDK 2258
La instalación del complemento personalizado falla debido a la compatibilidad de las versiones 2259
Historial de revisión 2262
Actualizaciones anteriores 2336
AWS Glosario 2340
..... mmcccxli

¿Qué es Amazon OpenSearch Service?

Amazon OpenSearch Service es un servicio gestionado que facilita la implementación, el funcionamiento y el escalado de OpenSearch clústeres en la AWS nube. Un dominio de OpenSearch servicio es sinónimo de OpenSearch clúster. Los dominios son clústeres con la configuración, los tipos de instancia, los recuentos de instancias y los recursos de almacenamiento que especifique. Amazon OpenSearch Service es compatible con OpenSearch Elasticsearch OSS heredado (hasta la 7.10, la versión final de código abierto del software). Al crear un dominio, tiene la opción de elegir qué motor de búsqueda utilizar.

OpenSearches un motor de búsqueda y análisis totalmente de código abierto para casos de uso como el análisis de registros, la supervisión de aplicaciones en tiempo real y el análisis del flujo de clics. Para obtener más información, consulte la Documentación de OpenSearch .

Amazon OpenSearch Service aprovisiona todos los recursos OpenSearch del clúster y lo lanza. También detecta y reemplaza automáticamente los nodos de OpenSearch servicio defectuosos, lo que reduce la sobrecarga asociada a las infraestructuras autogestionadas. Puede escalar el clúster con una única llamada al API o con algunos clics en la consola.



Para empezar a utilizar el OpenSearch Servicio, debe crear un dominio de OpenSearch Servicio, que equivale a un OpenSearch clúster. Cada instancia EC2 del clúster actúa como un nodo de OpenSearch servicio.

Puede utilizar la consola OpenSearch de servicio para instalar y configurar un dominio en cuestión de minutos. Si prefiere el acceso programático, puede usar la AWS CLIAWS SDKs, la o Terraform.

Características de Amazon OpenSearch Service

OpenSearch El servicio incluye las siguientes funciones:

Escalado

- Varias configuraciones de CPU, memoria y capacidad de almacenamiento, que se denominan tipos de instancias, que incluyen instancias rentables de Graviton
- Soporta hasta 1002 nodos de datos
- Hasta 25 PB de almacenamiento adjunto
- Almacenamiento en frío UltraWarmy rentable para datos de solo lectura

Seguridad

- AWS Identity and Access Management Control de acceso (IAM)
- Integración sencilla con Amazon VPC y grupos de seguridad de VPC
- Cifrado de datos en reposo y node-to-node cifrado
- Autenticación Amazon Cognito, HTTP básica o SAML para paneles OpenSearch
- Seguridad en el nivel de índice, de documento y de campo
- Registros de auditoría
- Tenencia múltiple de paneles

Estabilidad

- Numerosas ubicaciones geográficas para los recursos, denominadas regiones y zonas de disponibilidad
- Asignación de nodos en dos o tres zonas de disponibilidad de la misma región de AWS , lo que se conoce como Multi-AZ
- Nodos principales dedicados para aligerar las tareas de administración del clúster
- Instantáneas automatizadas para realizar copias de seguridad y restaurar los dominios de servicio OpenSearch

Flexibilidad

- Compatibilidad con SQL para la integración con aplicaciones de inteligencia empresarial (BI)

- Paquetes personalizados para mejorar los resultados de búsqueda

Integración con servicios populares

- Visualización de datos mediante paneles OpenSearch
- Integración con Amazon CloudWatch para monitorear las métricas OpenSearch del dominio del servicio y configurar alarmas
- Integración con AWS CloudTrail los dominios de OpenSearch servicio para auditar las llamadas a la API de configuración
- Integración con Amazon S3, Amazon Kinesis y Amazon DynamoDB para cargar datos de streaming en Service OpenSearch
- Alertas de Amazon SNS cuando los datos superan determinados umbrales

Cuándo usarlo en OpenSearch lugar de Amazon OpenSearch Service

Usa la siguiente tabla para ayudarte a decidir si Amazon OpenSearch Service aprovisionado o autogestionado OpenSearch es la opción correcta para ti.

OpenSearch	OpenSearch Servicio Amazon
• Su organización está dispuesta a supervisar y mantener manualmente los clústeres autoaprovisionados, y cuenta con personas con las habilidad es adecuadas para ello. • Desea tener un control total del código a nivel de compilación. • Su organización prefiere, o utiliza de forma exclusiva, el software de código abierto. • Tiene una estrategia multinube que requiere tecnologías que no son específicas de cada proveedor.	• No desea administrar, supervisar y mantener su infraestructura manualmente. • Desea formas sencillas de administrar los crecientes costos de análisis mediante la organizac ión de capas de datos en distintos niveles de almacenamiento, lo que permite aprovechar la durabilidad y el bajo costo de Amazon S3. • Desea aprovechar las integraciones con otros, Servicios de AWS como DynamoDB, Amazon DocumentDB (con compatibilidad con MongoDB), IAM y. CloudWatch CloudFormation

OpenSearch	OpenSearch Servicio Amazon
• Su equipo es capaz de abordar cualquier problema crítico de producció n. • Desea disponer de la flexibilidad necesaria para utilizar, modificar y ampliar el producto de la forma que desee. • Desea tener acceso inmediato a las nuevas características tan pronto como se publiquen.	• Desea acceder fácilmente a la asistencia de si tiene problemas de mantenimiento preventivo o Soporte de producción. • Desea aprovechar características como la autorreparación, el mantenimiento proactivo, la resiliencia y las copias de seguridad.

Versiones compatibles de Elasticsearch y OpenSearch

OpenSearch El servicio admite las siguientes versiones de: 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 y 1.0

OpenSearch El servicio es compatible con las siguientes versiones de la versión antigua de 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 y 1.5

Recomendamos actualizar a la última OpenSearch versión disponible para aprovechar al máximo el OpenSearch Servicio, en términos de precio-rendimiento, variedad de funciones y mejoras de seguridad.

Soporte extendido y estándar

AWS proporciona correcciones de errores y actualizaciones de seguridad para las versiones con soporte estándar. Para las versiones con soporte extendido, AWS ofrece correcciones de seguridad críticas durante al menos 12 meses después de que finalice el soporte estándar, a una tarifa fija por hora de instancia normalizada (NIH). El NIH se basa en el tamaño de la instancia y en las horas de uso.

Los cargos de soporte extendido se aplican automáticamente cuando un dominio ejecuta una versión que ya no cuenta con el soporte estándar. Para evitar estos cargos, actualice a una versión compatible.

En las siguientes tablas, se muestra el calendario de finalización del soporte para las versiones anteriores de Elasticsearch OpenSearch y las versiones anteriores.

OpenSearch El servicio admite varias versiones OpenSearch y versiones antiguas de código abierto de Elasticsearch. Para algunas versiones, ya publicamos las fechas de fin del soporte estándar y ampliamos el soporte. Le recomendamos que actualice a la última OpenSearch versión disponible para aprovechar al máximo el OpenSearch Servicio en términos de precio-rendimiento, variedad de funciones y mejoras de seguridad. Las siguientes tablas proporcionan listas de Elasticsearch y sus OpenSearch versiones y sus programas de soporte.

El calendario de finalización del soporte para las versiones de Elasticsearch es el siguiente:

Versión del software	Fecha de finalización del soporte estándar	Fecha de finalización del soporte extendido
Elasticsearch versiones 1.5 y 2.3	7 de noviembre de 2025	7 de noviembre de 2026
Elasticsearch versiones de la 5.1 a la 5.5	7 de noviembre de 2025	7 de noviembre de 2026
Elasticsearch versión 5.6	7 de noviembre de 2025	7 de noviembre de 2028
Elasticsearch versiones de la 6.0 a la 6.7	7 de noviembre de 2025	7 de noviembre de 2026

Versión del software	Fecha de finalización del soporte estándar	Fecha de finalización del soporte extendido
Elasticsearch versión 6.8	No se ha anunciado	No se ha anunciado
Elasticsearch versiones de la 7.1 a la 7.8	7 de noviembre de 2025	7 de noviembre de 2026
Elasticsearch versión 7.9	No se ha anunciado	No se ha anunciado
Elasticsearch versión 7.10	No se ha anunciado	No se ha anunciado

El calendario de finalización del soporte para OpenSearch las versiones es el siguiente:

Versión del software	Fecha de finalización del soporte estándar	Fecha de finalización del soporte extendido
OpenSearch versiones de la 1.0 a la 1.2	7 de noviembre de 2025	7 de noviembre de 2026
OpenSearch versiones 1.3	No se ha anunciado	No se ha anunciado

Versión del software	Fecha de finalización del soporte estándar	Fecha de finalización del soporte extendido
OpenSearch versiones 2.3 a 2.9	7 de noviembre de 2025	7 de noviembre de 2026
OpenSearch versiones 2.11 y versiones superiores	No se ha anunciado	No se ha anunciado

Soporte estándar y soporte extendido para OpenSearch Elasticsearch

AWS proporciona correcciones de errores y actualizaciones de seguridad periódicas para las versiones incluidas en Standard Support. Para las versiones con soporte extendido, AWS proporciona correcciones de seguridad críticas durante un periodo de al menos 12 meses después de la finalización del soporte estándar, por una tarifa fija adicional por cada hora de instancia normalizada (NIH). El NIH se calcula como un factor del tamaño de la instancia (por ejemplo, mediana o grande) y del número de horas de la instancia (consulte la sección sobre el cálculo de los cargos por soporte prolongado que aparece a continuación para ver un ejemplo). Los cargos de soporte extendido se aplican automáticamente cuando un dominio ejecuta una versión para la que el soporte estándar ha finalizado. Puede actualizar a una versión reciente que aún esté cubierta por el soporte estándar para evitar los cargos por soporte extendido. Para obtener más información sobre los cargos por soporte extendido, consulte la página de precios. Para obtener información general sobre el soporte extendido, consulte las Preguntas frecuentes sobre soporte extendido.

Cálculo de los cargos por soporte extendido

A los dominios que ejecuten versiones con soporte ampliado se les cobrará una hora de fee/ Normalized instancia (NIH) adicional, por ejemplo, 0,0065$ en la región EE.UU. Este (Virginia del Norte). El NIH se calcula como un factor del tamaño de la instancia (p. ej., mediana o grande) y del número de horas de la instancia. Por ejemplo, si ejecuta una instancia de m7g.medium.search durante 24 horas en la región EE. UU. Este (Virginia del Norte), cuyo precio es de 0,068 dólares/

hora de instancia (bajo demanda), normalmente tendrá que pagar 1632 USD (0,068 USD x 24). Si ejecuta una versión con soporte extendido, pagará 0,0065 USD/NIH adicionales, que se calculan como 0,0065 USD x 24 (número de horas de instancia) x 2 (factor de normalización de tamaño; 2 para instancias medianas), lo que equivale a 0,312 USD por soporte extendido durante 24 horas. El importe total que pagará durante 24 horas será la suma del costo de uso de la instancia estándar y el costo del soporte extendido, que es de 1944 USD (1,632 USD + 0,312 USD). En la siguiente tabla, se muestra el factor de normalización para distintos tamaños de instancias en Service. OpenSearch

Tamaño de instancia	Factor de normalización	
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	

Tamaño de instancia	Factor de normalización	
32xlarge	256	

Precios de Amazon OpenSearch Service

En el caso del OpenSearch servicio, usted paga por cada hora de uso de una instancia EC2 y por el tamaño acumulado de cualquier volumen de almacenamiento de EBS adjunto a sus instancias. También se aplican los cargos estándar de transferencia de AWS datos.

Sin embargo, existen algunas excepciones notables en la transferencia de datos. Si un dominio utiliza varias zonas de disponibilidad, el OpenSearch servicio no factura el tráfico entre las zonas de disponibilidad. Una parte importante de la transferencia de datos se produce dentro de un dominio durante la asignación y el reequilibrio de las particiones. OpenSearch No repare los contadores ni las facturas de este tráfico. Del mismo modo, el OpenSearch servicio no factura la transferencia de datos entre UltraWarm/cold nodes y Amazon S3.

Para obtener información completa sobre los precios, consulta los precios OpenSearch de Amazon Service. Para obtener información acerca de los gastos generados por los cambios de configuración, consulte the section called “Cargos por los cambios de configuración”.

Servicios relacionados

OpenSearch El servicio se utiliza normalmente con los siguientes servicios:

Amazon CloudWatch

OpenSearch Los dominios de servicio envían automáticamente las métricas para que CloudWatch puedas supervisar el estado y el rendimiento del dominio. Para obtener más información, consulte Supervisión de las métricas de los OpenSearch clústeres con Amazon CloudWatch.

CloudWatch Los registros también pueden ir en la otra dirección. Puede configurar CloudWatch los registros para transmitir datos al OpenSearch Servicio para su análisis. Para obtener más información, consulte the section called “Carga de datos de streaming desde Amazon CloudWatch”.

AWS CloudTrail

Úselo AWS CloudTrail para obtener un historial de las llamadas a la API de configuración del OpenSearch servicio y los eventos relacionados de su cuenta. Para obtener más información, consulte Supervisión de las llamadas a la API de Amazon OpenSearch Service con AWS CloudTrail.

Amazon Kinesis

Kinesis es un servicio administrado para el procesamiento de datos de streaming en tiempo real a una escala masiva. Para más información, consulte the section called “Cargar datos de streaming desde Amazon Kinesis Data Streams” y the section called “Carga de datos de streaming desde Amazon Data Firehose”.

Amazon S3

Amazon Simple Storage Service (Amazon S3) proporciona almacenamiento para Internet. Esta guía proporciona el código de muestra de Lambda para la integración con Amazon S3. Para obtener más información, consulte the section called “Carga de datos de streaming desde Amazon S3”.

AWS IAM

AWS Identity and Access Management (IAM) es un servicio web que puede utilizar para administrar el acceso a los dominios de su OpenSearch servicio. Para obtener más información, consulte the section called “Gestión de identidad y acceso”.

AWS Lambda

AWS Lambda es un servicio informático que le permite ejecutar código sin aprovisionar ni administrar servidores. Esta guía proporciona código de muestra de Lambda para transmitir datos desde DynamoDB, Amazon S3 y Kinesis. Para más información, consulte the section called “Cargando datos de streaming en el OpenSearch servicio”.

Amazon DynamoDB

Amazon DynamoDB es un servicio de base de datos NoSQL totalmente administrado que ofrece un rendimiento rápido y predecible, así como una perfecta escalabilidad. Para obtener más información sobre la transmisión de datos al OpenSearch Servicio, consultethe section called “Carga de datos de streaming desde Amazon DynamoDB”.

Amazon Quick

Puede visualizar los datos del OpenSearch Servicio mediante los paneles de control rápidos. Para obtener más información, consulta Uso de Amazon OpenSearch Service con Quick en la Guía rápida del usuario.

> **Note**
>
> OpenSearch incluye algunos códigos de Elasticsearch con licencia de Apache de Elasticsearch B.V. y otros códigos fuente. Elasticsearch B.V. no es el origen de ese otro código fuente. ELASTICSEARCH es una marca comercial registrada de Elasticsearch B.V.

Configuración de Amazon OpenSearch Service

Concesión de permisos

En entornos de producción, le recomendamos que utilice políticas más específicas. Para obtener más información sobre la administración del acceso, consulte la sección Gestión del acceso para AWS los recursos en la Guía del usuario de IAM.

Para dar acceso, agregue permisos a los usuarios, grupos o roles:

- Usuarios y grupos en AWS IAM Identity Center:

  Cree un conjunto de permisos. Siga las instrucciones de Creación de un conjunto de permisos en la Guía del usuario de AWS IAM Identity Center .
- Usuarios gestionados en IAM a través de un proveedor de identidades:

  Cree un rol para la federación de identidades. Siga las instrucciones descritas en Creación de un rol para un proveedor de identidad de terceros (federación) en la Guía del usuario de IAM.
- Usuarios de IAM:
  - Cree un rol que el usuario pueda aceptar. Siga las instrucciones descritas en Creación de un rol para un usuario de IAM en la Guía del usuario de IAM.
  - (No recomendado) Adjunte una política directamente a un usuario o agregue un usuario a un grupo de usuarios. Siga las instrucciones descritas en Adición de permisos a un usuario (consola) de la Guía del usuario de IAM.

Concesión de acceso programático

Los usuarios necesitan acceso programático si quieren interactuar con personas AWS ajenas a. Consola de administración de AWS La forma de conceder el acceso programático depende del tipo de usuario que acceda. AWS

Para conceder acceso programático a los usuarios, elija una de las siguientes opciones.

¿Qué usuario necesita acceso programático?	Para	Mediante
IAM	(Recomendado) Utilice las credenciales de la consola como credenciales temporales para firmar las solicitudes programáticas dirigidas al AWS CLI AWS SDKs, o. AWS APIs	Siga las instrucciones de la interfaz que desea utilizar: • Para ello AWS CLI, consulte Iniciar sesión para el desarrollo AWS local en la Guía del AWS Command Line Interface usuario. • Para ello AWS SDKs, consulte Iniciar sesión para el desarrollo AWS local en la Guía de referencia de AWS SDKs and Tools.
Identidad del personal (Usuarios administrados en el IAM Identity Center)	Utilice credenciales temporales para firmar las solicitudes programáticas dirigidas al AWS CLI, AWS SDKs, o AWS APIs.	Siga las instrucciones de la interfaz que desea utilizar: • Para ello AWS CLI, consulte Configuración del AWS CLI uso AWS IAM Identity Center en la Guía del AWS Command Line Interface usuario. • Para AWS SDKs ver las herramientas y AWS APIs, consulte la autenticación del Centro de Identidad de IAM en la Guía de referencia de herramientas AWS SDKs y herramientas.
IAM	Utilice credenciales temporales para firmar las solicitudes programáticas dirigidas al	Siga las instrucciones de Uso de credenciales temporales

¿Qué usuario necesita acceso programático?	Para	Mediante
	AWS CLI AWS SDKs, o. AWS APIs	con AWS recursos de la Guía del usuario de IAM.
IAM	(No recomendado) Utilice credenciales de larga duración para firmar las solicitudes programáticas dirigidas al AWS CLI AWS SDKs, o. AWS APIs	Siga las instrucciones de la interfaz que desea utilizar: • Para ello AWS CLI, consulte Autenticación con credencia les de usuario de IAM en la Guía del AWS Command Line Interface usuario. • Para obtener AWS SDKs información sobre las herramientas, consulte Autenticarse con credencia les de larga duración en la Guía de referencia de herramientas AWS SDKs y herramientas. • Para ello AWS APIs, consulte Administrar las claves de acceso para los usuarios de IAM en la Guía del usuario de IAM.

Instale y configure el AWS CLI

Si desea utilizar el OpenSearch Servicio APIs, debe instalar la última versión de AWS Command Line Interface (AWS CLI). No necesita el AWS CLI OpenSearch Servicio desde la consola y puede empezar sin la CLI siguiendo los pasos que se indican enCómo empezar a usar Amazon OpenSearch Service.

Para configurar el AWS CLI

1. Para instalar la versión más reciente de AWS CLI para macOS, Linux o Windows, consulte Instalación o actualización de la última versión de AWS CLI.
2. Para configurar AWS CLI y configurar de forma segura su acceso a Servicios de AWS, incluido el OpenSearch Servicio, consulte Configuración rápida con `aws configure`.
3. Para verificar la configuración, introduzca el siguiente DataBrew comando en la línea de comandos.

   ```
   aws opensearch help
   ```

   AWS CLI los comandos utilizan el valor predeterminado Región de AWS de la configuración, a menos que lo configure con un parámetro o un perfil. Para configurarlo Región de AWS con un parámetro, puede añadirlo a cada comando. `--region`

   Para configurar su perfil, añada primero un perfil Región de AWS con nombre en el `~/.aws/config` archivo o en el `%UserProfile%/.aws/config` archivo (para Microsoft Windows). Siga los pasos en Perfiles denominados para la AWS CLI. A continuación, defina sus ajustes Región de AWS y los demás con un comando similar al del ejemplo siguiente.

   ```
   [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
   ```

Abra la consola de .

La mayoría de los temas orientados a la consola de esta sección comienzan en la consola de OpenSearch servicio. Si aún no ha iniciado sesión en su cuenta Cuenta de AWS, abra la consola de OpenSearch servicio y pase a la siguiente sección para empezar OpenSearch a utilizar el servicio.

Cómo empezar a usar Amazon OpenSearch Service

Para empezar, debe registrarse para obtener una cuenta de Cuenta de AWS si no dispone de una. Después de configurar una cuenta, completa el tutorial de introducción a Amazon OpenSearch Service. Consulte los siguientes temas introductorios si desea obtener más información y conocer el servicio:

- Crear un dominio.
- Ajuste el tamaño del dominio en función de la carga de trabajo.
- Controle el acceso a su dominio mediante una política de acceso al dominio o un control de acceso detallado.
- Indexe los datos manualmente o desde otros AWS servicios.
- Utilice los OpenSearch paneles de control para buscar sus datos y crear visualizaciones.
- Conozca más opciones avanzadas para crear un dominio. Para más información, consulte Creación y administración de dominios.
- Descubra cómo administrar los índices de un dominio. Para obtener más información, consulte Administrar índices.
- Prueba uno de los tutoriales para trabajar con Amazon OpenSearch Service. Para obtener más información, consulte Tutoriales.

Para obtener información sobre cómo migrar a OpenSearch Service desde un OpenSearch clúster autogestionado, consulte. the section called “Migración al servicio OpenSearch ”

Para obtener información detallada, consulte Creación y administración de dominios y los demás temas de esta guía. Para obtener información sobre la migración a OpenSearch Service desde un clúster autogestionado OpenSearch , consulte. the section called “Migración al servicio OpenSearch ”

Puede completar los siguientes pasos mediante la consola de OpenSearch servicio AWS CLI, el o el AWS SDK. Para obtener información sobre la instalación y configuración del AWS CLI, consulte la Guía del AWS Command Line Interface usuario.

Crear un dominio OpenSearch de Amazon Service

> **Important**
>
> Este es un tutorial conciso para configurar un dominio de Amazon OpenSearch Service de prueba. No utilice este proceso para crear dominios de producción. Para ver una versión completa del mismo proceso, consulte Creación y administración de dominios.

Un dominio de OpenSearch servicio es sinónimo de OpenSearch clúster. Los dominios son clústeres con la configuración, los tipos de instancia, los recuentos de instancias y los recursos de almacenamiento que especifique. Puede crear un dominio de OpenSearch servicio mediante la consola AWS CLI, el o el AWS SDKs.

Para crear un dominio OpenSearch de servicio mediante la consola

1. Ve a Iniciar sesión en la consola https://aws.amazon.comy selecciona Iniciar sesión en la consola.
2. En Analytics, selecciona Amazon OpenSearch Service.
3. Seleccione Crear un dominio.
4. Proporcione un nombre para el dominio. Los ejemplos de este tutorial utilizan el nombre Películas.
5. Seleccione Creación estándar para el método de creación del dominio.

   > **Note**
   >
   > Para configurar rápidamente un dominio de producción con las prácticas recomendadas, puede elegir Creación sencilla. Para el desarrollo y las pruebas de este tutorial, utilizaremos la Creación estándar.

6. Para las plantillas, seleccione Desarrollo/pruebas.
7. Para la opción de implementación, seleccione Dominio con modo de espera.
8. En Versión, seleccione la versión más reciente.
9. Por ahora, ignore las secciones de Nodos de datos, Almacenamiento de datos en caliente y en frío, Nodos maestros dedicados, Configuración de instantáneas y Puntos de conexión personalizados.

10. Para simplificar este tutorial, utilice un dominio de acceso público. En Red, seleccione Acceso público.
11. En la configuración de control de acceso detallado, mantenga seleccionada la casilla Habilitar el control de acceso detallado. Seleccione Crear usuario maestro e ingrese un nombre de usuario y una contraseña.
12. Por el momento, ignore las secciones Autenticación SAML y Autenticación de Amazon Cognito.
13. Para la Política de acceso, seleccione Utilizar únicamente control de acceso detallado. En este tutorial, el control de acceso detallado controla la autenticación, no la política de acceso al dominio.
14. Omita el resto de la configuración y elija Crear. Los dominios nuevos suelen tardar entre 15 y 30 minutos en inicializarse, pero pueden tardar más en función de la configuración. Una vez inicializado su dominio, selecciónelo para abrir su panel de configuración. Tome nota del punto de conexión del dominio en Información general (por ejemplo, `https://search-my-domain.us-east-1.es.amazonaws.com`), que deberá usar en el siguiente paso.

A continuación: Suba los datos a un dominio OpenSearch de servicio para su indexación

Sube datos a Amazon OpenSearch Service para su indexación

> **Important**
>
> Este es un tutorial conciso para subir una pequeña cantidad de datos de prueba a Amazon OpenSearch Service. Para obtener más información sobre cómo cargar datos en un dominio de producción, consulte Indexación de datos.

Puede cargar datos a un dominio de OpenSearch servicio mediante la línea de comandos o la mayoría de los lenguajes de programación.

El siguiente ejemplo requiere utilizar curl, (un cliente HTTP común), para que el proceso sea más cómodo y rápido. Los clientes como curl no pueden realizar la firma de solicitudes necesaria si sus políticas de acceso especifican roles o usuarios de IAM. Para realizar correctamente este proceso, debe utilizar un control de acceso detallado con un nombre de usuario principal y una contraseña, como los configuró en el Paso 1.

Puede instalar curl en Windows y utilizarlo desde el símbolo del sistema, pero recomendamos utilizar una herramienta como Cygwin o Windows Subsystem for Linux. La distribución de macOS y la mayoría de las distribuciones de Linux vienen con curl preinstalado.

Opción 1: cargar un solo documento

Ejecute el siguiente comando para agregar un único documento al dominio Películas:

```
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'
```

En el comando, proporcione el nombre de usuario y la contraseña que creó en el Paso 1.

Para obtener una explicación detallada de este comando y de cómo realizar solicitudes firmadas a OpenSearch Service, consulteIndexación de datos.

Opción 2: cargar varios documentos

Para cargar un archivo JSON que contiene varios documentos a un dominio OpenSearch de Service

1. Cree un archivo local denominado `bulk_movies.json`. Pegue el siguiente contenido en el archivo y agregue una línea nueva al final:

```
{ "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. Ejecute el siguiente comando en el directorio local en el que está almacenado el archivo para cargarlo en el dominio Películas:

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

Para obtener más información acerca del formato de archivo masivo, consulte Indexación de datos.

Siguiente: buscar documentos

Buscar documentos en Amazon OpenSearch Service

Para buscar documentos en un dominio de Amazon OpenSearch Service, usa la API OpenSearch de búsqueda. Como alternativa, puedes usar los OpenSearch paneles de control para buscar documentos en el dominio.

Para buscar documentos desde la línea de comandos

Ejecute el siguiente comando para buscar el dominio Películas para la palabra mars:

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

Si utilizó datos masivos en la página anterior, intente buscar rebel en su lugar.

Verá una respuesta parecida a la siguiente:

```
{
  "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!"
        }
      }
    ]
  }
}
```

Busque documentos mediante paneles OpenSearch

OpenSearch Dashboards es una popular herramienta de visualización de código abierto diseñada para funcionar con ella. OpenSearch Proporciona una interfaz de usuario útil para que pueda buscar y monitorear índices.

Para buscar documentos de un dominio de OpenSearch servicio mediante paneles

1. Navegue hasta la URL de los OpenSearch paneles de control de su dominio. Puedes encontrar la URL en el panel de control del dominio, en la consola OpenSearch de servicio. La URL tiene este formato:

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

2. Inicie sesión con su nombre de usuario principal y contraseña.
3. Para utilizar Dashboards, debe crear al menos un patrón de índice. Dashboards utiliza estos patrones para identificar los índices que desea analizar. Abra el panel de navegación de la izquierda, seleccione Gestión de pilas, seleccione Patrones de índice y luego seleccione Crear patrón de índice. Para este tutorial, escriba Películas.
4. seleccione Siguiente paso y, a continuación, seleccione Crear patrón de índice. Una vez creado el patrón, puede ver los diversos campos de documento, como `actor` y `director`.
5. Regrese a la pestaña Patrones de índice y asegúrese de que `movies` se establezca como valor predeterminado. Si no lo es, seleccione el patrón y elija el icono de estrella para convertirlo en el predeterminado.
6. Para comenzar a buscar los datos, abra el panel de navegación izquierdo y elija Descubrir.
7. En la barra de búsqueda, escriba mars, si cargó un solo documento, o rebel, si cargó varios documentos y, a continuación, presione Enter. Puede buscar otros términos, como nombres de actores o directores.

Siguiente: eliminar un dominio

Eliminar un dominio OpenSearch de Amazon Service

Dado que el dominio Películas de este tutorial es para hacer pruebas, debe eliminarlo cuando termine de experimentar para evitar incurrir en gastos.

Para eliminar un dominio OpenSearch de servicio de la consola

1. Inicia sesión en la consola OpenSearch de Amazon Service.
2. En Dominios, seleccione el dominio películas.
3. Seleccione Eliminar y confirme la eliminación.

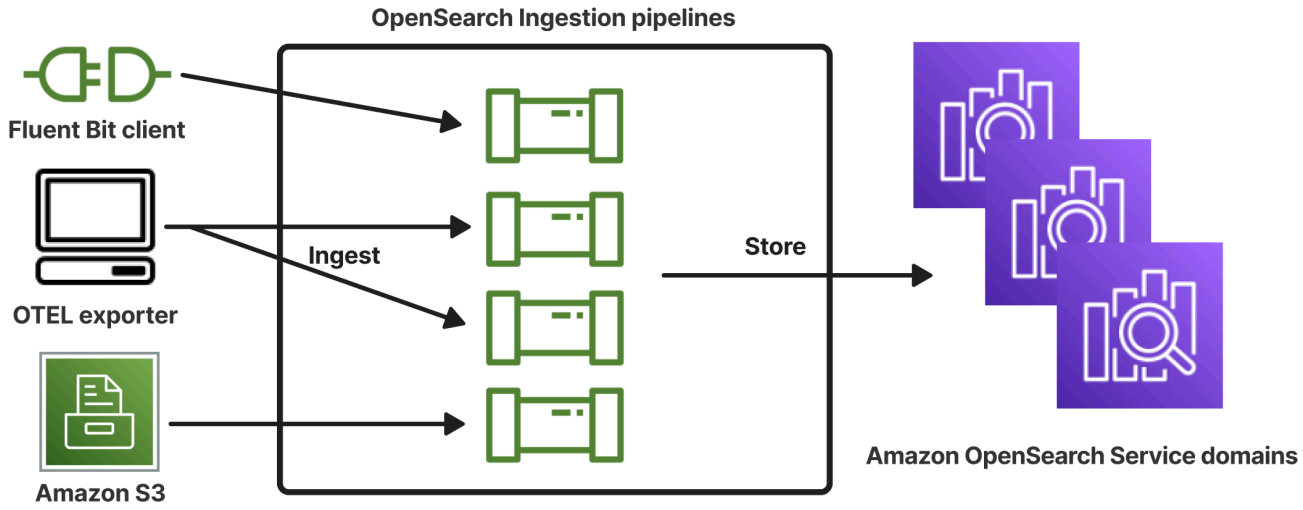
Información general sobre Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion es un recopilador de datos sin servidor totalmente gestionado que transmite registros, métricas y datos de rastreo en tiempo real a los dominios de Amazon OpenSearch Service y a las colecciones OpenSearch sin servidor.

Con OpenSearch Ingestion, ya no necesitará herramientas de terceros como Logstash o Jaeger para ingerir datos. Usted configura sus generadores de datos para que envíen datos a OpenSearch Ingestion y esta los envía automáticamente a su dominio o colección especificados. También puede transformar los datos antes de entregarlos.

Como OpenSearch Ingestion no tiene servidores, no tiene que administrar la infraestructura, parchear el software ni escalar los clústeres manualmente. Puede aprovisionar los canales de ingestión directamente en el Consola de administración de AWS, e OpenSearch Ingestion se encarga del resto.

Como componente de Amazon OpenSearch Service, OpenSearch Ingestion funciona con Data Prepper, un recopilador de datos de código abierto que filtra, enriquece, transforma, normaliza y agrega datos para su análisis y visualización posteriores.



Conceptos clave de Amazon OpenSearch Ingestion

Antes de empezar a usar OpenSearch Ingestion, es útil entender estos conceptos clave.

Canalización

Desde el punto de vista de OpenSearch la ingesta, una canalización hace referencia a un único recopilador de datos aprovisionado que se crea en Service. OpenSearch Puede considerarlo como el archivo de configuración de YAML completo, que incluye una o más subcanalizaciones. Para ver los pasos para crear una canalización de ingesta, consulte the section called “Creación de canalizaciones”.

Subcanalización

Las subcanalizaciones se definen dentro de un archivo de configuración YAML. Cada subcanalización es una combinación de una sola fuente, un búfer, cero o más procesadores y uno o más receptores. Puede definir varias subcanalizaciones en un único archivo YAML, cada una con fuentes, procesadores y receptores únicos. Para facilitar la supervisión con CloudWatch y otros servicios, te recomendamos que especifiques un nombre de canalización que sea distinto de todas sus subcanalizaciones.

Puede encadenar varias subcanalizaciones en un único archivo YAML, de modo que la fuente de una subcanalización sea otra subcanalización y su receptor sea una tercera subcanalización. Para ver un ejemplo, consulta the section called “OpenTelemetry Colector”.

origen

El componente de entrada de una subcanalización. Define el mecanismo mediante el cual una canalización consume los registros. La fuente puede consumir eventos ya sea recibiéndolos a través de HTTPS o leyendo desde puntos de conexión externos, como Amazon S3. Hay dos tipos de fuentes: basadas en push y basadas en pull. Las fuentes basadas en push, como HTTP y los OTel registros, transmiten los registros a los puntos finales de ingestión. Las fuentes basadas en extracción, como OTel trace y S3, extraen datos de la fuente.

Procesadores

Unidades de procesamiento intermedias que pueden filtrar, transformar y enriquecer los registros en el formato deseado antes de publicarlos en el receptor. El procesador es un componente opcional de una canalización. Si no define un procesador, los registros se publican en el formato definido en la fuente. Puede tener más de un procesador. Una canalización ejecuta los procesadores en el orden en el que los define.

Receptor

El componente de salida de una subcanalización. Define uno o más destinos en los que una subcanalización publica los registros. OpenSearch La ingestión admite los dominios OpenSearch

de servicio como receptores. También admite subcanalizaciones como receptores. Esto significa que puedes encadenar varias subcanalizaciones en una sola canalización de OpenSearch ingestión (archivo YAML). OpenSearch Los clústeres autogestionados no se admiten como receptores.

Búfer

La parte de un procesador que actúa como capa entre la fuente y el receptor. No puede configurar manualmente un búfer dentro de la canalización. OpenSearch La ingestión utiliza una configuración de búfer predeterminada.

Ruta

La parte de un procesador que permite a los autores de canalizaciones enviar únicamente los eventos que cumplan determinadas condiciones a distintos receptores.

Una definición de subcanalización válida debe contener una fuente y un receptor. Para obtener más información sobre cada uno de estos elementos de la canalización, consulte Referencia de configuración.

Ventajas de Amazon OpenSearch Ingestion

OpenSearch La ingestión tiene los siguientes beneficios principales:

- Elimina la necesidad de administrar manualmente una canalización autoaprovisionada.
- Escala automáticamente las canalizaciones en función de los límites de capacidad que defina.
- Mantiene la canalización actualizada con parches de seguridad y de errores.
- Ofrece la opción de conectar las canalizaciones a su nube privada virtual (VPC) para añadir un nivel de seguridad adicional.
- Le permite detener e iniciar canalizaciones para controlar los costos.
- Proporciona esquemas de configuración de canalizaciones para los casos de uso más habituales a fin de ayudarle a empezar a trabajar más rápido.
- Le permite interactuar programáticamente con sus canalizaciones a través de las distintas API AWS SDKs y de la API de ingestión. OpenSearch
- Admite la supervisión del rendimiento en Amazon CloudWatch y el registro de errores en CloudWatch Logs.

Limitaciones de Amazon OpenSearch Ingestion

OpenSearch La ingestión tiene las siguientes limitaciones:

- Solo puede ingerir datos en dominios que ejecuten la OpenSearch versión 1.0 o una versión posterior, o Elasticsearch 6.8 o una versión posterior. Si utilizas la fuente de OTel rastreo, te recomendamos que utilices Elasticsearch 7.9 o una versión posterior para poder usar el complemento Dashboards. OpenSearch
- Si una canalización se escribe en un dominio de OpenSearch servicio que se encuentra dentro de una VPC, la canalización debe crearse en el Región de AWS mismo dominio.
- Solo puede configurar un único origen de datos dentro de una definición de canalización.
- No puedes especificar OpenSearch clústeres autogestionados como receptores.
- No puede especificar un punto de conexión personalizado como receptor. Aún puede escribir en un dominio que tenga habilitados los puntos de conexión personalizados, pero debes especificar su punto de conexión estándar.
- No puede especificar los recursos dentro de las regiones registradas como orígenes o receptores.
- Existen algunas limitaciones en los parámetros que se pueden incluir en una configuración de canalización. Para más información, consulte the section called “Requisitos y restricciones de configuración”.

Versiones de Data Prepper admitidas

OpenSearch Actualmente, Ingestion es compatible con las siguientes versiones principales de Data Prepper:

- 2.x

Al crear una canalización con el Editor de código, utilice la opción de `version` necesaria para especificar la versión principal de Data Prepper que va a utilizar. Por ejemplo,. `version: "2"` OpenSearch Ingestion recupera la última versión secundaria compatible de esa versión principal y aprovisiona la canalización con esa versión.

Si no utilizas el editor de código para crear tu canalización, OpenSearch Ingestion aprovisiona automáticamente tu canalización con la última versión compatible.

Actualmente, OpenSearch Ingestion aprovisiona las canalizaciones con la versión 2.7 de Data Prepper. Para obtener información, consulte las notas de la versión 2.7. Ingestion no admite todas las versiones secundarias de una versión principal concreta. OpenSearch

Al actualizar la configuración de una canalización, si hay soporte para una nueva versión secundaria de Data Prepper, OpenSearch Ingestion actualiza automáticamente la canalización a la última versión secundaria compatible de la versión principal especificada en la configuración de la canalización. Por ejemplo, es posible que tengas una configuración `version: "2"` de canalización e OpenSearch Ingestion aprovisionó inicialmente la canalización con la versión 2.6.0. Cuando se añade la compatibilidad con la versión 2.7.0 y realizas un cambio en la configuración de la canalización, OpenSearch Ingestion actualiza la canalización a la versión 2.7.0. Este proceso mantiene su canalización actualizada con las últimas correcciones de errores y mejoras de rendimiento. OpenSearch Ingestion no puede actualizar la versión principal de tu canalización a menos que cambies manualmente la `version` opción en la configuración de la canalización. Para obtener más información, consulte the section called “Actualización de las canalizaciones”.

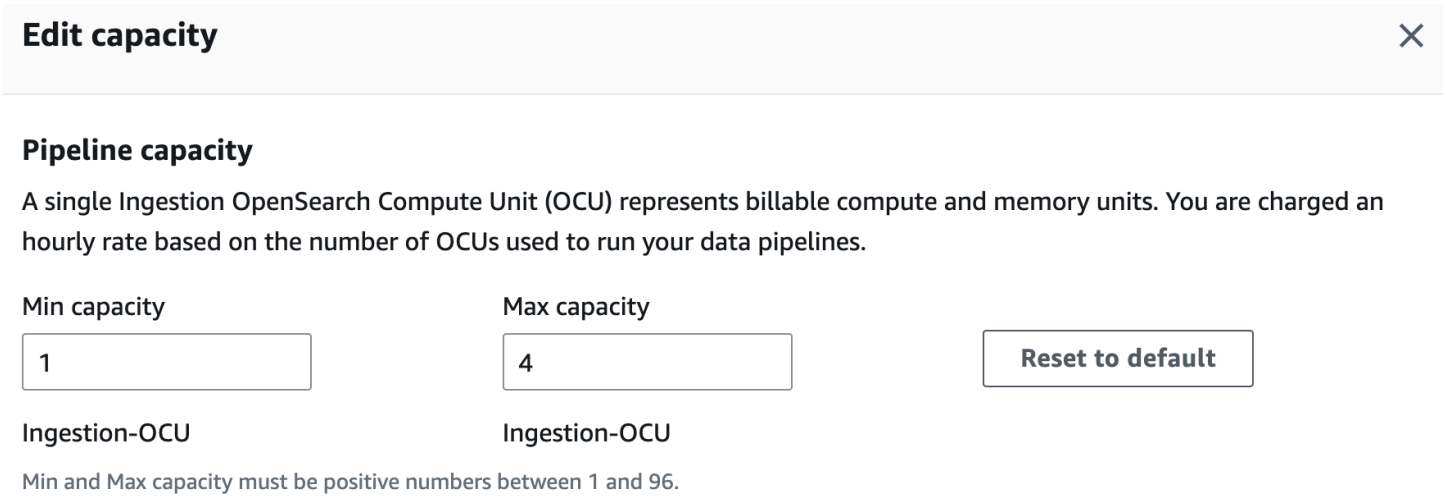
Escalar las canalizaciones en Amazon OpenSearch Ingestion

OpenSearch Ingestion escala automáticamente la capacidad de canalización en función de las unidades de OpenSearch cómputo de ingestión (ingestión) mínimas y máximas especificadas. OCUs Esto elimina la necesidad de aprovisionamiento y administración manuales.

Cada OCU de ingestión es una combinación de aproximadamente 15 GiB de memoria y 2 v. CPUs Puede especificar los valores de OCU mínimo y máximo para una canalización, e OpenSearch Ingestion escalará automáticamente la capacidad de la canalización en función de estos límites.

Especifique los siguientes valores al crear una canalización:

- Capacidad mínima: la canalización puede reducir la capacidad hasta este número de ingestas. OCUs La capacidad mínima especificada también es la capacidad inicial de una canalización.
- Capacidad máxima: la tubería puede aumentar la capacidad hasta este número de OCUs ingestas.



Asegúrese de que la capacidad máxima de la canalización sea lo suficientemente alta como para administrar los picos de carga de trabajo y que la capacidad mínima sea lo suficientemente baja como para minimizar los costos cuando la canalización no esté ocupada. En función de tu configuración, OpenSearch Ingestión escala automáticamente el número de ingestas de tu canalización OCUs para procesar la carga de trabajo de ingesta. En un momento específico, solo se te cobrará por la ingestión OCUs que esté siendo utilizada activamente por tu canal.

La capacidad asignada a tu canalización de OpenSearch ingestion se amplía y reduce en función de los requisitos de procesamiento de tu canalización y de la carga generada por la aplicación cliente. Cuando la capacidad es limitada, OpenSearch Ingestion se amplía mediante la asignación de más unidades de cómputo (GiB de memoria). Cuando tu canalización procese cargas de trabajo más pequeñas o no procese datos en absoluto, puede reducir la cantidad de ingestión mínima configurada. OCUs

Puedes especificar un mínimo de 1 OCU de ingestión, un máximo de 96 ingestas para las canalizaciones sin estado y un máximo de 48 ingestas OCUs para las canalizaciones con estado. OCUs Recomendamos un mínimo de 2 ingestas para las fuentes push. OCUs Cuando el almacenamiento en búfer persistente está activado, puedes especificar un mínimo de 2 y un máximo de 384 ingestas. OCUs

Dada una canalización de registro estándar con una sola fuente, un patrón grok simple y un receptor, cada unidad de cómputo puede admitir hasta 2 MiB por segundo. En el caso de las canalizaciones de registro más complejas con varios procesadores, cada unidad de cómputo puede soportar una carga de ingesta menor. En función de la capacidad de la canalización y el uso de los recursos, se inicia el proceso de escalado OpenSearch de la ingestión.

Para garantizar una alta disponibilidad, la ingesta OCUs se distribuye entre las zonas de disponibilidad (). AZs El número AZs depende de la capacidad mínima que especifique.

Por ejemplo, si especifica un mínimo de 2 unidades de cómputo, la ingesta OCUs que esté en uso en un momento dado se distribuirá uniformemente entre 2 AZs. Si especifica un mínimo de 3 o más unidades de cómputo, la ingesta OCUs se distribuirá uniformemente en 3. AZs Le recomendamos que aprovisione al menos dos entradas OCUs para garantizar una disponibilidad del 99,9% en sus canalizaciones de ingesta.

No se te facturará la ingesta OCUs cuando una canalización esté en los estados,, y. `Create failed` `Creating` `Deleting` `Stopped`

Para ver instrucciones sobre cómo configurar y recuperar los ajustes de capacidad de una canalización, consulte the section called “Creación de canalizaciones”.

OpenSearch Precios de ingestión

En un momento específico, solo pagas por la cantidad de ingestión OCUs que se asigne a una canalización, independientemente de si hay datos circulando por la canalización. OpenSearch Ingestion se adapta inmediatamente a tus cargas de trabajo al aumentar o reducir la capacidad de la canalización en función del uso.

Para obtener información completa sobre los precios, consulta los precios OpenSearch de Amazon Service.

Soportado Regiones de AWS

OpenSearch La ingestión está disponible en un subconjunto de Regiones de AWS ese OpenSearch servicio en el que está disponible. Para ver una lista de las regiones admitidas, consulta los puntos de conexión y las cuotas de Amazon OpenSearch Service en. Referencia general de AWS

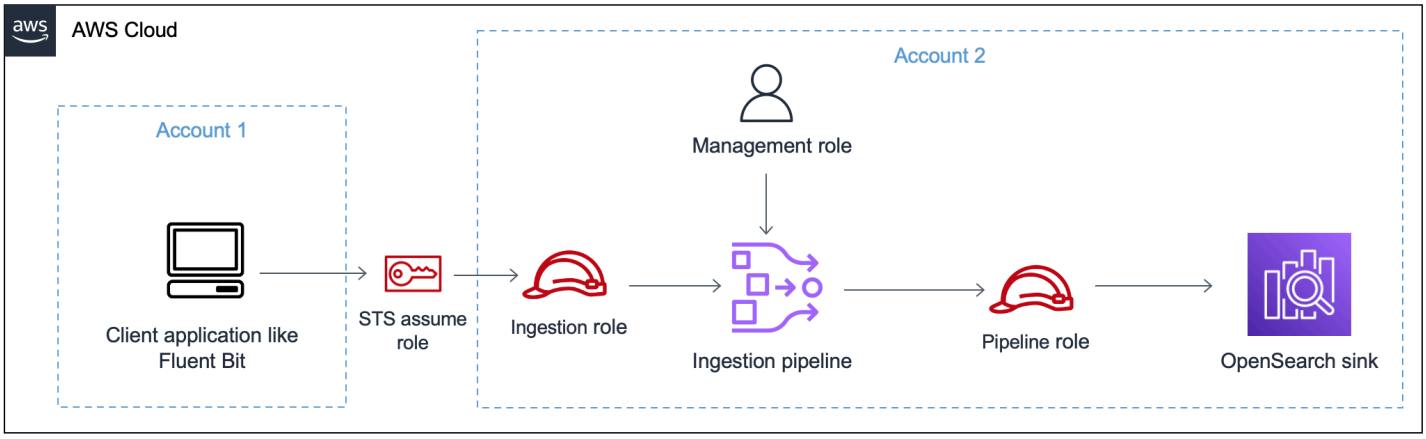
Configuración de roles y usuarios en Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion utiliza diversos modelos de permisos y funciones de IAM para permitir que las aplicaciones de origen escriban en las canalizaciones y que las canalizaciones escriban en los receptores. Antes de empezar a incorporar datos, debe crear uno o más roles de IAM con permisos específicos en función de su caso de uso.

Como mínimo, se requieren los siguientes roles para configurar una canalización exitosa.

Name	Description (Descripción)
Rol de canalización	El rol de canalización proporciona los permisos necesarios para que una canalización lea desde el origen y escriba en el receptor del dominio o de la colección. Puede crear el rol de canalización manualmente o puede hacer que OpenSearch Ingestion lo cree por usted.
Rol de ingesta	El rol de ingestión contiene el permiso `osis:Ingest` para el recurso de canalización. Este permiso permite que los orígenes basados en push incorporen datos en una canalización.

La siguiente imagen muestra una configuración de canalización típica, en la que un origen de datos como Amazon S3 o Fluent Bit escribe en una canalización de una cuenta diferente. En este caso, el cliente debe asumir el rol de incorporación para poder acceder a la canalización. Para obtener más información, consulte the section called “Incorporación entre cuentas”.



Para obtener una guía de configuración sencilla, consulte the section called “Tutorial: incorporar datos a un dominio”.

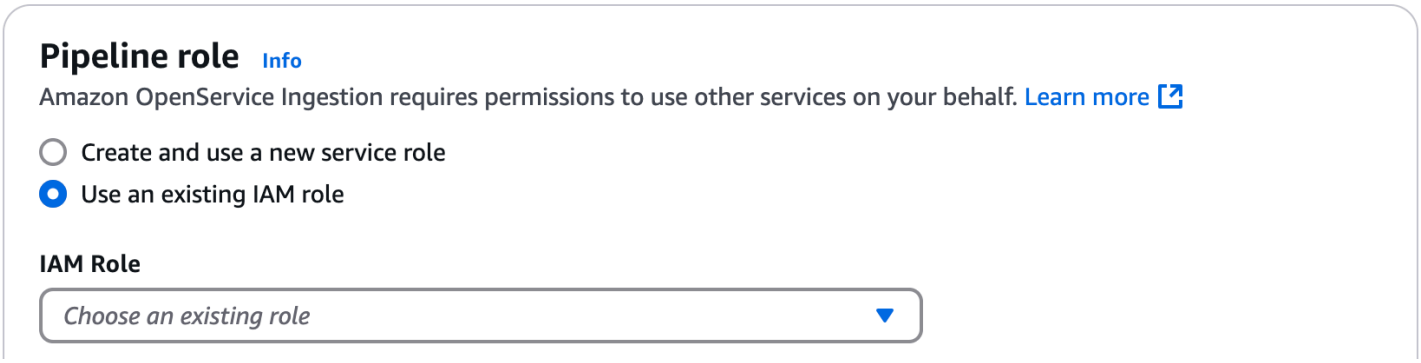
Temas

- the section called “Rol de canalización”
- the section called “Rol de ingesta”
- the section called “Incorporación entre cuentas”

Rol de canalización

Una canalización necesita ciertos permisos para leer de su fuente y escribir en su receptor. Estos permisos dependen de la aplicación cliente o de la Servicio de AWS que esté escribiendo en la canalización y de si el receptor es un dominio de OpenSearch servicio, una colección OpenSearch sin servidor o Amazon S3. Además, es posible que una canalización necesite permisos para extraer información de la aplicación de origen (si el origen es un complemento basado en la extracción) y permisos para escribir en una cola de mensajes fallidos de S3, si está configurada.

Al crear una canalización, tiene la opción de especificar una función de IAM existente que haya creado manualmente o hacer que OpenSearch Ingestion cree automáticamente la función de canalización en función de la fuente y el receptor que haya seleccionado. La siguiente imagen muestra cómo especificar el rol de canalización en la Consola de administración de AWS.



Temas

- Automatización de la creación de roles de canalización
- Creación manual del rol de canalización

Automatización de la creación de roles de canalización

Puede elegir que OpenSearch Ingestion cree el rol de canalización por usted. Identifica automáticamente los permisos que requiere el rol en función del origen y los receptores configurados. Crea un rol de IAM con el prefijo `OpenSearchIngestion-` y el sufijo que introduzcas. Por ejemplo, si lo introduce `PipelineRole` como sufijo, OpenSearch Ingestión crea un rol denominado. `OpenSearchIngestion-PipelineRole`

La creación automática del rol de canalización simplifica el proceso de configuración y reduce la probabilidad de que se produzcan errores de configuración. Al automatizar la creación de roles, puede evitar la asignación manual de permisos y garantizar que se apliquen las políticas correctas

sin correr el riesgo de errores de configuración de seguridad. Esto también ahorra tiempo y mejora el cumplimiento de las normas de seguridad al aplicar las mejores prácticas y, al mismo tiempo, garantizar la coherencia en las distintas implementaciones en proceso.

Solo puede hacer que OpenSearch Ingestion cree automáticamente el rol de canalización en el. Consola de administración de AWS Si utilizas la AWS CLI API de OpenSearch ingestión o una de ellas SDKs, debes especificar una función de canalización creada manualmente.

Para que OpenSearch Ingestion cree el rol por usted, seleccione Crear y usar un nuevo rol de servicio.

> **Important**
>
> Aún tiene que modificar manualmente la política de acceso al dominio o a la colección para conceder el acceso a la función de canalización. En el caso de los dominios que utilicen un control de acceso detallado, también debe asignar el rol de canalización a un rol de backend. Puede realizar estos pasos antes o después de crear la canalización.
> Para obtener instrucciones, consulte los siguientes temas:
>
> - Configurar el acceso a los datos para el dominio
> - Configurar el acceso a los datos y a la red para la colección

Creación manual del rol de canalización

Es posible que prefiera crear el rol de canalización de forma manual si necesita tener más control sobre los permisos para cumplir requisitos específicos de seguridad o conformidad. La creación manual le permite personalizar los roles para adaptarlos a las estrategias de administración de acceso o de infraestructura existentes. También puede optar por la configuración manual para integrar el rol con otro Servicios de AWS o asegurarse de que se ajuste a sus necesidades operativas únicas.

Para elegir un rol de canalización creado manualmente, seleccione Usar un rol de IAM existente y elija un rol existente. El rol debe tener todos los permisos necesarios para recibir datos del origen seleccionado y escribir en el receptor seleccionado. En las siguientes secciones se describe cómo crear manualmente un rol de canalización.

Temas

- Permisos para leer desde un origen

- Permisos para escribir en un receptor de dominio
- Permisos para escribir en un receptor de colecciones
- Permisos para escribir en Amazon S3 o en una cola de mensajes fallidos

Permisos para leer desde un origen

Una canalización de OpenSearch ingestión necesita permiso para leer y recibir datos de la fuente especificada. Por ejemplo, para un origen de Amazon DynamoDB, necesita permisos como `dynamodb:DescribeTable` y `dynamodb:DescribeStream`. Para ver ejemplos de políticas de acceso a roles de canalización para fuentes comunes, como Amazon S3, Fluent Bit y OpenTelemetry Collector, consultethe section called “Integración de canalizaciones”.

Permisos para escribir en un receptor de dominio

Una canalización de OpenSearch ingestión necesita permiso para escribir en un dominio de OpenSearch servicio que esté configurado como receptor. Estos permisos incluyen la capacidad de describir el dominio y enviarle solicitudes HTTP. Estos permisos son los mismos para los dominios públicos y de VPC. Para obtener instrucciones acerca de cómo crear un rol de canalización y especificarlo en la política de acceso a dominios, consulte Permitir que las canalizaciones accedan a los dominios.

Permisos para escribir en un receptor de colecciones

Una canalización de OpenSearch ingestión necesita permiso para escribir en una colección OpenSearch sin servidor que esté configurada como sumidero. Estos permisos incluyen la capacidad de describir la colección y enviarle solicitudes HTTP.

En primer lugar, asegúrese de que su política de acceso a los roles de canalización conceda los permisos necesarios. A continuación, incluya este rol en una política de acceso a los datos y asígnele permisos para crear índices, actualizar índices, describir índices y escribir documentos dentro de la colección. Para obtener instrucciones para completar cada uno de estos pasos, consulte Allowing pipelines to access collections.

Permisos para escribir en Amazon S3 o en una cola de mensajes fallidos

Si especifica Amazon S3 como destino receptor de su canalización, o si habilita una cola de mensajes fallidos (DLQ), el rol de canalización debe permitirle acceder al bucket de S3 que especifique como destino.

Adjunte una política de permisos independiente al rol de canalización que otorgue acceso a la DLQ. Como mínimo, debe otorgarse la acción `S3:PutObject` al rol en el recurso del bucket:

JSON

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

Rol de ingesta

La función de ingestión es una función de IAM que permite a los servicios externos interactuar de forma segura con una canalización de ingestión y enviarlos a ella. OpenSearch En el caso de los orígenes basados en push, como Amazon Security Lake, este rol debe conceder permisos para introducir datos en la canalización, incluido `osis:Ingest`. En el caso de las fuentes basadas en extracciones, como Amazon S3, la función debe permitir que OpenSearch Ingestion la asuma y acceda a los datos con los permisos necesarios.

Temas

- Rol de ingesta para orígenes basados en push
- Rol de ingesta de orígenes basados en extracciones
- Incorporación entre cuentas

Rol de ingesta para orígenes basados en push

En el caso de los orígenes basados en push, los datos se envían o insertan en la canalización de ingesta desde otro servicio, como Amazon Security Lake o Amazon DynamoDB. En este escenario, el rol de ingesta necesita, como mínimo, el permiso `osis:Ingest` para interactuar con la canalización.

La siguiente política de acceso de IAM demuestra cómo conceder este permiso al rol de ingesta:

JSON

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

Rol de ingesta de orígenes basados en extracciones

En el caso de las fuentes basadas en extracciones, la canalización de OpenSearch ingestión extrae o recupera activamente datos de una fuente externa, como Amazon S3. En este caso, la canalización debe asumir un rol de canalización de IAM que conceda los permisos necesarios para acceder al origen de datos. En estos escenarios, el rol de ingesta es sinónimo del rol de canalización.

La función debe incluir una relación de confianza que permita a OpenSearch Ingestion asumirla y permisos específicos de la fuente de datos. Para obtener más información, consulte the section called “Permisos para leer desde un origen”.

Incorporación entre cuentas

Es posible que tengas que introducir datos en una canalización desde otra cuenta Cuenta de AWS, como una cuenta de aplicación. Para configurar la incorporación entre cuentas, defina un rol de incorporaco+pm dentro de la misma cuenta que la canalización y establezca una relación de confianza entre el rol de ingesta y la cuenta de la aplicación:

JSON

```
{
```

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

A continuación, configure la aplicación para que asuma el rol de incorporación. La cuenta de aplicación debe conceder al rol de aplicación AssumeRolepermisos para el rol de ingestión en la cuenta de canalización.

Para ver ejemplos de pasos detallados y ejemplos de políticas de IAM, consulte the section called "Provisión de acceso de ingesta entre cuentas".

Otorgar a Amazon OpenSearch Ingestion pipelines acceso a los dominios

Una canalización OpenSearch de Amazon Ingestion necesita permiso para escribir en el dominio de OpenSearch servicio que está configurado como receptor. Para proporcionar acceso, debe configurar un rol AWS Identity and Access Management (IAM) con una política de permisos restrictiva que limite el acceso al dominio al que una canalización envía datos. Por ejemplo, es posible que desee limitar una canalización de incorporación únicamente al dominio y los índices necesarios para respaldar su caso de uso.

> **Important**
>
> Puede elegir crear el rol de canalización manualmente o puede hacer que OpenSearch Ingestion lo cree por usted durante la creación del canalización. Si eliges la creación automática de roles, OpenSearch Ingestion añadirá todos los permisos necesarios a la política de acceso a los roles de canalización en función de la fuente y el receptor que elijas. Crea un rol de canalización en IAM con el prefijo `OpenSearchIngestion-` y el sufijo que introduzca. Para obtener más información, consulte the section called "Rol de canalización". Si ha pedido a OpenSearch Ingestion que cree el rol de canalización por usted, aún tendrá que incluirlo en la política de acceso al dominio y asignarlo a un rol de back-end (si el dominio usa un control de acceso detallado), antes o después de crear la canalización. Para obtener instrucciones, consulte el paso 2.

Temas

- Paso 1: crear el rol de canalización
- Paso 2: Configurar el acceso a los datos del dominio

Paso 1: crear el rol de canalización

El rol debe tener una política de permisos adjunta que le permita enviar datos al receptor del dominio. También debe tener una relación de confianza que permita a OpenSearch Ingestion asumir la función. Para ver instrucciones sobre cómo adjuntar una política a un rol, consulte Adición de permisos de identidad de IAM en la Guía del usuario de IAM.

En el siguiente ejemplo de política, se muestra el privilegio mínimo que se puede proporcionar en el rol de una canalización para escribir en un solo 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/*"
        }
    ]
}
```

Si piensa reutilizar el rol para escribir en varios dominios, puede ampliar la política sustituyendo el nombre de dominio por un carácter comodín (*).

El rol debe tener la siguiente relación de confianza, que permita a OpenSearch Ingestion asumir el rol de canalización:

JSON

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

Paso 2: Configurar el acceso a los datos del dominio

Para que una canalización escriba datos en un dominio, el dominio debe tener una política de acceso de nivel de dominio que permita al rol de canalización acceder a ellos.

El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización denominado `pipeline-role` escriba datos en el dominio denominado `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/*"
        }
    ]
```

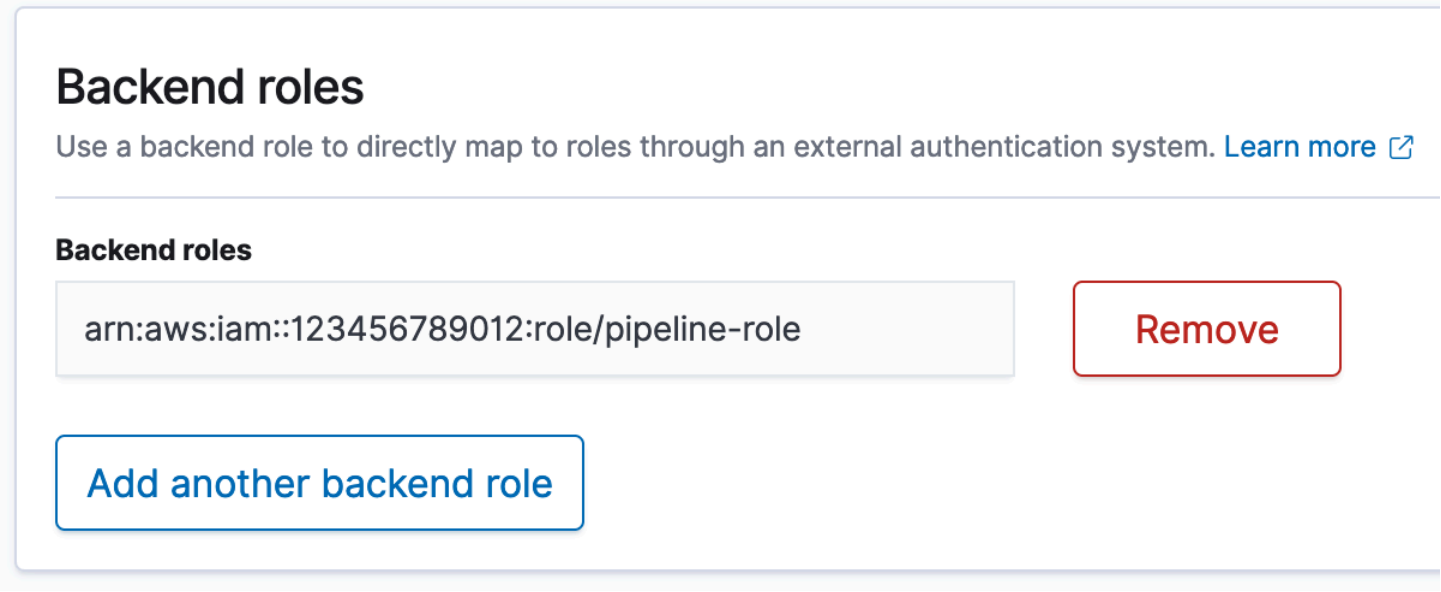
```
}
```

Asignar el rol de canalización (solo para dominios que utilicen un control de acceso detallado)

Si su dominio utiliza un control de acceso detallado para la autenticación, debe seguir algunos pasos adicionales para proporcionar a su canalización acceso a un dominio. Los pasos varían en función de la configuración del dominio:

- Escenario 1: función de maestro y función de canalización diferentes: si utiliza un nombre de recurso de Amazon (ARN) de IAM como usuario maestro y es diferente de la función de canalización, debe asignar la función de canalización a la función de backend OpenSearch`all_access`. Esto agrega el rol de canalización como usuario maestro adicional. Para obtener más información, consulte Usuarios maestros adicionales.
- Escenario 2: Usuario maestro en la base de datos de usuarios interna: si tu dominio usa un usuario maestro en la base de datos de usuarios interna y la autenticación básica HTTP para los OpenSearch paneles, no puedes pasar el nombre de usuario y la contraseña maestros directamente a la configuración de la canalización. En su lugar, asigne la función de canalización a la función de OpenSearch `all_access` backend. Esto agrega el rol de canalización como usuario maestro adicional. Para obtener más información, consulte Usuarios maestros adicionales.
- Escenario 3: el mismo rol de maestro y rol de canalización (poco común): si utiliza un ARN de IAM como usuario maestro y es el mismo ARN que utiliza como rol de canalización, no necesita realizar ninguna otra acción. La canalización tiene los permisos necesarios para escribir en el dominio. Este escenario es poco común porque la mayoría de los entornos utilizan un rol de administrador o algún otro rol como rol de maestro.

La siguiente imagen muestra cómo asignar el rol de canalización a un rol de backend:



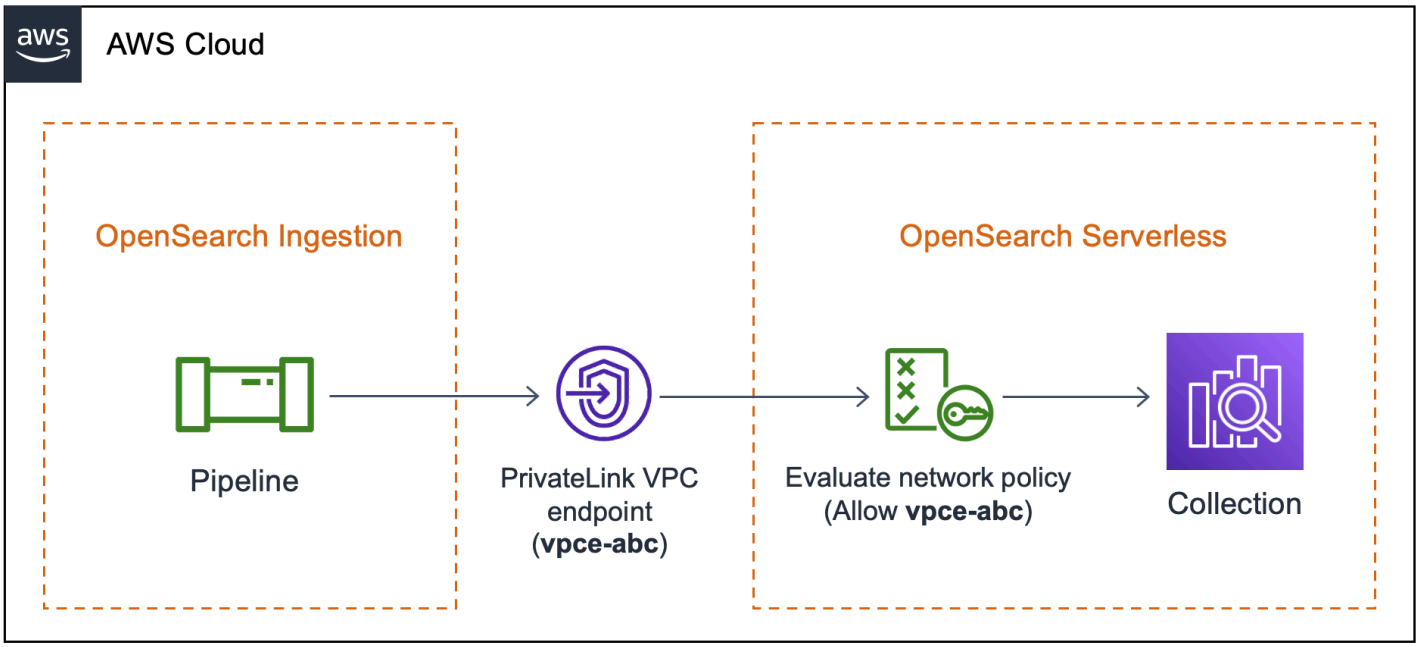
Otorgar a Amazon OpenSearch Ingestion pipelines acceso a las colecciones

Una canalización OpenSearch de Amazon Ingestion puede escribir en una colección pública OpenSearch sin servidor o en una colección de VPC. Para proporcionar acceso a la colección, debe configurar una función de canalización AWS Identity and Access Management (IAM) con una política de permisos que conceda el acceso a la colección. La canalización asume esta función para firmar las solicitudes que se envían al receptor de recopilación OpenSearch Serverless.

> **Important**
>
> Puedes elegir crear el rol de canalización manualmente o puedes hacer que OpenSearch Ingestion lo cree por ti durante la creación del canalización. Si eliges la creación automática de roles, OpenSearch Ingestion añadirá todos los permisos necesarios a la política de acceso a los roles de canalización en función de la fuente y el receptor que elijas. Crea un rol de canalización en IAM con el prefijo `OpenSearchIngestion-` y el sufijo que introduzca. Para obtener más información, consulte the section called “Rol de canalización”.
> Si ha pedido a OpenSearch Ingestion que cree el rol de canalización por usted, igual tendrá que incluirlo en la política de acceso a los datos de la recopilación, ya sea antes o después de crear el canalización. Para obtener instrucciones, consulte el paso 2.

Durante la creación de la canalización, OpenSearch Ingestion crea una AWS PrivateLink conexión entre la canalización y la colección OpenSearch Serverless. Todo el tráfico de la canalización pasa

por este punto de conexión de VPC y se enruta a la colección. Para llegar a la colección, el punto de conexión debe tener acceso a la colección mediante una política de acceso a la red.



Temas

- Paso 1: crear el rol de canalización
- Paso 2: Configurar el acceso a los datos y a la red para la colección

Paso 1: crear el rol de canalización

El rol debe tener una política de permisos adjunta que le permita enviar datos al receptor de la colección. También debe tener una relación de confianza que permita a OpenSearch Ingestion asumir el rol. Para ver instrucciones sobre cómo adjuntar una política a un rol, consulte Adición de permisos de identidad de IAM en la Guía del usuario de IAM.

En el siguiente ejemplo de política, se muestra el privilegio mínimo que se puede proporcionar en la política de acceso al rol de una canalización para escribir en las colecciones:

JSON

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

```
            "Sid": "Statement1",
            "Effect": "Allow",
            "Action": [
                "aoss:APIAccessAll",
                "aoss:BatchGetCollection",
                "aoss:CreateSecurityPolicy",
                "aoss:GetSecurityPolicy",
                "aoss:UpdateSecurityPolicy"
            ],
            "Resource": "*"
        }
    ]
}
```

El rol debe tener la siguiente relación de confianza, que permita a OpenSearch Ingestion asumirlo:

JSON

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

Paso 2: Configurar el acceso a los datos y a la red para la colección

Cree una colección OpenSearch sin servidor con la siguiente configuración. Para ver instrucciones sobre cómo crear una colección, consulte the section called “Creación de colecciones”.

Política de acceso a los datos

Cree una política de acceso a los datos para la colección que conceda los permisos necesarios para el rol de canalización. Por ejemplo:

```
[
  {
    "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**
>
> En el elemento `Principal`, debe especificar el nombre de recurso de Amazon (ARN) de un rol de canalización.

Política de acceso a la red

Cada colección que cree en OpenSearch Serverless tiene asociada al menos una política de acceso a la red. Las políticas de acceso a la red determinan si se puede acceder a la colección a través de Internet desde redes públicas o si se debe acceder a ella de forma privada. Para obtener más información sobre las políticas de red, consulte the section called “Acceso a la red”.

Dentro de una política de acceso a la red, solo puede especificar puntos finales de OpenSearch VPC administrados sin servidor. Para obtener más información, consulte the section called “Punto final de VPC del plano de datos”. Sin embargo, para que la canalización escriba en la colección, la política también debe conceder acceso al punto final de la VPC que OpenSearch Ingestion crea automáticamente entre la canalización y la colección. Por lo tanto, si eliges una colección

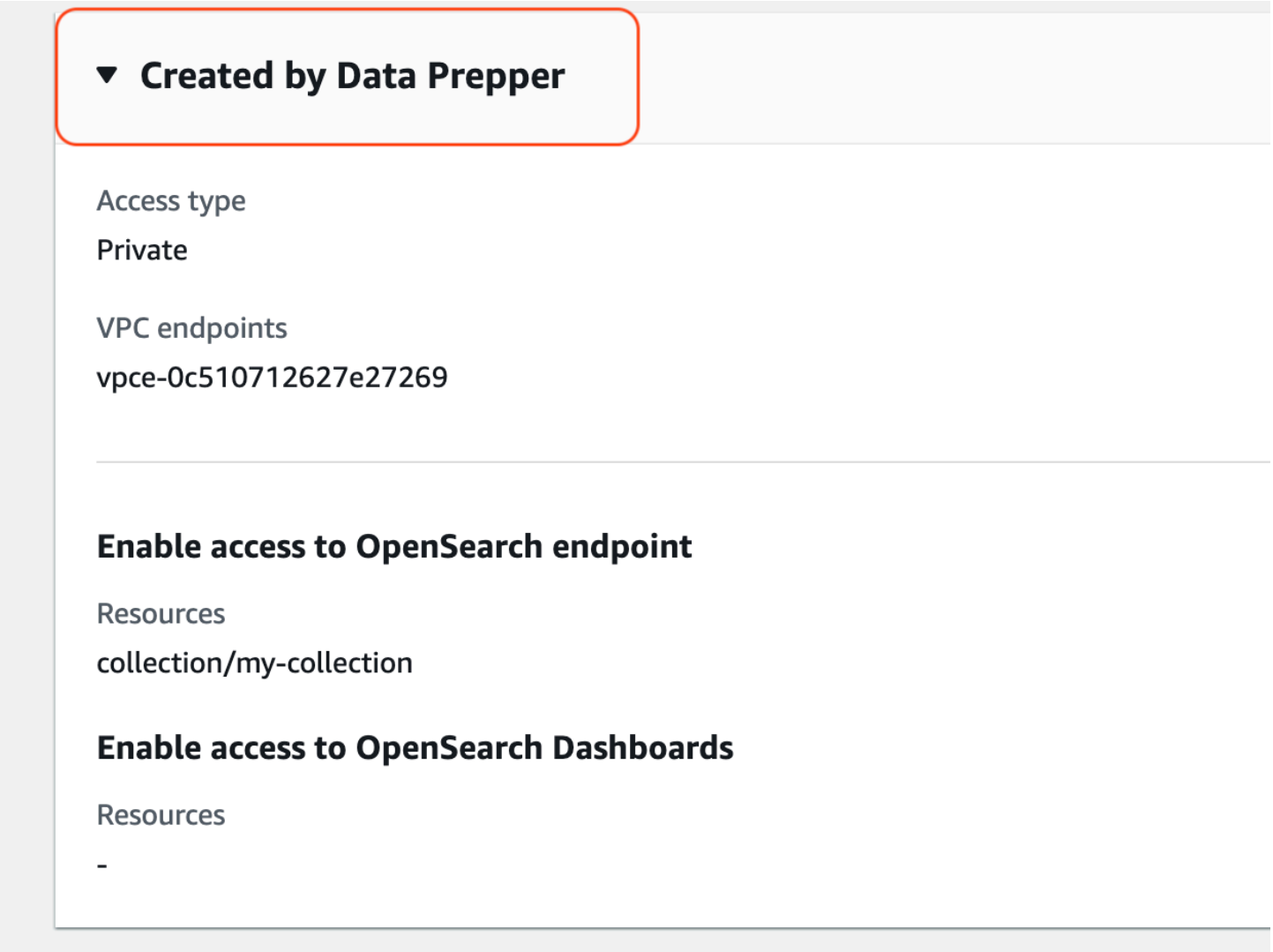
OpenSearch sin servidor como destino de una canalización, debes introducir el nombre de la política de red asociada en el campo Nombre de la política de red.

Durante la creación de la canalización, OpenSearch Ingestion comprueba la existencia de la política de red especificada. Si no existe, OpenSearch Ingestion la crea. Si existe, OpenSearch Ingestión la actualiza añadiéndole una nueva regla. La regla concede acceso al punto de conexión de VPC que conecta la canalización y la colección.

Por ejemplo:

```
{
   "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"
}
```

En la consola, cualquier regla que OpenSearch Ingestion añada a las políticas de red se denomina Creada por Data Prepper:



> **Note**
>
> En general, una regla que especifica el acceso público a una colección anula una regla que especifica el acceso privado. Por lo tanto, si la política ya tenía configurado el acceso público, esta nueva regla que añade OpenSearch Ingestion en realidad no cambia el comportamiento de la política. Para obtener más información, consulte the section called "Prioridad política".

Si detiene o elimina la canalización, OpenSearch Ingestión elimina el punto final de la VPC entre la canalización y la colección. También modifica la política de red para eliminar el punto de conexión de VPC de la lista de puntos de conexión permitidos. Si reinicia la canalización, se vuelve a crear el punto de conexión de VPC y se vuelve a actualizar la política de red con el ID del punto de conexión.

Primeros pasos con Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion admite la ingesta de datos en dominios de OpenSearch servicios gestionados y colecciones OpenSearch sin servidor. Los tutoriales siguientes muestran los pasos básicos para poner en marcha una canalización.

El primer tutorial muestra cómo utilizar Amazon OpenSearch Ingestion para configurar una canalización sencilla e ingerir datos en un dominio de Amazon OpenSearch Service.

El segundo tutorial muestra cómo utilizar Amazon OpenSearch Ingestion para configurar una canalización sencilla e ingerir datos en una colección de Amazon OpenSearch Serverless.

> **Note**
>
> La creación de canalizaciones fallará si no configura los permisos correctos. Consulte the section called “Configuración de roles y usuarios” para comprender mejor las funciones necesarias antes de crear una canalización.

Temas

- Tutorial: Ingerir datos en un dominio mediante Amazon OpenSearch Ingestion
- Tutorial: Ingerir datos en una colección mediante Amazon OpenSearch Ingestion

Tutorial: Ingerir datos en un dominio mediante Amazon OpenSearch Ingestion

En este tutorial, se muestra cómo utilizar Amazon OpenSearch Ingestion para configurar una canalización sencilla e ingerir datos en un dominio de Amazon OpenSearch Service. Una canalización es un recurso que OpenSearch Ingestion aprovisiona y administra. Puede usar una canalización para filtrar, enriquecer, transformar, normalizar y agregar datos para el análisis y la visualización posteriores en OpenSearch Service.

Este tutorial recorre los pasos básicos para poner en marcha una canalización de forma rápida. Para obtener instrucciones más completas, consulte the section called “Creación de canalizaciones”.

En este tutorial, deberá completar los siguientes pasos:

1. Crear un dominio.

2. Crear una canalización.
3. Incorporar algunos datos de muestra.

En este tutorial creará los recursos siguientes:

- Un dominio llamado `ingestion-domain` al que la canalización escribe.
- Una canalización llamada `ingestion-pipeline`

Permisos necesarios

Para completar este tutorial, su usuario o rol debe tener una política basada en la identidad adjunta con los siguientes permisos mínimos. Estos permisos le permiten crear un rol de canalización y adjuntar una política (`iam:Create*` y `iam:Attach*`), crear o modificar un dominio (`es:*`) y trabajar con canalizaciones (`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"
         ]
```

```
      }
   ]
}
```

Paso 1: crear el rol de canalización

En primer lugar, cree una función que asuma la canalización para acceder al receptor del dominio OpenSearch de Service. Incluirá esta función en la configuración de la canalización más adelante en este tutorial.

Cómo crear el rol de canalización

1. Abre la AWS Identity and Access Management consola en https://console.aws.amazon.com/iamv2/.
2. Seleccione Políticas y, a continuación, seleccione Crear política.
3. En este tutorial, incorporará datos a un dominio llamado `ingestion-domain`, que creará en el siguiente paso. Seleccione JSON y pegue la siguiente política en el editor. Sustituya *`your-account-id`* por su Identificador de cuenta y modifique la región si es necesario.

   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/*"
        }
    ]
}
```

Si desea escribir datos en un dominio existente, sustituya `ingestion-domain` por el nombre de su dominio.

> **Note**
>
> Para simplificar este tutorial, utilizaremos una política de acceso amplia. No obstante, en entornos de producción, se recomienda que se aplique una política de acceso más restrictiva a su rol de canalización. Para ver un ejemplo de política que proporciona los permisos mínimos necesarios, consulte the section called “Concesión de acceso a dominios para canalizaciones”.

4. Seleccione Siguiente, seleccione Siguiente y asigne a su política el nombre pipeline-policy.
5. Seleccione Crear política.
6. A continuación, cree un rol y añádale la política. Seleccione Roles y, a continuación, Crear rol.
7. Seleccione Política de confianza personalizada y pegue la siguiente política en el editor:

   JSON

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

8. Seleccione Siguiente. A continuación, busque y seleccione pipeline-policy (que acaba de crear).
9. Seleccione Siguiente y asigne un nombre al rol PipelineRole.
10. Elija Crear rol.

Recuerde el nombre de recurso de Amazon (ARN) del rol (por ejemplo, `arn:aws:iam::your-account-id:role/PipelineRole`). Lo necesitará cuando cree la canalización.

Paso 2: crear un dominio

Primero, cree un dominio denominado `ingestion-domain` al que desee incorporar datos.

Ve a la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa y crea un dominio que cumpla los siguientes requisitos:

- Está ejecutando la OpenSearch versión 1.0 o posterior, o Elasticsearch 7.4 o una versión posterior
- Utiliza el acceso público
- No utiliza el control de acceso detallado

> **Note**
>
> Estos requisitos están pensados para garantizar la simplicidad de este tutorial. En los entornos de producción, puede configurar un dominio con acceso a la VPC y and/or utilizar un control de acceso detallado. Para utilizar un control de acceso detallado, consulte Map the pipeline role.

El dominio debe tener una política de acceso que conceda permisos a la función de `OpenSearchIngestion-PipelineRole` IAM, que el OpenSearch Servicio creará para usted en el siguiente paso. La canalización asumirá este rol para enviar datos al receptor de dominios.

Asegúrese de que el dominio tenga la siguiente política de acceso a nivel de dominio, que otorga al rol de canalización acceso al dominio. Sustituya la Región y el ID de cuenta por los suyos:

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/*"
    }
  ]
}
```

Para obtener más información sobre cómo crear políticas de acceso a nivel de dominio, consulte the section called “Políticas basadas en recursos”.

Si ya ha creado un dominio, modifique su política de acceso actual para proporcionar los permisos anteriores a `OpenSearchIngestion-PipelineRole`.

Paso 3: crear una canalización

Ahora que tiene un dominio, puede crear una canalización.

Creación de una canalización

1. En la consola de Amazon OpenSearch Service, selecciona Pipelines en el panel de navegación izquierdo.
2. Seleccione Create pipeline.
3. Seleccione la canalización en blanco y, a continuación, elija Seleccionar esquema.
4. En este tutorial, crearemos una canalización simple que utiliza el complemento de origen HTTP. El complemento acepta datos de registro en un formato de matriz JSON. Especificaremos un único dominio OpenSearch de servicio como receptor e incorporaremos todos los datos al `application_logs` índice.

   En el menú Fuente, seleccione HTTP. Para la ruta, ingrese /logs.
5. Para simplificar este tutorial, configuraremos el acceso público para la canalización. En las opciones de red de origen, seleccione Acceso público. Para obtener información acerca de la configuración del acceso a VPC, consulte the section called “Configuración del acceso a la VPC para canalizaciones”.
6. Elija Siguiente.
7. En Procesador, introduzca la fecha y seleccione Agregar.
8. Habilite Desde la hora de recepción. Deje el resto de opciones con sus valores predeterminados.
9. Elija Siguiente.

10. Configure los detalles del destino. Para el tipo OpenSearch de recurso, elija Clúster gestionado. A continuación, elija el dominio de OpenSearch servicio que creó en la sección anterior.

    En Nombre de índice, ingresa application_logs. OpenSearch La ingestión crea automáticamente este índice en el dominio si aún no existe.

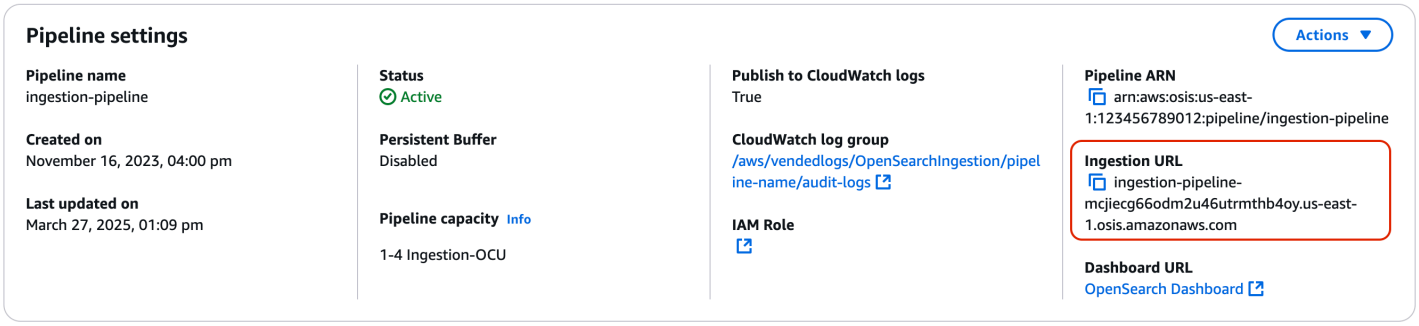
11. Elija Siguiente.
12. Asigne a la canalización el nombre ingestion-pipeline. Deje la configuración de capacidad con sus valores predeterminados.
13. Para el rol de canalización, seleccione Crear y usar un nuevo rol de servicio. El rol de canalización proporciona los permisos necesarios para que una canalización escriba en el destino del dominio y lea desde orígenes basados en extracción. Al seleccionar esta opción, permite que OpenSearch Ingestion cree el rol por usted, en lugar de crearlo manualmente en IAM. Para obtener más información, consulte the section called “Configuración de roles y usuarios”.
14. En el sufijo del nombre de la función de servicio, introduzca. PipelineRole En IAM, el rol tendrá el formato `arn:aws:iam::`*`your-account-id`*`:role/OpenSearchIngestion-`**`PipelineRole`**.
15. Elija Siguiente. Revise la configuración de la canalización y elija Crear canalización. La canalización tarda entre 5 y 10 minutos en activarse.

Paso 4: incorporar algunos datos de muestra

Cuando el estado de la canalización sea `Active`, puede empezar a incorporarle datos. Debe firmar todas las solicitudes HTTP que se envíen a la canalización mediante la versión 4 de Signature. Use una herramienta HTTP como Postman o awscurl para enviar algunos datos a la canalización. Al igual que ocurre con la indexación de datos directamente en un dominio, la incorporación de datos a una canalización siempre requiere un rol de IAM o una clave de acceso y una clave secreta de IAM.

> **Note**
>
> La entidad principal que firma la solicitud debe tener el permiso de IAM `osis:Ingest`.

Primero, obtenga la URL de incorporación en la página de Configuración de canalización:



A continuación, incorpore algunos datos de muestra. La siguiente solicitud utiliza awscurl para enviar un único archivo de registro a la canalización:

```
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
```

Debería ver una respuesta `200 OK`. Si recibe un error de autenticación, es posible que se deba a que está incorporando datos de una cuenta distinta a la de la canalización. Consulte the section called “Solución de problemas de permisos”.

Ahora, consulte el índice `application_logs` para asegurarse de que la entrada de registro se haya incorporado correctamente:

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

Respuesta de ejemplo:

```
{
   "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"
            }
         }
      ]
   }
}
```


Solución de problemas de permisos

Si has seguido los pasos del tutorial y sigues viendo errores de autenticación cuando intentas ingerir datos, puede que se deba a que la función que escribe en una canalización Cuenta de AWS no es la propia canalización. En este caso, debe crear y asumir un rol que le permita específicamente incorporar datos. Para obtener instrucciones, consulte the section called “Provisión de acceso de ingesta entre cuentas”.

Recursos relacionados

Este tutorial presenta un caso práctico sencillo de incorporación de un único documento a través de HTTP. En escenarios de producción, configurará las aplicaciones cliente (como Fluent Bit, Kubernetes o OpenTelemetry Collector) para que envíen datos a una o más canalizaciones. Es probable que sus canalizaciones sean más complejas que en el ejemplo sencillo de este tutorial.

Para empezar a configurar sus clientes e incorporar datos, consulte los siguientes recursos:

- Creación y administración de canalizaciones
- Configurar sus clientes para que envíen datos a Ingestion OpenSearch
- Documentación de Data Prepper

Tutorial: Ingerir datos en una colección mediante Amazon OpenSearch Ingestion

En este tutorial, se muestra cómo utilizar Amazon OpenSearch Ingestion para configurar una canalización sencilla e ingerir datos en una colección de Amazon OpenSearch Serverless. Una canalización es un recurso que OpenSearch Ingestion aprovisiona y administra. Puede usar una canalización para filtrar, enriquecer, transformar, normalizar y agregar datos para el análisis y la visualización posteriores en OpenSearch Service.

Para ver un tutorial que muestra cómo incorporar datos en un dominio de OpenSearch servicio aprovisionado, consulte. the section called “Tutorial: incorporar datos a un dominio”

En este tutorial, deberá completar los siguientes pasos:

1. Crear una colección.
2. Crear una canalización.
3. Incorporar algunos datos de muestra.

En este tutorial creará los recursos siguientes:

- Una colección llamada `ingestion-collection` donde escribirá la canalización
- Una canalización llamada `ingestion-pipeline-serverless`

Permisos necesarios

Para completar este tutorial, su usuario o rol debe tener una política basada en identidad adjunta con los siguientes permisos mínimos. Estos permisos le permiten crear un rol de canalización y adjuntar una política (`iam:Create*` y `iam:Attach*`), crear o modificar una colección (`aoss:*`) y trabajar con canalizaciones (`osis:*`).

Además, se requieren varios permisos de IAM para crear automáticamente la función de canalización y pasarla a OpenSearch Ingestion para que pueda escribir datos en la colección.

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"
         ]
      }
   ]
}
```

Paso 1: Crear una colección

A continuación, cree una colección para incorporar datos. El nombre que le pondremos a la colección será `ingestion-collection`.

1. Ve a la consola de Amazon OpenSearch Service de https://console.aws.amazon.com/aos/tucasa.
2. Seleccione Colecciones del panel de navegación de la izquierda y elija Crear colección.

3. Nombre la colección ingestion-collection.
4. En Seguridad, elija Creación estándar.
5. En Configuración de acceso a la red, cambie el tipo de acceso a Público.
6. Mantenga todas las demás configuraciones en sus valores predeterminados y elija Siguiente.
7. Ahora, configure una política de acceso a los datos para la colección. Anule la selección de Coincidir automáticamente la configuración de la política de acceso.
8. Seleccione JSON para el Método de definición y pegue la siguiente política en el editor. Esta política realiza dos tareas:

   - Permite que el rol de canalización escriba en la colección.
   - Le permite leer de la colección. Más adelante, después de incorporar algunos datos de muestra a la canalización, consultará la colección para garantizar que los datos se hayan incorporado y escrito correctamente en el índice.

```
[
  {
    "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 los `Principal` elementos para incluir tu Cuenta de AWS ID. Para la segunda entidad principal, especifique un usuario o rol que pueda utilizar para consultar la colección más tarde.
10. Elija Siguiente. Asigne un nombre a la política de acceso pipeline-collection-accessy vuelva a seleccionar Siguiente.
11. Revise la configuración de la colección y seleccione Enviar.

Paso 2: Crear una canalización

Ahora que tiene una colección, puede crear una canalización.

Creación de una canalización

1. En la consola de Amazon OpenSearch Service, selecciona Pipelines en el panel de navegación izquierdo.
2. Seleccione Create pipeline.
3. Seleccione la canalización en blanco y, a continuación, elija Seleccionar esquema.
4. En este tutorial, crearemos una canalización simple que utiliza el complemento de origen HTTP. El complemento acepta datos de registro en un formato de matriz JSON. Especificaremos una única colección OpenSearch sin servidor como sumidero e incorporaremos todos los datos al índice. `my_logs`

   En el menú Fuente, seleccione HTTP. Para la ruta, ingrese /logs.
5. Para simplificar este tutorial, configuraremos el acceso público para la canalización. En las opciones de red de origen, seleccione Acceso público. Para obtener información acerca de la configuración del acceso a VPC, consulte the section called “Configuración del acceso a la VPC para canalizaciones”.
6. Elija Siguiente.
7. En Procesador, introduzca la fecha y seleccione Agregar.
8. Habilite Desde la hora de recepción. Deje el resto de opciones con sus valores predeterminados.
9. Elija Siguiente.
10. Configure los detalles del destino. Para el tipo de OpenSearch recurso, elija Colección (sin servidor). A continuación, elija la colección de OpenSearch servicios que creó en la sección anterior.

Deje el nombre de la política de red como predeterminado. En Nombre de índice, escribe my_logs. OpenSearch Ingestión crea automáticamente este índice en la colección si aún no existe.

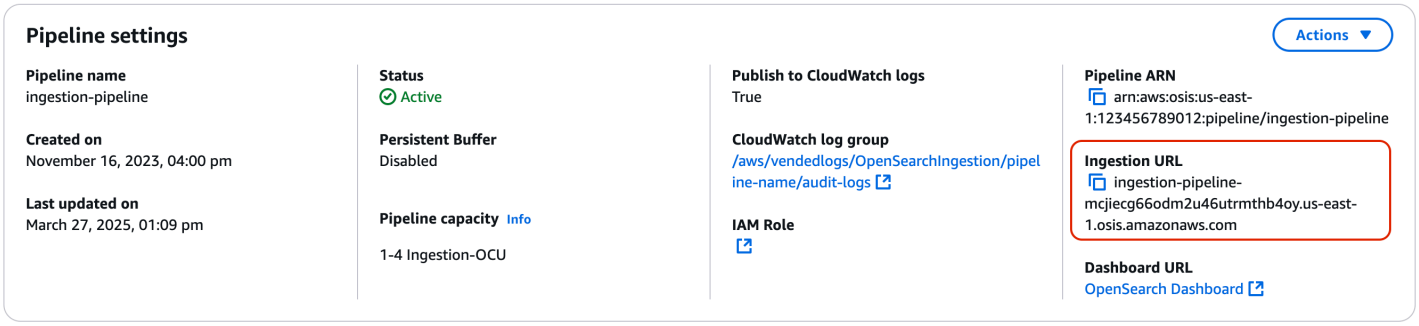
11. Elija Siguiente.
12. Asigne un nombre a la canalización ingestion-pipeline-serverless. Deje la configuración de capacidad con sus valores predeterminados.
13. Para el rol de canalización, seleccione Crear y usar un nuevo rol de servicio. El rol de canalización proporciona los permisos necesarios para que una canalización escriba en el destino de colección y lea desde orígenes basados en extracción. Al seleccionar esta opción, permite que OpenSearch Ingestion cree el rol por usted, en lugar de crearlo manualmente en IAM. Para obtener más información, consulte the section called “Configuración de roles y usuarios”.
14. En el sufijo del nombre del rol de servicio, introduzca. PipelineRole En IAM, el rol tendrá el formato `arn:aws:iam::`*`your-account-id`*`:role/OpenSearchIngestion-`**`PipelineRole`**.
15. Elija Siguiente. Revise la configuración de la canalización y elija Crear canalización. La canalización tarda entre 5 y 10 minutos en activarse.

Paso 3: Ingiera algunos datos de muestra

Cuando el estado de la canalización sea `Active`, puede empezar a incorporarle datos. Debe firmar todas las solicitudes HTTP que se envíen a la canalización mediante la versión 4 de Signature. Use una herramienta HTTP como Postman o awscurl para enviar algunos datos a la canalización. Al igual que ocurre con la indexación de datos directamente en una colección, la incorporación de datos a una canalización siempre requiere un rol de IAM o una clave de acceso y una clave secreta de IAM.

> **Note**
>
> La entidad principal que firma la solicitud debe tener el permiso de IAM `osis:Ingest`.

Primero, obtenga la URL de incorporación en la página de Configuración de canalización:



A continuación, envíe algunos datos de muestra a la ruta de ingesta. La siguiente solicitud de muestra utiliza awscurl para enviar un único archivo de registro a la canalización:

```
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
```

Debería ver una respuesta `200 OK`.

Ahora, consulte el índice `my_logs` para asegurarse de que la entrada del registro se haya incorporado correctamente:

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

Respuesta de ejemplo:

```
{
   "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"
            }
         }
      ]
   }
}
```


Recursos relacionados

Este tutorial presenta un caso práctico sencillo de incorporación de un único documento a través de HTTP. En los escenarios de producción, configurará las aplicaciones cliente (como Fluent Bit, Kubernetes o OpenTelemetry Collector) para que envíen datos a una o más canalizaciones. Es probable que sus canalizaciones sean más complejas que en el ejemplo sencillo de este tutorial.

Para empezar a configurar sus clientes e incorporar datos, consulte los siguientes recursos:

- Creación y administración de canalizaciones
- Configurar sus clientes para que envíen datos a Ingestion OpenSearch
- Documentación de Data Prepper

Descripción general de las funciones de canalización en Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion aprovisiona las canalizaciones, que constan de una fuente, un búfer, cero o más procesadores y uno o más receptores. Las canalizaciones de incorporación funcionan con Data Prepper como motor de datos. Para obtener información general sobre los distintos componentes de una canalización, consulte the section called “Conceptos clave”.

En las siguientes secciones se ofrece información general sobre algunas de las funciones más utilizadas en Amazon OpenSearch Ingestion.

> **Note**
>
> No es una lista exhaustiva de las características disponibles para las canalización. Para obtener una documentación completa de todas las funciones de canalización disponibles, consulte Documentación de Data Prepper. Tenga en cuenta que OpenSearch Ingestion impone algunas restricciones a los complementos y las opciones que puede utilizar. Para obtener más información, consulte the section called “Complementos y opciones compatibles”.

Temas

- Almacenamiento en búfer persistente
- División
- Encadenar
- Colas de mensajes fallidos
- Administración de índices
- End-to-end reconocimiento
- Contrapresión del origen

Almacenamiento en búfer persistente

Un búfer persistente almacena los datos en un búfer basado en disco en varias zonas de disponibilidad para mejorar la durabilidad de los datos. Puede utilizar el almacenamiento en búfer

persistente para ingerir datos de todos los orígenes basados en push compatibles sin configurar un búfer independiente. Estas fuentes incluyen HTTP y OpenTelemetry para registros, rastreos y métricas. Para habilitar el almacenamiento en búfer persistente, seleccione Habilitar búfer persistente al crear o actualizar una canalización. Para obtener más información, consulte the section called "Creación de canalizaciones".

OpenSearch La ingesta determina de forma dinámica la cantidad de datos que se OCUs van a utilizar para el almacenamiento en búfer persistente, teniendo en cuenta la fuente de datos, las transformaciones de transmisión y el destino del receptor. Como asigna una parte OCUs al almacenamiento en búfer, es posible que tengas que aumentar el mínimo y el máximo para mantener el mismo rendimiento de ingestión. OCUs Las canalizaciones conservan los datos en el búfer hasta 72 horas.

Si habilita el almacenamiento en búfer persistente para una canalización, los tamaños máximos de carga útil de solicitud predeterminados son los siguientes:

- Orígenes HTTP: 10 MB
- OpenTelemetry fuentes: 4 MB

En el caso de los orígenes HTTP, puede aumentar el tamaño máximo de carga útil a 20 MB. El tamaño de la carga útil de la solicitud incluye toda la solicitud HTTP, que normalmente contiene varios eventos. Cada evento no puede superar los 3,5 MB.

Las canalizaciones con el almacenamiento en búfer persistente dividen las unidades de canalización configuradas en unidades de computación y unidades de búfer. Si una canalización utiliza un procesador con un uso intensivo de la CPU, como grok, key-value o split string, asigna las unidades en una proporción de 1:1. buffer-to-compute De lo contrario, las asigna en una proporción de 3:1, lo que siempre favorece las unidades de cómputo.

Por ejemplo:

- Canalización con grok y 2 unidades como máximo: 1 unidad de computación y 1 unidad de búfer
- Canalización con grok y 5 unidades como máximo: 3 unidades de computación y 2 unidades de búfer
- Canalización sin procesadores y 2 unidades como máximo: 1 unidad de computación y 1 unidad de búfer
- Canalización sin procesadores y 4 unidades como máximo: 1 unidad de computación y 3 unidades de búfer

- Canalización con grok y 5 unidades como máximo: 2 unidades de computación y 3 unidades de búfer

De forma predeterminada, las canalizaciones utilizan y para cifrar los datos del búfer. Clave propiedad de AWS Estas canalizaciones no necesitan ningún permiso adicional para el rol de canalización.

Como alternativa, se puede especificar una clave administrada por el cliente y agregar los siguientes permisos de IAM al rol de canalización:

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"
        }
    ]
}
```

Para más información, consulte las claves administradas por el cliente en la Guía para desarrolladores de AWS Key Management Service .

> **Note**
>
> Si deshabilita el almacenamiento en búfer persistente, su canalización se ejecuta por completo con el almacenamiento en búfer en memoria.

División

Puede configurar una canalización de OpenSearch ingestión para dividir los eventos entrantes en una subcanalización, lo que le permitirá realizar diferentes tipos de procesamiento en el mismo evento entrante.

El siguiente ejemplo de canalización divide los eventos entrantes en dos subcanalizaciones. Cada subcanalización utiliza su propio procesador para enriquecer y manipular los datos y, a continuación, los envía a distintos índices. 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"
```

Encadenar

Puede encadenar varias subcanalizaciones para procesar y enriquecer los datos por partes. En otras palabras, puede enriquecer un evento entrante con determinadas capacidades de procesamiento en una subcanalización, luego enviarlo a otra subcanalización para enriquecerla aún más con un procesador diferente y, finalmente, enviarlo a su receptor. OpenSearch

En el siguiente ejemplo, la `log_pipeline` subcanalización enriquece un evento de registro entrante con un conjunto de procesadores y, a continuación, envía el evento a un índice denominado. OpenSearch `enriched_logs` La canalización envía el mismo evento a la `log_advanced_pipeline` subcanalización, que lo procesa y lo envía a un índice diferente OpenSearch denominado. `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"
```

Colas de mensajes fallidos

Las colas con letras muertas (DLQs) son los destinos de los eventos que una canalización no logra registrar en un receptor. En OpenSearch Ingestion, debe especificar un bucket de Amazon S3 con los permisos de escritura adecuados para usarlo como DLQ. Puede añadir una configuración de DLQ a cada receptor de una canalización. Cuando una canalización detecta errores de escritura, crea objetos de DLQ en el bucket de S3 configurado. Los objetos DLQ existen dentro de un archivo JSON como una matriz de eventos fallidos.

Una canalización escribe eventos en la DLQ cuando se cumple alguna de las siguientes condiciones:

- Se agotó el número máximo de reintentos del OpenSearch sumidero. OpenSearch La ingestión requiere un mínimo de 16 para esta configuración.
- El receptor rechaza los eventos debido a una condición de error.

Configuración

Para configurar una cola de mensajes fallidos para una subcanalización, elija Habilitar DLQ de S3 al configurar el destino de su receptor. A continuación, especifique los ajustes necesarios para la cola. Para obtener más información, consulte Configuración en la documentación de Data Prepper DLQ.

Los archivos que se escriben en este DLQ de S3 tienen el siguiente patrón de nomenclatura:

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

Para obtener instrucciones acerca de cómo configurar manualmente el rol de canalización para permitir el acceso al bucket de S3 en el que escribe la DLQ, consulte the section called “Permisos para escribir en Amazon S3 o en una cola de mensajes fallidos”.

Ejemplo

Considere el siguiente ejemplo de archivo DLQ:

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

Este es un ejemplo de datos que no se pudieron escribir en el receptor y que se envían al bucket DLQ de S3 para su posterior análisis:

```
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"
```

Administración de índices

Amazon OpenSearch Ingestion cuenta con numerosas funciones de administración de índices, entre las que se incluyen las siguientes.

Creación de índices

Puede especificar un nombre de índice en un colector de canalización e OpenSearch Ingestion crea el índice cuando aprovisiona la canalización. Si ya existe un índice, la canalización lo usa para indexar los eventos entrantes. Si detiene y reinicia una canalización, o si actualiza su configuración de YAML, la canalización intentará crear nuevos índices si aún no existen. Una canalización nunca puede eliminar un índice.

En el siguiente ejemplo, los receptores crean dos índices cuando se aprovisiona la canalización:

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

Creación de nombres y patrones de índice

Puede generar nombres de índice dinámicos mediante variables de los campos de los eventos entrantes. En la configuración del receptor, utilice el formato `string${}` para indicar la interpolación de cadenas y utilice un puntero JSON para extraer los campos de los eventos. Las opciones para `index_type` son `custom` o `management_disabled`. Como el `index_type` valor predeterminado es `custom` para los OpenSearch dominios y `management_disabled` para las colecciones OpenSearch sin servidor, se puede dejar sin configurar.

Por ejemplo, la siguiente canalización selecciona el campo `metadataType` entre los eventos entrantes para generar nombres de índice.

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

La siguiente configuración sigue generando un índice nuevo cada día o cada hora.

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

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

El nombre del índice también puede ser una cadena simple con un patrón de fecha y hora como sufijo, por ejemplo `my-index-%{yyyy.MM.dd}`. Cuando el receptor envía datos a OpenSearch, reemplaza el patrón de fecha y hora por la hora UTC y crea un índice nuevo para cada día, por ejemplo. `my-index-2022.01.25` Para obtener más información, consulta la DateTimeFormatterclase.

Este nombre de índice también puede ser una cadena formateada (con o sin un sufijo de patrón de fecha y hora), como `my-${index}-name`. Cuando el receptor envía datos a OpenSearch, reemplaza la `"${index}"` parte por el valor en caso de que se esté procesando. Si el formato lo es `"${index1/index2/index3}"`, reemplaza el campo `index1/index2/index3` por su valor en el evento.

Generando documento IDs

Una canalización puede generar un identificador de documento al indexar OpenSearch los documentos en él. Puede deducir estos documentos a IDs partir de los campos incluidos en los eventos entrantes.

En este ejemplo, se utiliza el campo `uuid` de un evento entrante para generar un identificador de documento.

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

En el ejemplo siguiente, el procesador Add entries combina los campos `uuid` y `other_field` del evento entrante para generar un identificador de documento.

La `create` acción garantiza que los documentos idénticos IDs no se sobrescriban. La canalización elimina los documentos duplicados sin ningún tipo de reintento ni ningún evento de DLQ. Esta es una

expectativa razonable para los autores de canalización que utilizan esta acción, ya que el objetivo es evitar actualizar los documentos existentes.

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

Es posible que desee establecer el identificador del documento de un evento en un campo de un subobjeto. En el siguiente ejemplo, el complemento OpenSearch sink utiliza el subobjeto `info/id` para generar un identificador de documento.

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

Si se produce el siguiente evento, la canalización generará un documento con el campo `_id` establecido en `json001`:

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

Generar enrutamiento IDs

Puedes usar la `routing_field` opción del complemento OpenSearch sink para establecer el valor de una propiedad de enrutamiento de documentos (`_routing`) en el valor de un evento entrante.

El enrutamiento admite la sintaxis de punteros JSON, por lo que también están disponibles campos anidados, no solo campos de nivel superior.

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

Si se produce el siguiente evento, el complemento genera un documento con el campo `_routing` establecido en `abcd`:

```
{
   "id":"123",
   "metadata":{
      "id":"abcd",
      "fieldA":"valueA"
   },
   "fieldB":"valueB"
}
```

Para obtener instrucciones sobre cómo crear plantillas de índice que las canalizaciones puedan utilizar durante la creación del índice, consulte Plantillas de índice.

End-to-end reconocimiento

OpenSearch La ingestión garantiza la durabilidad y la fiabilidad de los datos al rastrear su entrega desde su origen hasta los sumideros de los oleoductos sin estado mediante el uso de acuse de recibo. end-to-end

> **Note**
>
> Actualmente, solo el complemento fuente de S3 admite el reconocimiento. end-to-end

Con el end-to-end acuse de recibo, el complemento Pipeline Source crea un conjunto de acuses de recibo para monitorear un lote de eventos. Recibe un reconocimiento positivo cuando esos eventos se envían correctamente a sus receptores, o un reconocimiento negativo cuando alguno de los eventos no se ha podido enviar a sus receptores.

En caso de fallo o caída de un componente de la canalización, o si un origen no recibe un reconocimiento, se agota el tiempo de espera y toma las medidas necesarias, como volver a intentarlo o registrar el fallo. Si la canalización tiene configurados varios receptores o subcanalizaciones, las confirmaciones a nivel de evento solo se envían después de que el evento se envíe a todos los receptores de todas las subcanalizaciones. Si un receptor tiene un DLQ configurado, los acuses de end-to-end recibo también rastrean los eventos escritos en el DLQ.

Para activar la end-to-end confirmación, expanda Opciones adicionales en la configuración de origen de Amazon S3 y seleccione Habilitar la confirmación de end-to-end mensajes.

Contrapresión del origen

Una canalización puede sufrir una contrapresión cuando está ocupada procesando datos o si sus receptores están temporalmente inactivos o tardan en ingerir datos. OpenSearch La ingestión tiene diferentes formas de gestionar la contrapresión en función del complemento fuente que utilice la canalización.

Origen de HTTP

Las canalizaciones que utilizan el complemento origen de HTTP gestionan la contrapresión de forma diferente en función del componente de la canalización que esté congestionado:

- Búferes: cuando los búferes están llenos, la canalización comienza a devolver el estado HTTP `REQUEST_TIMEOUT` con el código de error 408 al punto de conexión de origen. A medida que se van liberando los búferes, la canalización vuelve a procesar los eventos HTTP.
- Subprocesos de origen: cuando todos los subprocesos de origen HTTP están ocupados ejecutando solicitudes y el tamaño de la cola de solicitudes sin procesar ha superado el número máximo de solicitudes permitido, la canalización comienza a devolver el estado HTTP `TOO_MANY_REQUESTS` con el código de error 429 al punto de conexión de origen. Cuando la cola de solicitudes cae por debajo del tamaño máximo permitido, la canalización vuelve a procesar las solicitudes.

OTel fuente

Cuando los búferes están llenos para las canalizaciones que utilizan OpenTelemetry fuentes (OTel registros, OTel métricas y OTel rastreo), la canalización comienza a devolver el estado HTTP `REQUEST_TIMEOUT` con el código de error 408 al punto final de origen. A medida que se van liberando los búferes, la canalización vuelve a procesar los eventos.

Fuente de S3

Cuando los búferes están llenos para las canalizaciones con un origen de S3, las canalizaciones dejan de procesar las notificaciones de SQS. A medida que se van liberando los búferes, la canalización vuelve a procesar las notificaciones.

Si un sumidero no funciona o no puede ingerir datos y el acuse de end-to-end recibo está activado para la fuente, la canalización deja de procesar las notificaciones de SQS hasta que reciba un acuse de recibo correcto de todos los sumideros.

Creación de canalizaciones OpenSearch de Amazon Ingestion

Una canalización es el mecanismo que Amazon OpenSearch Ingestion utiliza para mover los datos desde su origen (de donde provienen los datos) hasta su receptor (adonde van los datos). En OpenSearch Ingestion, el receptor siempre será un único dominio de Amazon OpenSearch Service, mientras que la fuente de los datos podrían ser clientes como Amazon S3, Fluent Bit o OpenTelemetry Collector.

Para obtener más información, consulte Pipelines en la OpenSearch documentación.

Temas

- Requisitos previos y rol de IAM requerido
- Permisos de IAM necesarios
- Especificar la versión de la canalización
- Especificación de la ruta de ingesta
- Creación de canalizaciones
- Seguimiento del estado de creación de la canalización
- Uso de esquemas

Requisitos previos y rol de IAM requerido

Para crear una canalización OpenSearch de ingestión, debes disponer de los siguientes recursos:

- Una función de IAM, denominada función de canalización, que OpenSearch Ingestión asume para poder trabajar en el receptor. Puedes crear esta función con antelación o puedes hacer que OpenSearch Ingestion la cree automáticamente mientras creas la canalización.

- Un dominio OpenSearch de servicio o una colección OpenSearch sin servidor que sirva de receptor. Si escribes en un dominio, debe ejecutar la OpenSearch versión 1.0 o una versión posterior, o Elasticsearch 7.4 o una versión posterior. El receptor debe tener una política de acceso que conceda los permisos adecuados a su rol de canalización de IAM.

Si desea obtener instrucciones para crear estos recursos, consulte los siguientes temas:

- the section called “Concesión de acceso a dominios para canalizaciones”
- the section called “Concesión de acceso a colecciones para canalizaciones”

> **Note**
>
> Si escribe en un dominio que usa un control de acceso detallado, debe completar algunos pasos adicionales. Consulte the section called “Asignar el rol de canalización (solo para dominios que utilicen un control de acceso detallado)”.

Permisos de IAM necesarios

OpenSearch Ingestion usa los siguientes permisos de IAM para crear canalizaciones:

- `osis:CreatePipeline`: Crear una canalización.
- `osis:ValidatePipeline`: Comprobar si la configuración de la canalización es válida.
- `iam:CreateRoleyiam:AttachPolicy`: haga que OpenSearch Ingestion cree automáticamente el rol de canalización por usted.
- `iam:PassRole`— Transfiera la función de canalización a OpenSearch Ingestion para que pueda escribir datos en el dominio. Este permiso debe estar en el recurso de rol de canalización, o simplemente coloque * si planea usar diferentes roles en cada canalización.

Por ejemplo, la siguiente política concede permiso para crear una canalización:

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 también incluye un permiso denominado`osis:Ingest`, que es necesario para enviar solicitudes firmadas a la canalización mediante la versión 4 de Signature. Para obtener más información, consulte the section called “Creación de un rol de ingesta”.

> **Note**
>
> Además, el primer usuario que cree una canalización en una cuenta debe tener permisos para realizar la acción `iam:CreateServiceLinkedRole`. Para más información, consulte recurso de rol de canalización.

Para obtener más información sobre cada permiso, consulte las acciones, los recursos y las claves de condición de la OpenSearch ingestión en la Referencia de autorización de servicios.

Especificar la versión de la canalización

Al crear una canalización con el editor de configuración, debe especificar la versión principal de Data Prepper en la que se ejecutará la canalización. Para especificar la versión, incluya la opción de `version` en la configuración de la canalización:

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

Al elegir Crear, OpenSearch Ingestion determina la última versión secundaria disponible de la versión principal que especifique y aprovisiona la canalización con esa versión. Por ejemplo, si lo especificas `version: "2"` y la última versión compatible de Data Prepper es la 2.1.1, OpenSearch Ingestion aprovisiona tu canalización con la versión 2.1.1. No mostramos públicamente la versión secundaria que está ejecutando la canalización.

Para actualizar la canalización cuando haya disponible una nueva versión principal de Data Prepper, edite la configuración de la canalización y especifique la nueva versión. No puede cambiar una canalización a una versión anterior.

> **Note**
>
> OpenSearch Ingestion no ofrece soporte inmediato a las nuevas versiones de Data Prepper tan pronto como se publican. Habrá cierto intervalo entre el momento en que una nueva versión esté disponible públicamente y el momento en OpenSearch que Ingestion la admita. Además, es posible que OpenSearch Ingestion no admita por completo determinadas versiones principales o secundarias de forma explícita. Para obtener una lista completa, consulte the section called “Versiones de Data Prepper admitidas”.

Cada vez que realices un cambio en la canalización que inicie una blue/green implementación, OpenSearch Ingestion puede actualizarla a la última versión secundaria de la versión principal que esté configurada actualmente para la canalización. Para obtener más información, consulte. the section called “Implementaciones azul/verde para actualizaciones de canalización” OpenSearch La ingestión no puede cambiar la versión principal de tu canalización a menos que actualices explícitamente la `version` opción en la configuración de la canalización.

Especificación de la ruta de ingesta

Para las fuentes basadas en la extracción, como el OTel rastreo y OTel las métricas, OpenSearch Ingestion requiere la `path` opción adicional en la configuración de la fuente. La ruta es una cadena, por ejemplo`/log/ingest`, que representa la ruta del URI para la ingesta. Esta ruta define el URI que usa para enviar datos a la canalización.

Por ejemplo, supongamos que especifica la siguiente ruta para una canalización con un origen HTTP:



Al introducir datos de ingesta en la canalización, debe especificar el siguiente punto de conexión en la configuración de su cliente: `https://pipeline-name-abc123.us-west-2.osis.amazonaws.com/my/test_path`.

La ruta debe empezar con una barra (/) y puede contener los caracteres especiales '-', '_', '.' y '/', así como el marcador de posición `${pipelineName}`. Si usas `${pipelineName}` (por ejemplo`/${pipelineName}/test_path`), OpenSearch Ingestion reemplaza la variable por el nombre de la subcanalización asociada.

Creación de canalizaciones

En esta sección se describe cómo crear canalizaciones OpenSearch de ingestión mediante la consola de OpenSearch servicio y el. AWS CLI

Consola

Para crear una canalización, inicia sesión en la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/osis/home y selecciona Create pipeline.

Seleccione una canalización en blanco o elija un esquema de configuración. Los esquemas incluyen una canalización preconfigurada para una variedad de casos de uso comunes. Para obtener más información, consulte the section called “Uso de esquemas”.

Elija Seleccionar esquema.

Configurar el origen

1. Si parte de una canalización vacía, seleccione un origen del menú desplegable. Las fuentes disponibles pueden incluir otras Servicios de AWS o OpenTelemetry HTTP. Para obtener más información, consulte the section called “Integración de canalizaciones”.
2. En función del origen que elija, configure los ajustes adicionales correspondientes. Por ejemplo, para utilizar Amazon S3 como origen, debe especificar la URL de la cola de Amazon SQS de la canalización que recibe los mensajes. Para obtener una lista de complementos de origen compatibles y enlaces a su documentación, consulte the section called “Complementos y opciones compatibles”.
3. Para algunos orígenes, debe especificar Opciones de red de origen. Seleccione Acceso a VPC o Acceso público. Si elige Acceso público, vaya al siguiente paso. Si elige Acceso a la VCP, configure los siguientes ajustes:

Opción	Description (Descripción)
Administración de puntos de conexión	Elija si quiere crear sus puntos de conexión de nube privada virtual (VPC) usted mismo o dejar que OpenSearch Ingestion los cree por usted. La administración de terminales utiliza de forma predeterminada los puntos finales gestionados por Ingestion. OpenSearch
VPC	Elija el ID de la VPC que desea utilizar. La VPC y la canalización deben estar en la misma Región de AWS.
Subredes	Elija una o más subredes. OpenSearch El servicio colocará un punto final de VPC e interfaces de red elásticas en las subredes.
Grupos de seguridad	Elija uno o más grupos de seguridad de VPC que permitan que la aplicación requerida llegue a la canalización de OpenSearch ingestión en los puertos (80 o 443) y protocolos (HTTP o HTTPs) expuestos por la canalización.
Opciones de conexión de VPC	Si el origen es un punto de conexión autoadministrado, conecte la canalización a una VPC. Elija una de las opciones de CIDR predeterm inadas que se proporcionan o utilice un CIDR personalizado.

Para obtener más información, consulte the section called “Configuración del acceso a la VPC para canalizaciones”.

4. Elija Siguiente.

Configurar el procesador

Agregue uno o más procesadores a su canalización. Los procesadores son componentes de una subcanalización que permiten filtrar, transformar y enriquecer los eventos antes de publicar los registros en el dominio o en el receptor de la colección. Para obtener una lista de los procesadores compatibles y los enlaces a su documentación, consulte the section called “Complementos y opciones compatibles”.

Puede elegir Acciones y agregar lo siguiente:

- Enrutamiento condicional: enruta los eventos a diferentes receptores en función de condiciones específicas. Para obtener más información, consulte Enrutamiento condicional.
- Subcanalización : cada subcanalización es una combinación de un solo origen, cero o más procesadores y un único receptor. Solo una subcanalización puede tener un origen externo. Todas las demás deben tener orígenes que sean otras subcanalizaciones dentro de la configuración general de la canalización. Una única configuración de canalización puede contener de 1 a 10 subcanalizaciones.

Elija Siguiente.

Configurar el receptor

Seleccione el destino en el que la canalización publica los registros. Cada subcanalización debe contener al menos un receptor. Puede agregar un máximo de 10 receptores a una canalización.

Para los OpenSearch sumideros, configura los siguientes campos:

Opción	Description (Descripción)
Nombre de la política de red	Si seleccionó una colección OpenSearch sin servidor, introduzca un nombre de política de red. OpenSearch La ingestión crea la política si no existe o la actualiza con una regla que concede acceso al punto final de la VPC que conecta la canalización y la colección. Para obtener más informaci

Opción	Description (Descripción)
(Solo receptores sin servidor)	ón, consulte the section called “Concesión de acceso a colecciones para canalizaciones”.
Nombre del índice	El nombre del índice al que la canalización envía los datos. OpenSearch La ingestión crea este índice si aún no existe.
Opciones de asignación de índices	Elija la forma en que la canalización almacena e indexa los documentos y sus campos en el OpenSearch colector. Si selecciona el mapeo dinámico, OpenSearch agrega campos automáticamente al indexar un documento. Si selecciona Personalizar asignación, ingrese una plantilla de asignación de índices. Para más información, consulte Plantillas de índices.
Habilite DLQ	Configure una cola de mensajes fallidos (DLQ) de Amazon S3 para la canalización. Para obtener más información, consulte the section called “Colas de mensajes fallidos”.
Ajustes adicionales	Configure las opciones avanzadas para el OpenSearch fregadero. Para obtener más información, consulte Opciones de configuración en la documentación de Data Prepper.

Para agregar un receptor de Amazon S3, seleccione Agregar receptor y Amazon S3. Para obtener más información, consulte the section called “Amazon S3 como destino”.

Elija Siguiente.

Configuración de la canalización

Configure los siguientes ajustes de canalización adicional:

Opción	Description (Descripción)
Nombre de la canalización	Un nombre único para la canalización.
Búfer persistente	Un búfer persistente almacena los datos en un búfer basado en disco en varias zonas de disponibilidad. Para obtener más información, consulte the section called “Almacenamiento en búfer persistente”.

Opción	Description (Descripción)
	Si habilita el búfer persistente, seleccione la clave AWS Key Management Service para cifrar los datos del búfer.
Capacidad de las canalizaciones	La capacidad mínima y máxima de la canalización, en unidades de OpenSearch cómputo de ingestión (OCUs). Para obtener más información, consulte the section called “Escala de canalizaciones”.
Rol de canalizac ión	El rol de IAM que proporciona los permisos necesarios para que la canalizac ión escriba en el receptor y lea desde orígenes basados en extracciones. Puede crear el rol usted mismo o hacer que OpenSearch Ingestion lo cree por usted en función del caso de uso que haya seleccionado. Para obtener más información, consulte the section called “Configuración de roles y usuarios”.
Etiquetas	Agregue una o más etiquetas a su canalización. Para obtener más informaci ón, consulte the section called “Etiquetado de canalizaciones”.
Opciones de publicación de registros	Habilita la publicación de registros de canalización en Amazon CloudWatch Logs. Le recomendamos que habilite la publicación de registros para poder solucionar más fácilmente los problemas de la canalización. Para obtener más información, consulte the section called “Monitoreo de registros de canalización”.

Elija Siguiente, revise la configuración de la canalización y elija Crear canalización.

OpenSearch Ingestion ejecuta un proceso asíncrono para crear la canalización. Una vez que el estado de la canalización sea `Active`, puede empezar a incorporar datos.

AWS CLI

El comando create-pipeline acepta la configuración de la canalización como una cadena o dentro de un archivo .yaml o .json. Si proporciona la configuración en forma de cadena, cada nueva línea debe ir acompañada de secuencias de escape `\n`. Por ejemplo, `"log-pipeline:\n source:\n http:\n processor:\n - grok:\n ...`

El siguiente comando de ejemplo crea una canalización con la siguiente configuración:

- Mínimo 4 ingestas, máximo 10 ingestas OCUs OCUs
- Aprovisionada en una nube privada virtual (VPC)
- Publicación de registros habilitada

```
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 La ingestión ejecuta un proceso asíncrono para crear la canalización. Una vez que el estado de la canalización sea `Active`, puede empezar a incorporar datos. Para comprobar el estado de la canalización, usa el comando. GetPipeline

OpenSearch API de ingestión

Para crear una canalización OpenSearch de ingestión mediante la API de OpenSearch ingestión, llama a la operación. CreatePipeline

Una vez que la canalización se haya creado correctamente, puede configurar su cliente y empezar a ingerir datos en su OpenSearch dominio de servicio. Para obtener más información, consulte the section called “Integración de canalizaciones”.

Seguimiento del estado de creación de la canalización

Puede realizar un seguimiento del estado de una canalización a medida que OpenSearch Ingestion la aprovisiona y la prepara para ingerir datos.

Consola

Una vez creada inicialmente una canalización, ésta pasa por varias etapas a medida que OpenSearch Ingestion la prepara para la ingesta de datos. Para ver las distintas etapas de creación de la canalización, seleccione el nombre de la canalización para ver la página de Configuración de la canalización. En Estado, seleccione Ver detalles.

Una canalización pasa por las siguientes etapas antes de estar disponible para incorporar datos:

- Validación: se valida la configuración de la canalización. Cuando se complete esta etapa, todas las validaciones se han realizado correctamente.
- Crear entorno: preparar y aprovisionar recursos. Cuando se complete esta etapa, se habrá creado el nuevo entorno de canalización.
- Implementar canalización: implementar la canalización. Cuando se complete esta etapa, la canalización se habrá implementado correctamente.
- Comprobar el estado de la canalización: comprobación del estado de la canalización. Cuando se complete esta etapa, todas las comprobaciones de estado se habrán aprobado.
- Habilitar tráfico: permitir que la canalización incorpore datos. Cuando se complete esta etapa, puede empezar a incorporar datos a la canalización.

CLI

Usa el get-pipeline-change-progresscomando para comprobar el estado de una canalización. La siguiente AWS CLI solicitud comprueba el estado de una canalización denominada`my-pipeline`:

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

Respuesta:

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

OpenSearch API de ingestión

Para realizar un seguimiento del estado de la creación de la canalización mediante la API OpenSearch de ingestión, llama a la GetPipelineChangeProgressoperación.

Uso de esquemas

En lugar de crear una definición de canalización desde cero, puede usar esquemas de configuración, que son plantillas preconfiguradas para escenarios de ingesta comunes, como los registros de Trace Analytics o Apache. Los esquemas de configuración lo ayudan a aprovisionar canalizaciones fácilmente sin tener que crear una configuración desde cero.

Consola

Cómo usar un esquema de canalización

1. Inicie sesión en la consola de OpenSearch Ingestion en osis/home. https://console.aws.amazon.com/aos/ Estarás en la página de Pipelines.
2. Seleccione Create pipeline.
3. Seleccione un esquema en la lista de casos de uso y, a continuación, elija Seleccionar esquema. La configuración de la canalización se completa con una subcanalización para el caso de uso que haya seleccionado.

   El esquema de la canalización no es válido tal como está. Debe especificar ajustes adicionales en función del origen seleccionado.

CLI

Para obtener una lista de todos los planos disponibles mediante el AWS CLI, envía una list-pipeline-blueprintssolicitud.

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

La solicitud devuelve una lista de todos los esquemas disponibles.

Para obtener información más detallada sobre un plano específico, utilice el get-pipeline-blueprintcomando:

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

Esta solicitud devuelve el contenido del esquema de la canalización de registro de 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 de ingestión

Para obtener información sobre los planos de canalización mediante la API OpenSearch de ingestión, usa las ListPipelineBlueprintsoperaciones y. GetPipelineBlueprint

Visualización de las canalizaciones OpenSearch de Amazon Ingestion

Puede ver los detalles de una canalización de Amazon OpenSearch Ingestion mediante la Consola de administración de AWS AWS CLI, la o la API de OpenSearch ingestión.

Consola

Para ver una canalización

1. Inicia sesión en la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/osis/home. Estarás en la página de Pipelines.
2. (Opcional) Para ver las canalizaciones con un estado determinado, seleccione Cualquier estado y seleccione un estado para filtrar.

   Una canalización puede tener los siguientes estados:

   - `Active`: la canalización está activa y lista para incorporar datos.
   - `Creating`: la canalización se está creando.
   - `Updating`: la canalización se está actualizando.
   - `Deleting`: la canalización se está eliminando.
   - `Create failed`: la canalización no se pudo crear.
   - `Update failed`: la canalización no se pudo actualizar.
   - `Stop failed`: la canalización no se pudo detener.
   - `Start failed`: la canalización no se pudo iniciar.
   - `Stopping`: la canalización se está deteniendo.
   - `Stopped`: la canalización está detenida y se puede reiniciar en cualquier momento.
   - `Starting`: la canalización se está iniciando.

No se te facturará la ingesta OCUs cuando una canalización esté en los estados`Create failed`, `CreatingDeleting`, y. `Stopped`

CLI

Para ver las canalizaciones mediante el AWS CLI, envía una solicitud de lista de canalizaciones:

```
aws osis list-pipelines
```

La solicitud devuelve una lista de todas las canalizaciones existentes:

```
{
    "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."
            }
        }
    ]
}
```

Para ver información sobre una sola canalización, use el comando get-pipeline:

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

La solicitud devuelve la información de configuración de la canalización especificada:

```
{
    "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 de ingestión

Para ver las canalizaciones OpenSearch de ingestión mediante la API de OpenSearch ingestión, llama a las operaciones and. ListPipelinesGetPipeline

Actualización de las canalizaciones OpenSearch de Amazon Ingestion

Puede actualizar las canalizaciones OpenSearch de Amazon Ingestion mediante la Consola de administración de AWS, la o la AWS CLI API de OpenSearch ingestión. OpenSearch Ingestion inicia una blue/green implementación al actualizar la configuración de una canalización. Para obtener más información, consulte the section called “Implementaciones azul/verde para actualizaciones de canalización”.

Temas

- Consideraciones
- Permisos necesarios
- Actualización de las canalizaciones
- Implementaciones azul/verde para actualizaciones de canalización

Consideraciones

Tenga en cuenta lo siguiente cuando actualice una canalización:

- No se puede actualizar el nombre ni la configuración de red de una canalización.
- Si su canalización escribe en un receptor de dominio de VPC, no puede volver atrás y cambiar el receptor a un dominio de VPC diferente una vez creada la canalización. Debe eliminar y volver a crear manualmente la canalización con el nuevo receptor. Aún puede cambiar el receptor de un dominio de VPC a un dominio público, de un dominio público a un dominio de VPC o de un dominio público a otro dominio público.
- Puede cambiar el colector de canalización en cualquier momento entre un dominio de OpenSearch servicio público y una colección OpenSearch sin servidor.
- Al actualizar la configuración de origen, procesador o receptor de una canalización, OpenSearch Ingestion inicia una implementación. blue/green Para obtener más información, consulte the section called “Implementaciones azul/verde para actualizaciones de canalización”.
- Al actualizar la configuración de origen, procesador o receptor de una canalización, OpenSearch Ingestion la actualiza automáticamente a la última versión secundaria compatible de la versión principal de Data Prepper que esté ejecutando la canalización. Este proceso mantiene su canalización actualizada con las últimas correcciones de errores y mejoras de rendimiento.
- Puede seguir actualizando su canalización cuando esté detenida.

Permisos necesarios

OpenSearch Ingestion utiliza los siguientes permisos de IAM para actualizar las canalizaciones:

- `osis:UpdatePipeline`: Actualizar una canalización.
- `osis:ValidatePipeline`: Comprobar si la configuración de la canalización es válida.
- `iam:PassRole`— Transfiera la función de canalización a OpenSearch Ingestion para que pueda escribir datos en el dominio. Este permiso solo es necesario si actualiza la configuración de la canalización, no si modifica otros ajustes, como la publicación de registros o los límites de capacidad.

Por ejemplo, la siguiente política concede permiso para actualizar una canalización:

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"
            ]
        }
    ]
}
```

Actualización de las canalizaciones

Puede actualizar las canalizaciones OpenSearch de Amazon Ingestion mediante la Consola de administración de AWS, la o la AWS CLI API de OpenSearch ingestión.

Consola

Cómo actualizar una canalización

1. Inicia sesión en la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/osis/home. Estarás en la página de Pipelines.
2. Seleccione una canalización para abrir su configuración. A continuación, seleccione una de las opciones de edición.
3. Cuando termine de realizar los cambios, seleccione Guardar.

CLI

Para actualizar una canalización mediante el AWS CLI, envía una solicitud de actualización de la canalización. En el siguiente ejemplo de solicitud, se carga un nuevo archivo de configuración y se actualizan los valores de capacidad mínima y máxima:

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

OpenSearch API de ingestión

Para actualizar una canalización OpenSearch de ingestión mediante la API de OpenSearch ingestión, llama a la operación. UpdatePipeline

Implementaciones azul/verde para actualizaciones de canalización

OpenSearch La ingestión inicia un proceso de despliegue azul o verde al actualizar la configuración de una canalización.

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/greenlas implementaciones en sí mismas no tienen ningún impacto en el rendimiento, pero el rendimiento puede cambiar si la configuración de la canalización cambia de una manera que altere el rendimiento.

OpenSearch La ingestión bloquea el blue/green autoescalado durante las implementaciones. Se le seguirá cobrando solo por el tráfico a la canalización anterior hasta que se redirija a la nueva canalización. Una vez que se haya redirigido el tráfico, solo se cobrará por la nueva canalización. Nunca se le cobrará por dos canalizaciones simultáneamente.

Al actualizar la configuración de origen, procesador o receptor de una canalización, OpenSearch Ingestion puede actualizar automáticamente la canalización a la última versión secundaria compatible de la versión principal en la que se esté ejecutando la canalización. Por ejemplo, es posible que tengas una configuración `version: "2"` de canalización e OpenSearch Ingestion aprovisionó inicialmente la canalización con la versión 2.1.0. Cuando se añade la compatibilidad con la versión 2.1.1 y realizas un cambio en la configuración de la canalización, OpenSearch Ingestion la actualiza a la versión 2.1.1.

Este proceso mantiene tu canalización actualizada con las últimas correcciones de errores y mejoras de rendimiento. OpenSearch Ingestion no puede actualizar la versión principal de tu canalización a menos que cambies manualmente la `version` opción en la configuración de la canalización.

Gestión de los costes del pipeline OpenSearch de Amazon Ingestion

Puede iniciar y detener las canalizaciones de ingestión en Amazon OpenSearch Ingestion para controlar el flujo de datos en función de sus necesidades. Al detener una canalización, se detiene el procesamiento de datos y, al mismo tiempo, se conservan las configuraciones, por lo que puede reiniciarla sin tener que volver a configurarla. Esto puede ayudar a optimizar los costos, administrar el uso de los recursos o solucionar problemas. Al detener una canalización, OpenSearch Ingestion no procesa los datos entrantes, pero los datos ingeridos anteriormente permanecen disponibles en él. OpenSearch

El inicio y la detención simplifican los procesos de configuración y eliminación de canalizaciones que usa para desarrollo, pruebas o actividades similares que no requieren disponibilidad continua. Mientras su canalización esté detenida, no se le cobrará ninguna hora de Ingestion OCU. Puede seguir actualizando las canalizaciones detenidas y estas reciben actualizaciones automáticas de las versiones secundarias y revisiones de seguridad.

Al detener e iniciar una canalización, se reprocesarán todos los datos desde el principio para las canalizaciones basadas en extracciones (DDB, S3, DocDB, etc.). Al detener una canalización, se eliminan todos los puntos finales de VPC gestionados por el servicio que haya creado la canalización. En el caso de las canalizaciones con puntos de enlace de VPC autogestionados, debe volver a crear el punto de enlace de VPC en su cuenta al reiniciar la canalización. Para obtener más información, consulte the section called “Puntos de conexión de VPC autogestionados”.

> **Note**
>
> Si su canalización tiene un exceso de capacidad pero necesita permanecer operativa, considera ajustar sus límites de capacidad máxima en lugar de detenerla y reiniciarla. Esto puede ayudar a administrar los costos y, al mismo tiempo, garantizar que la canalización continúe procesando los datos de manera eficiente. Para obtener más información, consulte the section called “Escala de canalizaciones”.

En los siguientes temas se explica cómo iniciar y detener las canalizaciones mediante la API, y la API de ingestión. Consola de administración de AWS AWS CLI OpenSearch

Temas

- Detener una canalización de Amazon OpenSearch Ingestion
- Inicio de una canalización de Amazon OpenSearch Ingestion

Detener una canalización de Amazon OpenSearch Ingestion

Para utilizar una canalización de OpenSearch ingestión o realizar tareas de administración, siempre hay que empezar con una canalización activa, después se detiene la canalización y, a continuación, se vuelve a iniciar la canalización. Mientras su canalización esté detenida, no se le cobrará ninguna hora de Ingestion OCU.

Consola

Cómo detener una canalización

1. Inicia sesión en la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/osis/home. Estarás en la página de Pipelines.
2. Elige un oleoducto. Puede realizar la operación de detención desde esta página o navegar a la página de detalles de la canalización de datos que desea detener.
3. En Acciones, seleccione Detener canalización.

   Si no se puede detener ni iniciar una canalización, la acción Detener canalización no estará disponible.

AWS CLI

Para detener una canalización mediante AWS CLI el comando stop-pipeline con los siguientes parámetros:

- `--pipeline-name`: el nombre de la canalización.

Example

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

OpenSearch API de ingestión

Para detener una canalización mediante la API OpenSearch de ingestión, llama a la StopPipelineoperación con el siguiente parámetro:

- `PipelineName`: el nombre de la canalización.

Inicio de una canalización de Amazon OpenSearch Ingestion

Siempre se inicia una canalización de OpenSearch ingestión empezando por una canalización que ya está detenida. La canalización mantiene sus ajustes de configuración como límites de capacidad, configuración de red y opciones de publicación de registros.

El reinicio de la canalización suele tardar varios minutos.

Consola

Cómo iniciar una canalización

1. Inicia sesión en la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/osis/home. Estarás en la página de Pipelines.
2. Elige un oleoducto. Puede realizar la operación de inicio desde esta página o navegar a la página de detalles de la canalización que desea iniciar.
3. En Acciones, seleccione Iniciar canalización.

AWS CLI

Para iniciar una canalización mediante el AWS CLI comando start-pipeline con los siguientes parámetros:

- `--pipeline-name`: el nombre de la canalización.

Example

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

OpenSearch API de ingestión

Para iniciar una canalización OpenSearch de ingestión mediante la API OpenSearch de ingestión, llama a la StartPipelineoperación con el siguiente parámetro:

- `PipelineName`: el nombre de la canalización.

Eliminar las canalizaciones OpenSearch de Amazon Ingestion

Puede eliminar una canalización de Amazon OpenSearch Ingestion mediante la Consola de administración de AWS AWS CLI, la o la API de OpenSearch ingestión. No puede eliminar una canalización si tiene un estado de `Creating` o `Updating`.

Consola

Cómo eliminar una canalización

1. Inicia sesión en la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/osis/home. Estarás en la página de Pipelines.
2. Seleccione la canalización que desee eliminar y seleccione Acciones, Eliminar.
3. Confirme la eliminación y elija Eliminar.

CLI

Para eliminar una canalización mediante el AWS CLI, envía una solicitud de borrado de canalización:

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

OpenSearch API de ingestión

Para eliminar una canalización OpenSearch de ingestión mediante la API OpenSearch de ingestión, llama a la DeletePipelineoperación con el siguiente parámetro:

- `PipelineName`: el nombre de la canalización.

Plugins y opciones compatibles para las canalizaciones de Amazon OpenSearch Ingestion

Amazon OpenSearch Ingestion admite un subconjunto de fuentes, procesadores y receptores en Data Prepper de código OpenSearch abierto. Además, OpenSearch Ingestion impone algunas restricciones a las opciones disponibles para cada complemento compatible. En las siguientes secciones se describen los complementos y las opciones asociadas que admite OpenSearch Ingestion.

> **Note**
>
> OpenSearch Ingestion no admite ningún complemento de búfer porque configura automáticamente un búfer predeterminado. Si incluye un búfer en la configuración de la canalización, recibirá un error de validación.

Temas

- Complementos compatibles
- Procesadores sin estado frente a procesadores con estado
- Requisitos y restricciones de configuración

Complementos compatibles

OpenSearch Ingestion es compatible con los siguientes complementos de Data Prepper:

Fuentes:

- DocumentDB
- DynamoDB

- HTTP
- Kafka
- Kinesis
- OpenSearch
- OTel logs
- OTel métricas
- OTel rastreo
- S3

Procesadores:

- Agregar entradas
- Aggregate

- Detector de anomalías
- AWS Lambda
- Convertir el tipo de entrada
- Copiar valores
- CSV
- Fecha
- Retraso
- Descomprimir
- Eliminar entradas
- Diseccionar
- Soltar eventos
- Aplanar
- IP geográfica
- Grok
- Valor de clave
- Lista a mapa
- Cadena en minúsculas
- Mapa a lista
- Evento de mutación (serie de procesadores)
- Cadena de mutación (serie de procesadores)
- Obfuscate
- OTel métricas
- OTel grupo de rastreo
- OTel rastrear
- Analizar Ion
- Analizar JSON
- Analizar XML
- Cambiar el nombre de las claves
- Seleccionar entradas
- Mapa de servicios

- Evento dividido
- Cadena dividida
- Convertidor de cadenas
- Cadena sustitutiva
- Reenviador por pares de seguimiento
- Translate
- Recortar cadena
- Truncar
- Cadena en mayúscula
- Agente usuario
- Escribir JSON

Receptores:

- OpenSearch(es compatible con OpenSearch Service, OpenSearch Serverless y Elasticsearch 6.8 o versiones posteriores)
- S3

Códecs de receptor:

- Avro
- NDJSON
- JSON
- Parquet

Procesadores sin estado frente a procesadores con estado

Los procesadores sin estado realizan operaciones como las transformaciones y el filtrado, mientras que los procesadores con estado realizan operaciones como las agregaciones, que recuerdan el resultado de la ejecución anterior. OpenSearch Ingestion es compatible con los procesadores con estado Aggregate y Service-map. Todos los demás procesadores compatibles no tienen estado.

En el caso de las canalizaciones que contienen únicamente procesadores sin estado, el límite máximo de capacidad es de 96 ingestion. OCUs Si una canalización contiene procesadores con

estado, el límite máximo de capacidad es de 48 tomas. OCUs Sin embargo, si una canalización tiene habilitado el almacenamiento en búfer persistente, puede tener un máximo de 384 ingestas solo OCUs con procesadores sin estado, o 192 ingestas si contiene procesadores con estado. OCUs Para obtener más información, consulte the section called “Escala de canalizaciones”.

End-to-end El reconocimiento solo se admite en procesadores sin estado. Para obtener más información, consulte the section called “End-to-end reconocimiento”.

Requisitos y restricciones de configuración

A menos que se especifique lo contrario a continuación, todas las opciones descritas en la referencia de configuración de Data Prepper para los complementos compatibles enumerados anteriormente están permitidas en OpenSearch los procesos de ingestión. En las siguientes secciones se explican las restricciones que OpenSearch Ingestion impone a determinadas opciones de los complementos.

> **Note**
>
> OpenSearch Ingestion no admite ningún complemento de búfer porque configura automáticamente un búfer predeterminado. Si incluye un búfer en la configuración de la canalización, recibirá un error de validación.

OpenSearch Ingestion configura y administra internamente muchas opciones, como y. `authentication` `acm_certificate_arn` Otras opciones, como `thread_count` `yrequest_timeout`, tienen un impacto en el rendimiento si se modifican manualmente. Por lo tanto, estos valores se establecen internamente para garantizar un rendimiento óptimo de sus canalizaciones.

Por último, algunas opciones no se pueden pasar a OpenSearch Ingestion, como `ism_policy_file` `ysink_template`, porque son archivos locales cuando se ejecutan en Data Prepper de código abierto. No se admiten estos valores.

Temas

- Opciones generales de canalización
- Procesador Grok
- Origen de HTTP
- OpenSearch sumidero
- OTel fuente de métricas, fuente de OTel rastreo y fuente de OTel registros

- OTel procesador de grupos de rastreo
- OTel procesador de trazas
- Procesador de mapas de servicio
- Fuente de S3

Opciones generales de canalización

OpenSearch Ingestion establece las siguientes opciones generales de canalización y no se admiten en las configuraciones de canalización:

- `workers`
- `delay`

Procesador Grok

No se admiten las siguientes opciones de procesador Grok:

- `patterns_directories`
- `patterns_files_glob`

Origen de HTTP

El complemento fuente HTTP tiene los siguientes requisitos y restricciones:

- La opción `path` es obligatoria. La ruta es una cadena, por ejemplo `/log/ingest`, que representa la ruta del URI para la ingesta de registros. Esta ruta define el URI que usa para enviar datos a la canalización. Por ejemplo, `https://log-pipeline.us-west-2.osis.amazonaws.com/log/ingest`. La ruta debe empezar con una barra (/) y puede contener los caracteres especiales '-', '_', '.' y '/', así como el marcador de posición `${pipelineName}`.
- OpenSearch Ingestion establece las siguientes opciones de fuente HTTP y no se admiten en las configuraciones de canalización:
  - `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 sumidero

El complemento OpenSearchsink tiene los siguientes requisitos y limitaciones.

- La opción `aws` es obligatoria y debe contener las siguientes opciones:
  - `sts_role_arn`
  - `region`
  - `hosts`
  - `serverless`(si el sumidero es una colección OpenSearch sin servidor)
- La opción `sts_role_arn` debe apuntar al mismo rol para cada receptor de un archivo de definición YAML.
- La `hosts` opción debe especificar un punto final de dominio OpenSearch de servicio o un punto final de colección OpenSearch sin servidor. No se puede especificar un punto de conexión personalizado para un dominio; este debe ser el punto de conexión estándar.
- Si la opción de `hosts` es un punto de conexión de la colección sin servidor, debe configurar la opción `serverless` en `true`. Además, si su archivo de definición YAML contiene la opción `index_type`, debe estar configurada en `management_disabled`, de lo contrario, la validación fallará.
- Las siguientes opciones no son compatibles:

- `username`
- `password`
- `cert`
- `proxy`
- `dlq_file`: si quiere descargar los eventos fallidos a una cola de mensajes fallidos (DLQ), debe usar la opción `dlq` y especificar un bucket de S3.
- `ism_policy_file`
- `socket_timeout`
- `template_file`
- `insecure`

OTel fuente de métricas, fuente de OTel rastreo y fuente de OTel registros

Los complementos fuente de OTel métricas, fuente de OTel rastreo y fuente de OTel registros tienen los siguientes requisitos y limitaciones:

- La opción `path` es obligatoria. La ruta es una cadena, por ejemplo `/log/ingest`, que representa la ruta del URI para la ingesta de registros. Esta ruta define el URI que usa para enviar datos a la canalización. Por ejemplo, `https://log-pipeline.us-west-2.osis.amazonaws.com/log/ingest`. La ruta debe empezar con una barra (/) y puede contener los caracteres especiales '-', '_', '.' y '/', así como el marcador de posición `${pipelineName}`.
- OpenSearch Ingestion establece las siguientes opciones y no se admiten en las configuraciones de canalización:
  - `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 procesador de grupos de rastreo

El procesador de grupos de OTel rastreo tiene los siguientes requisitos y limitaciones:

- La opción `aws` es obligatoria y debe contener las siguientes opciones:
  - `sts_role_arn`
  - `region`
  - `hosts`
- La `sts_role_arn` opción especifica la misma función que la función de canalización que se especifica en la configuración OpenSearch del receptor.
- No se admiten las opciones `username`, `password`, `cert`, y `insecure`.
- La opción `aws_sigv4` es obligatoria y se debe establecer en verdadero.
- No se admite la `serverless` opción incluida en el complemento de OpenSearch sumidero. El procesador Otel Trace Group no funciona actualmente con las colecciones OpenSearch Serverless.
- El número de procesadores `otel_trace_group` en el cuerpo de configuración de la canalización no puede ser superior a 8.

OTel procesador de trazas

El procesador de OTel trazas tiene los siguientes requisitos y limitaciones:

- El valor de la opción `trace_flush_interval` no puede superar los 300 segundos.

Procesador de mapas de servicio

El procesador de Service-map tiene los siguientes requisitos y limitaciones:

- El valor de la opción `window_duration` no puede superar los 300 segundos.

Fuente de S3

El complemento fuente de S3 tiene los siguientes requisitos y limitaciones:

- La opción `aws` es obligatoria y debe contener las opciones `region` y `sts_role_arn`.
- El valor de la opción `records_to_accumulate` no puede ser superior a 200.
- El valor de la opción `maximum_messages` no puede ser superior a 10.
- Si se especifica, la opción `disable_bucket_ownership_validation` se debe establecer en falso.
- Si se especifica, la opción `input_serialization` se debe establecer en `parquet`.

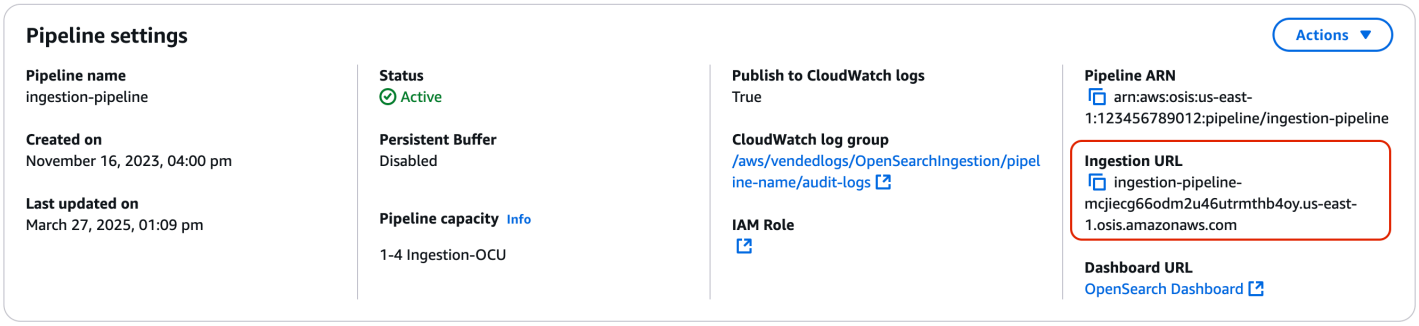
Integración de las canalizaciones OpenSearch de Amazon Ingestion con otros servicios y aplicaciones

Para incorporar datos correctamente a una canalización de Amazon OpenSearch Ingestion, debe configurar la aplicación cliente (la fuente) para enviar los datos al punto final de la canalización. Su fuente pueden ser clientes como Fluent Bit logs, OpenTelemetry Collector o un simple bucket de S3. La configuración exacta es diferente para cada cliente.

Las diferencias importantes durante la configuración de la fuente (en comparación con el envío de datos directamente a un dominio de OpenSearch servicio o a una recopilación OpenSearch sin servidor) son el nombre del AWS servicio (`osis`) y el punto final del host, que debe ser el punto final de la canalización.

Creación del punto de conexión de ingesta

Para ingerir datos a una canalización, envíelos al punto de conexión de ingesta. Para localizar la URL de ingesta, vaya a la página Configuración de canalización y copie la URL de ingesta.



Para crear el punto final de ingesta completo para las fuentes basadas en la extracción, como el OTel seguimiento y OTel las métricas, añade la ruta de ingesta desde la configuración de tu canalización a la URL de ingesta.

Por ejemplo, supongamos que su configuración de canalización tiene la siguiente ruta de ingesta:



El punto de conexión de ingesta completo, que se especifica en la configuración de su cliente, tendrá el siguiente formato: `https://ingestion-pipeline-abcdefg.us-east-1.osis.amazonaws.com/my/test_path`.

Creación de un rol de ingesta

Todas las solicitudes de OpenSearch ingestión deben estar firmadas con la versión 4 de Signature. Como mínimo, el rol que firma la solicitud debe tener permiso para realizar la `osis:Ingest` acción, lo que le permite enviar datos a una canalización de OpenSearch Ingestion.

Por ejemplo, la siguiente política AWS Identity and Access Management (IAM) permite que la función correspondiente envíe datos a una única canalización:

JSON

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

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

> **Note**
>
> Para usar el rol en todas las canalizaciones, sustituya el ARN del elemento `Resource` por un comodín (*).

Provisión de acceso de ingesta entre cuentas

> **Note**
>
> Solo puede proporcionar acceso de ingesta entre cuentas a las canalizaciones públicas, no a las canalizaciones de VPC.

Es posible que tengas que introducir datos en una canalización desde otra Cuenta de AWS, como una cuenta que aloje tu aplicación de origen. Si la entidad principal que escribe en una canalización está en una cuenta diferente a la de la propia canalización, debe configurar la entidad principal para que confíe en otro rol de IAM para incorporar los datos a la canalización.

Cómo configurar los permisos de ingesta entre cuentas

1. Crea el rol de ingesta con `osis:Ingest` permiso (descrito en la sección anterior) dentro de la Cuenta de AWS misma canalización. Para ver instrucciones, consulte Creación de roles de IAM.
2. Adjunte una política de confianza al rol de ingesta que permita que la entidad principal de otra cuenta lo asuma:

   JSON

```
{
```

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

3. En la otra cuenta, configure la aplicación cliente (por ejemplo, Fluent Bit) para que asuma el rol de ingesta. Para que esto funcione, la cuenta de la aplicación debe conceder permisos al usuario o rol de la aplicación para que asuma el rol de ingesta.

   El siguiente ejemplo de política basada en identidad permite que la entidad principal adjunta asuma el `ingestion-role` de la cuenta de canalización:

   JSON

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

A continuación, la aplicación cliente puede utilizar la AssumeRoleoperación para asumir `ingestion-role` e ingerir datos en la canalización asociada.

Uso de una canalización OpenSearch de ingestión con Atlassian Services

Puedes usar los complementos fuente de Atlassian Jira y Confluence para incorporar datos de los servicios de Atlassian a tu canal de ingestión. OpenSearch Estas integraciones le permiten crear una base de conocimientos unificada con capacidad de búsqueda al sincronizar proyectos completos de Jira y los espacios de Confluence y, al mismo tiempo, mantener la relevancia en tiempo real mediante la supervisión continua y la sincronización automática de las actualizaciones.

Integrating with Jira

Transforma tu experiencia con Jira con potentes funciones de búsqueda contextual integrando tu contenido de Jira en. OpenSearch El complemento de origen Atlassian Jira de Data Prepper le permite crear una base de conocimientos unificada y con capacidad de búsqueda mediante la sincronización de proyectos completos de Jira, al tiempo que mantiene la relevancia en tiempo real gracias a la supervisión continua y la sincronización automática de las actualizaciones. Esta integración permite la sincronización de datos con opciones de filtrado flexibles para proyectos, tipos de problemas y estados específicos, lo que garantiza que solo se importe la información que necesita.

Para garantizar una conectividad segura y fiable, el complemento admite varios métodos de autenticación, incluida la autenticación y OAuth2 la autenticación mediante claves de API básicas, con la seguridad adicional de gestionar las credenciales mediante un secreto almacenado en él. AWS Secrets Manager También cuenta con la renovación automática de los tokens para un acceso ininterrumpido, lo que garantiza un funcionamiento continuo. Esta integración, que se basa en la versión 2 de la API de Atlassian, permite a los equipos obtener información valiosa a partir de sus datos de Jira mediante las funciones de búsqueda avanzada OpenSearch de las que dispone.

Integrating with Confluence

Mejora las capacidades de colaboración y gestión del conocimiento de tu equipo integrando el contenido de Confluence de Atlassian mediante el complemento fuente Confluence de Data OpenSearch Prepper. Esta integración le permite crear un repositorio de conocimiento colectivo centralizado y con capacidad de búsqueda, lo que mejora la detección de información y la productividad del equipo. Al sincronizar el contenido de Confluence y supervisar continuamente las actualizaciones, el complemento garantiza que tu índice siga siendo completo. OpenSearch up-to-date

La integración ofrece opciones de filtrado flexibles, lo que le permite importar contenido de forma selectiva desde espacios o tipos de páginas específicos, adaptando el contenido sincronizado a las necesidades de su organización. El complemento admite tanto la clave de API básica como los métodos de OAuth2 autenticación, con la opción de gestionar las credenciales de forma segura mediante ellos. AWS Secrets Manager La característica de renovación automática de los tokens del complemento garantiza un acceso ininterrumpido y un funcionamiento sin problemas. Basada en la API de Confluence de Atlassian, esta integración permite a los equipos aprovechar las funciones de búsqueda avanzada que OpenSearch ofrecen en todo el contenido de Confluence, lo que mejora la accesibilidad y la utilización de la información en la organización.

Temas

- Requisitos previos
- Configurar un rol de canalización
- Configuración de una canalización de Jira Connector
- Configuración de una canalización del conector de Confluence
- Coherencia de datos
- Limitaciones
- Métricas incluidas para los conectores de CloudWatch Atlassian
- Conexión de una canalización de Amazon OpenSearch Ingestion a Atlassian, Jira o Confluence mediante la versión 2.0 OAuth

Requisitos previos

Antes de crear tu canal de OpenSearch ingestión, sigue los siguientes pasos:

1. Prepara las credenciales de tu sitio de Jira seleccionando una de las siguientes opciones. OpenSearch La ingestión solo requiere la `ReadOnly` autorización del contenido.
   a. Opción 1: clave de API: inicie sesión en su cuenta de Atlassian y utilice la información del siguiente tema para generar su clave de API:
      - Gestione los tokens de API para su cuenta de Atlassian
   b. Opción 2: OAuth2 — Inicia sesión en tu cuenta de Atlassian y utiliza la información en ella. the section called “Conexión de una canalización de Amazon OpenSearch Ingestion a Atlassian, Jira o Confluence mediante la versión 2.0 OAuth”
2. Cree un secreto de AWS Secrets Manager para almacenar los detalles de las credenciales que creó en el paso anterior: Realice las siguientes elecciones a medida que sigue el procedimiento:
   - En Secret type (Tipo de secreto), elija Other type of secret (Otro tipo de secreto).
   - Para los pares clave/valor, cree los siguientes pares, según el tipo de autorización seleccionado:

   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
}
```

Una vez creado el secreto, copie el nombre de recurso de Amazon (ARN) del secreto. Lo incluirá en la política de permisos de los roles en proceso.

Configurar un rol de canalización

El rol transferido en la canalización debe incluir la siguiente política de lectura y escritura en el secreto creado en la sección de requisitos previos.

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"
        }
    ]
}
```

El rol también debe incluir una política de acceso y escritura en el servidor que elijas. Por ejemplo, si eliges OpenSearch como destino, la política es similar a la siguiente:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "OpenSearchWritePolicy",
            "Effect": "Allow",
            "Action": "aoss:*",
            "Resource": "arn:aws:aoss:us-
east-1:111122223333:collection/collection-id"
        }
    ]
}
```

Configuración de una canalización de Jira Connector

Puede utilizar un esquema de Atlassian Jira preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Sustituya *placeholder values* por su propia información.

```
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 “Compresión de solicitudes 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"
```

Atributos clave del origen de Jira:

1. hosts: la URL en las instalaciones o en la nube de Jira. En general, se ve similar a `https://your-domain-name.atlassian.net/`.
2. acknowledgments: para garantizar la entrega de los datos hasta el receptor.
3. authentication: describe cómo desea que la canalización acceda a su instancia de Jira. Elija `Basic` o especifique `OAuth2` los atributos clave correspondientes que hagan referencia a las claves de su AWS secreto.
4. filter: esta sección lo ayuda a seleccionar qué parte de sus datos de Jira desea extraer y sincronizar.
   a. project: enumere las claves del proyecto que desea sincronizar en la sección `include`. De lo contrario, enumere los proyectos que desea excluir en la sección `exclude`. Proporcione solo una de las opciones de inclusión o exclusión en un momento dado.
   b. issue_type: tipos de problemas específicos que desea sincronizar. Siga un patrón similar a `include` o `exclude` que se adapte a sus necesidades. Tenga en cuenta que los archivos adjuntos aparecerán como enlaces de anclaje al archivo adjunto original, pero el contenido del archivo adjunto no se extraerá.
   c. status: filtro de estado específico que desea aplicar a la consulta de extracción de datos. Si especifica `include`, solo se sincronizarán los tickets con esos estados. Si especifica `exclude`, se sincronizarán todos los tickets, excepto los que tengan los estados de exclusión enumerados.

Configuración de una canalización del conector de Confluence

Puede utilizar un esquema de Atlassian Confluence preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

```
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 “Compresión de solicitudes
 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"
```

Atributos clave del origen de Confluence:

1. hosts: la URL en las instalaciones o en la nube de Confluence. En general, se ve similar a `https://your-domain-name.atlassian.net/`.
2. acknowledgments: para garantizar la entrega de los datos hasta el receptor.
3. authentication: describe cómo desea que la canalización acceda a su instancia de Confluence. Elija `Basic` o `0Auth2` y especifique los atributos clave correspondientes que hacen referencia a las claves de su AWS secreto.
4. filter: esta sección lo ayuda a seleccionar qué parte de sus datos de Confluence desea extraer y sincronizar.
   a. space: enumere las claves de espacio que desea sincronizar en la sección `include`. De lo contrario, enumere los espacios que desea excluir en la sección `exclude`. Proporcione solo una de las opciones de inclusión o exclusión en un momento dado.
   b. page_type: tipos de página específicos (como página, entrada de blog o archivos adjuntos) que desea sincronizar. Siga un patrón similar a `include` o `exclude` que se adapte a sus necesidades. Tenga en cuenta que los archivos adjuntos aparecerán como enlaces de anclaje al archivo adjunto original, pero el contenido del archivo adjunto no se extraerá.

Coherencia de datos

Según los filtros especificados en el archivo YAML de la canalización, los proyectos (o espacios) seleccionados se extraerán una vez y se sincronizarán completamente con el receptor de destino. Luego, la supervisión continua de los cambios capturará los cambios a medida que se produzcan y actualizará los datos del receptor. Una excepción es que la supervisión de cambios sincroniza solo las acciones `create` y `update`, no las acciones `delete`.

Limitaciones

- Las acciones de eliminación de usuarios no se sincronizarán. Una vez registrados en el receptor, los datos permanecerán allí. Las actualizaciones sobrescribirán el contenido existente con nuevos cambios si se especifica la asignación de ID en la configuración del receptor.
- Las instancias locales que utilizan versiones anteriores del software de Atlassian que no son compatibles con lo siguiente no APIs son compatibles con esta fuente:
  - API de búsqueda de Jira, versión 3

    - `rest/api/3/search`
    - `rest/api/3/issue`
- Confluence
    - `wiki/rest/api/content/search`
    - `wiki/rest/api/content`
    - `wiki/rest/api/settings/systemInfo`

Métricas incluidas para los conectores de CloudWatch Atlassian

Tipo: métricas del conector de Jira

origen	Métrica	Tipo de métrica
acknowledgementSetSuccesses.contar	Contador	Si los acuses de recibo están habilitados, esta métrica proporciona el número de tickets que se sincronizaron correctamente.
acknowledgementSetFailures.count	Contador	Si las confirmaciones están habilitadas, esta métrica proporciona el número de tickets que no se pudieron sincronizar.
crawlingTime.avg	Timer	El tiempo que se tardó en revisar todos los cambios nuevos.
ticketFetchLatency.avg	Timer	El promedio de latencia de la API de búsqueda de tickets.
ticketFetchLatency.max	Timer	La latencia máxima de la API de búsqueda de tickets.
ticketsRequested.count	Contador	Número de solicitudes de recogida de entradas realizadas.

origen	Métrica	Tipo de métrica
ticketRequestedFailed.count	Contador	Número de solicitudes de recuperación de entradas fallidas.
ticketRequestedSuccess.count	Contador	Número de solicitudes de recuperación de entradas realizadas correctamente.
searchCallLatency.avg	Timer	Latencia media de llamadas a la API de búsqueda.
searchCallLatency.max	Timer	Latencia máxima de llamadas a la API de búsqueda.
searchResultsFound.count	Contador	Número de elementos encontrados en una llamada de búsqueda determinada.
searchRequestFailed.count	Contador	Los errores en las llamadas a la API de búsqueda cuentan.
authFailures.count	Contador	Número de errores de autenticación.

Tipo: métricas del conector de Confluence

origen	Métrica	Tipo de métrica
acknowledgementSet	Contador	Si las confirmaciones están habilitadas, esta métrica indica el número de páginas que se sincronizaron correctamente.

origen	Métrica	Tipo de métrica
Successes.count		
acknowledgementSetFailures.count	Contador	Si las confirmaciones están habilitadas, esta métrica proporciona el número de páginas que no se pudieron sincronizar.
crawlingTime.avg	Timer	El tiempo que se tardó en revisar todos los cambios nuevos.
pageFetchLatency.avg	Timer	Latencia de la API de obtención de contenido (media).
pageFetchLatency.max	Timer	Latencia de la API de obtención de contenido (máxima).
pagesRequested.count	Contador	Número de invocaciones de la API de obtención de contenido.
pageRequestFailed.count	Contador	Número de solicitudes fallidas de la API de obtención de contenido.
pageRequestedSuccess.count	Contador	Número de solicitudes satisfactorias de la API de obtención de contenido.
searchCallLatency.avg	Timer	Latencia media de llamadas a la API de búsqueda.

origen	Métrica	Tipo de métrica
searchCallLatency.max	Timer	Latencia máxima de llamadas a la API de búsqueda.
searchResultsFound.count	Contador	Número de elementos encontrados en una llamada de búsqueda determinada.
searchRequestsFailed.count	Contador	Los errores en las llamadas a la API de búsqueda cuentan.
authFailures.count	Contador	Número de errores de autenticación.

Conexión de una canalización de Amazon OpenSearch Ingestion a Atlassian, Jira o Confluence mediante la versión 2.0 OAuth

Usa la información de este tema como ayuda para configurar y conectar una canalización de Amazon OpenSearch Ingestion a una cuenta de Jira o Confluence mediante OAuth la autenticación 2.0. Realiza esta tarea cuando estés completando el proceso de uso de una canalización the section called “Requisitos previos” de OpenSearch ingestión con Atlassian Services, pero decidas no usar las credenciales de la clave de la API.

Temas

- Crea una aplicación de integración OAuth 2.0
- Generación y actualización de un token de acceso para desarrolladores de Atlassian

Crea una aplicación de integración OAuth 2.0

Usa el siguiente procedimiento para ayudarte a crear una aplicación de integración OAuth 2.0 en el sitio web para desarrolladores de Atlassian.

Para crear una aplicación de integración OAuth 2.0

1. Inicia sesión en tu cuenta de desarrollador de Atlassian en https://developer.atlassian.com/console/ myapps/.
2. Elige Create, integración 2.0. OAuth
3. En Nombre, escriba un nombre para identificar el propósito de la aplicación.
4. Seleccione la casilla Acepto regirme por las condiciones del desarrollador de Atlassian y seleccione Crear.
5. En el panel de navegación de la izquierda, seleccione Autorización y, a continuación, seleccione Agregar.
6. En URL de devolución de llamada, ingrese cualquier URL, como **`https://www.amazon.com`** o **`https://www.example.com`** y seleccione Guardar cambios.
7. En el panel de navegación de la izquierda, seleccione la página Permisos y, a continuación, en la fila correspondiente a la API de Jira, seleccione Agregar. Luego, seleccione Configurar. Seleccione todos los permisos de lectura de ámbitos clásicos (la lista aparece abajo) y, a continuación, seleccione “Guardar”.
8. Seleccione la pestaña Ámbitos granulares y, a continuación, seleccione Editar ámbitos para abrir el cuadro de diálogo Editar la API de Jira.
9. Seleccione los permisos del complemento de origen que está utilizando:

   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. Seleccione Save.

Para obtener información relacionada, consulta Implementación de la OAuth versión 2.0 (3LO) y Determinar los alcances necesarios para una operación en el sitio web para desarrolladores de Atlassian.

Generación y actualización de un token de acceso para desarrolladores de Atlassian

Utilice el siguiente procedimiento para generar y actualizar un token de acceso para desarrolladores de Atlassian en el sitio web para desarrolladores de Atlassian.

Para generar y actualizar un token de acceso de Jira

1. Inicia sesión en tu cuenta de desarrollador de Atlassian en myapps/. https://developer.atlassian.com/console/
2. Seleccione la aplicación que ha creado en the section called “Crea una aplicación de integración OAuth 2.0”.
3. En el panel de navegación de la izquierda, seleccione Autorización.
4. Copie el valor detallado de la URL de autorización de la API de Atlassian de la parte inferior de la página y péguelo en el editor de texto que prefiera.

   El formato de URL es el siguiente:

   ```
   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. Para `state=YOUR_USER_BOUND_VALUE`, cambie el valor del parámetro por el que prefiera, como state="**`sample_text`**".

   Para obtener más información, consulte ¿Para qué se utiliza el parámetro de estado? en el sitio web para desarrolladores de Atlassian.
6. Tenga en cuenta que en la sección `scope` se enumeran los ámbitos detallados que seleccionó en una tarea anterior. Por ejemplo: `scope=read%3Ajira-work%20read%3Ajira-user%20offline_access`

   `offline_access` indica que desea generar un `refresh_token`.
7. Abra una ventana del navegador Web e introduzca la URL de autorización que haya copiado en la barra de direcciones de la ventana del navegador.
8. Cuando se abra la página de destino, compruebe que la información es correcta y, a continuación, seleccione Aceptar para que lo redirijan a su página de inicio de Jira o Confluence.
9. Una vez que se haya cargado la página de inicio, copia la URL de esta página. Contiene el código de autorización de su aplicación. Utilice este código para generar su token de acceso. La sección completa siguiente a `code=` es el código de autorización.

10. Utilice el siguiente comando cURL para generar el token de acceso. Sustituya *placeholder values* por su propia información.

> **Tip**
>
> También puede usar un servicio de terceros, como 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 respuesta a este comando incluye los valores de `access_code` y `refresh_token`.

Uso de una canalización OpenSearch de ingestión con Amazon Aurora

Puede usar una canalización de OpenSearch ingestión con Amazon Aurora para exportar los datos existentes y transmitir los cambios (como crear, actualizar y eliminar) a los dominios y colecciones de Amazon OpenSearch Service. La canalización de OpenSearch ingestión incorpora una infraestructura de captura de datos de cambios (CDC) para proporcionar una forma de alta escala y baja latencia de transmitir datos de forma continua desde Amazon Aurora. Aurora MySQL y Aurora PostgreSQL son compatibles.

Existen dos formas de utilizar Amazon Aurora como origen para procesar datos: con o sin una instantánea inicial completa. Una instantánea inicial completa es una instantánea de las tablas especificadas y esta instantánea se exporta a Amazon S3. A partir de ahí, una canalización de OpenSearch ingestión los envía a un índice de un dominio o los divide en varios índices de un dominio. Para mantener la OpenSearch coherencia de los datos en Amazon Aurora, la canalización sincroniza todos los eventos de creación, actualización y eliminación de las tablas de los clústeres de Amazon Aurora con los documentos guardados en el OpenSearch índice o los índices.

Cuando utiliza una instantánea inicial completa, la canalización de OpenSearch ingestión primero ingiere la instantánea y, a continuación, comienza a leer los datos de las transmisiones de cambios

de Amazon Aurora. Con el tiempo, recupera y mantiene la coherencia de los datos casi en tiempo real entre Amazon Aurora y OpenSearch.

También puede utilizar la integración de OpenSearch Ingestion con Amazon Aurora para realizar un seguimiento de los cambios, capturar datos e ingerir todas las actualizaciones de Aurora. OpenSearch Elija esta opción si ya tiene una instantánea completa de algún otro mecanismo o si solo quiere capturar todos los cambios en los datos del clúster de Amazon Aurora.

Si elige esta opción, debe configurar el registro binario para Aurora MySQL o configurar la replicación lógica para Aurora PostgreSQL en el clúster.

Temas

- Aurora MySQL
- Aurora PostgreSQL

Aurora MySQL

Complete los siguientes pasos para configurar una canalización OpenSearch de ingestión con Amazon Aurora para Aurora MySQL.

Temas

- Requisitos previos de Aurora MySQL
- Paso 1: configurar el rol de canalización
- Paso 2: crear la canalización
- Coherencia de datos
- Asignación de tipos de datos
- Limitaciones
- Alarmas recomendadas CloudWatch

Requisitos previos de Aurora MySQL

Antes de crear su canalización OpenSearch de ingestión, lleve a cabo los siguientes pasos:

1. Cree un grupo de parámetros de clúster de base de datos de Aurora en Amazon Aurora para configurar el registro binario.

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

Además, compruebe que el parámetro `binlog_transaction_compression` no esté establecido en `ON` y que el parámetro `binlog_row_value_options` no esté establecido en `PARTIAL_JSON`.

2. Seleccione o cree un clúster de base de datos Aurora MySQL y asocie el grupo de parámetros creado en el paso anterior al clúster de base de datos.
3. Configure la retención de registros binarios en 24 horas o más.
4. Configure la autenticación con nombre de usuario y contraseña en su clúster de Amazon Aurora mediante la administración de contraseñas con Aurora y AWS Secrets Manager. También puedes crear una username/password combinación creando un secreto de Secrets Manager.
5. Si utiliza la función de instantáneas iniciales completa, cree una función de IAM AWS KMS key y una función de IAM para exportar datos de Amazon Aurora a Amazon S3.

   El rol de IAM debe tener la siguiente política de permisos:

   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/*"
```

```
            ]
        }
    ]
}
```

El rol también debe tener las siguientes relaciones de confianza:

JSON

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

6. Seleccione o cree un dominio de OpenSearch servicio o una colección OpenSearch sin servidor. Para obtener más información, consulte Creación de dominios OpenSearch de servicio y Creación de colecciones.
7. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster de base de datos de Amazon Aurora en su dominio o colección.

Paso 1: configurar el rol de canalización

Una vez configurados los requisitos previos de la canalización de Amazon Aurora, configure el rol de la canalización que se usará en la configuración de la canalización. Agregue también los siguientes permisos para el origen de Amazon Aurora al rol:

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:*"
    ]
    }
    ]
    }
```

Paso 2: crear la canalización

Configure una canalización OpenSearch de ingestión similar a la siguiente. La canalización de ejemplo especifica un clúster de Amazon Aurora como origen.

```
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
```

Puede utilizar un esquema de Amazon Aurora preconfigurado para crear esta canalización. Para obtener más información, consulte Uso de esquemas.

A fin de utilizar Amazon Aurora como origen, debe configurar el acceso a la VPC para la canalización. La VPC que elija debe ser la misma que usa su origen de Amazon Aurora. A continuación, elija una o varias subredes y uno o varios grupos de seguridad de VPC. Tenga en cuenta que la canalización necesita acceso de red a una base de datos Aurora MySQL, por lo que también debe comprobar que el clúster de Aurora esté configurado con un grupo de seguridad de VPC que permita el tráfico entrante desde el grupo de seguridad de VPC de la canalización al puerto de la base de datos. Para obtener más información, consulte Control de acceso con grupos de seguridad.

Si utiliza la Consola de administración de AWS para crear su canalización, también debe adjuntarla a la VPC para poder utilizar Amazon Aurora como fuente. Para ello, busque la sección Configuración de la red, seleccione la casilla Adjuntar a la VPC y elija el CIDR de una de las

opciones predeterminadas proporcionadas o seleccione una opción propia. Puede usar cualquier CIDR de un espacio de direcciones privado, tal como se define en RFC 1918 Best Current Practice.

Para proporcionar un CIDR personalizado, seleccione Otros en el menú desplegable. Para evitar una colisión en las direcciones IP entre OpenSearch Ingestión y Amazon Aurora, asegúrese de que el CIDR de VPC de Amazon Aurora sea diferente del CIDR de Ingestión. OpenSearch

Para obtener más información, consulte Configuring VPC access for a pipeline.

Coherencia de datos

La canalización garantiza la coherencia de los datos mediante el sondeo continuo o la recepción de cambios del clúster de Amazon Aurora y la actualización de los documentos correspondientes del OpenSearch índice.

OpenSearch La ingestión apoya el end-to-end reconocimiento para garantizar la durabilidad de los datos. Cuando una canalización lee instantáneas o transmisiones, crea particiones de forma dinámica para el procesamiento paralelo. La canalización marca una partición como completa cuando recibe un acuse de recibo después de ingerir todos los registros del dominio o la OpenSearch colección. Si quieres incorporarlos a una colección de búsquedas OpenSearch sin servidor, puedes generar un identificador de documento en la canalización. Si quieres incorporarlo a una colección de series temporales OpenSearch sin servidor, ten en cuenta que la canalización no genera un identificador de documento, por lo que debes omitirlo `document_id: "${getMetadata(\"primary_key\")}"` en la configuración del colector de canalizaciones.

Una canalización OpenSearch de ingestión también asigna las acciones de los eventos entrantes a las correspondientes acciones de indexación masiva para facilitar la ingesta de documentos. Esto mantiene la coherencia de los datos, de modo que cada cambio de datos en Amazon Aurora se concilia con los cambios en OpenSearch los documentos correspondientes.

Asignación de tipos de datos

OpenSearch La canalización de ingestión asigna los tipos de datos de MySQL a representaciones que son adecuadas para el consumo de los dominios o colecciones de OpenSearch servicios. Si no hay ninguna plantilla de mapeo definida OpenSearch, determina OpenSearch automáticamente los tipos de campo con un mapeo dinámico en función del primer documento enviado. También puede definir de forma explícita los tipos de campo que mejor se adapten a sus necesidades OpenSearch mediante una plantilla de mapeo.

La siguiente tabla muestra los tipos de datos de MySQL y los tipos de OpenSearch campo correspondientes. La columna Tipo de OpenSearch campo predeterminado muestra el tipo de campo

correspondiente OpenSearch si no se ha definido un mapeo explícito. En este caso, determina OpenSearch automáticamente los tipos de campo con un mapeo dinámico. La columna Tipo de OpenSearch campo recomendado es el tipo de campo correspondiente que se recomienda especificar de forma explícita en una plantilla de mapeo. Estos tipos de campos están más alineados con los tipos de datos de MySQL y, por lo general, pueden habilitar mejores funciones de búsqueda disponibles en OpenSearch.

Tipo de datos de MySQL	Tipo OpenSearch de campo predeterminado	Tipo OpenSearch de campo recomendado
BIGINT	long	long
BIGINT UNSIGNED	long	largo sin signo
BIT	long	byte, corto, entero o largo según el número de bits
DECIMAL	text	doble o palabra
DOUBLE	float	double
FLOAT	float	float
INT	long	entero
INT UNSIGNED	long	long
MEDIUMIN	long	entero
MEDIUMIN UNSIGNED	long	entero
NUMERIC	text	doble o palabra
SMALLINT	long	short

Tipo de datos de MySQL	Tipo OpenSearch de campo predeterminado	Tipo OpenSearch de campo recomendado
SMALLINT UNSIGNED	long	entero
TINYINT	long	byte
TINYINT UNSIGNED	long	short
BINARIO	text	binario
BLOB	text	binario
CHAR	text	text
ENUM	text	palabra clave
LONGBLOB	text	binario
LONGTEXT	text	text
MEDIUMBLOB	text	binario
MEDIUMTEXT	text	text
SET	text	palabra clave
TEXT	text	text
TINYBLOB	text	binario
TINYTEXT	text	text
VARBINARY	text	binario
VARCHAR	text	text

Tipo de datos de MySQL	Tipo OpenSearch de campo predeterminado	Tipo OpenSearch de campo recomendado
DATE	largo (en milisegundos de época)	date
DATETIME	largo (en milisegundos de época)	date
TIME	largo (en milisegundos de época)	date
TIMESTAM	largo (en milisegundos de época)	date
YEAR	largo (en milisegundos de época)	date
GEOMETR	texto (en formato WKT)	geo_shape
GEOMETR OLLECTIO	texto (en formato WKT)	geo_shape
LINESTRIN G	texto (en formato WKT)	geo_shape
MULTILINE STRING	texto (en formato WKT)	geo_shape
MULTIPOI T	texto (en formato WKT)	geo_shape
MULTIPOL GON	texto (en formato WKT)	geo_shape
POINT	texto (en formato WKT)	geo_point o geo_shape
POLYGON	texto (en formato WKT)	geo_shape
JSON	text	objeto

Te recomendamos que configures la cola de mensajes sin salida (DLQ) en tu canal de ingestión. OpenSearch Si has configurado la cola, el OpenSearch servicio envía a la cola todos los documentos fallidos que no se puedan ingerir debido a errores de mapeo dinámico.

Si las asignaciones automáticas fallan, puede usar `template_type` y `template_content` en su configuración de canalización para definir reglas de asignación explícitas. Como alternativa, puede crear plantillas de asignación directamente en su dominio o colección de búsqueda antes de iniciar la canalización.

Limitaciones

Tenga en cuenta las siguientes limitaciones al configurar una canalización OpenSearch de ingestión para Aurora MySQL:

- La integración solo admite una base de datos MySQL por canalización.
- Actualmente, la integración no admite la ingesta de datos entre regiones; el clúster y el OpenSearch dominio de Amazon Aurora deben estar en el mismo lugar. Región de AWS
- Actualmente, la integración no admite la ingesta de datos entre cuentas; el clúster de Amazon Aurora y la canalización de OpenSearch ingestión deben estar en el mismo lugar. Cuenta de AWS
- Asegúrese de que el clúster de Amazon Aurora tenga habilitada la autenticación mediante Secrets Manager, que es el único mecanismo de autenticación admitido.
- La configuración de canalización existente no se puede actualizar para ingerir datos de una base and/or de datos diferente o de una tabla diferente. Para actualizar la base de datos o el nombre de la tabla de una canalización, debe detener la canalización y reiniciarla con una configuración actualizada o crear una canalización nueva.
- Por lo general, las instrucciones en lenguaje de definición de datos (DDL) no son compatibles. La coherencia de datos no se mantendrá si:
  - Se cambian las claves principales (add/delete/rename).
  - Las tablas se eliminan o se truncan.
  - Se cambian los nombres de las columnas o los tipos de datos.
- Si las tablas MySQL que se van a sincronizar no tienen definidas las claves principales, no se garantiza la coherencia de datos. Deberá definir correctamente la `document_id` opción personalizada en la configuración del OpenSearch sumidero para poder updates/deletes sincronizarla OpenSearch.
- No se admiten las referencias a claves externas con acciones de eliminación en cascada y pueden provocar inconsistencias en los datos entre Aurora MySQL y. OpenSearch
- Versiones compatibles: Aurora MySQL 3.05.2 y versiones posteriores.

Alarmas recomendadas CloudWatch

Se recomiendan las siguientes CloudWatch métricas para supervisar el rendimiento del proceso de ingestión. Estas métricas pueden ayudarlo a identificar la cantidad de datos procesados a partir de exportaciones y flujos, los errores al procesar eventos de transmisión y exportación y el número de documentos escritos en el destino. Puedes configurar CloudWatch alarmas para realizar una acción cuando una de estas métricas supere un valor específico durante un período de tiempo específico.

Métrica	Description (Descripción)
pipeline-name .RDS. Se han modificado las credenciales	Esta métrica indica la frecuencia con la que se rotan los secretos. AWS
pipeline-name .rds. executorRefreshErrors	Esta métrica indica errores al actualizar AWS los secretos.
pipeline-name .rds. exportRecordsTotal	Esta métrica indica el número de registros exportados desde Amazon Aurora.
pipeline-name .rds. exportRecordsProcessed	Esta métrica indica la cantidad de registros procesados por la canalización OpenSearch de ingestión.
pipeline-name .rds. exportRecordProcessingErrores	Esta métrica indica el número de errores de procesamiento en una canalización de OpenSearch ingestión al leer los datos de un clúster de Amazon Aurora.
pipeline-name .rds.	Esta métrica indica el número total de registros de exportación procesados correctamente.

Métrica	Description (Descripción)
exportRecordsSuccessTotal	
pipeline-name .rds.exportRecordsFailedTotal	Esta métrica indica el número total de registros que no se han podido procesar.
pipeline-name .RDS. Bytes recibidos	Esta métrica indica el número total de bytes recibidos por una canalización de ingestión. OpenSearch
pipeline-name .rds.bytes procesados	Esta métrica indica el número total de bytes procesados por una canalización de ingestión. OpenSearch
pipeline-name .rds.streamRecordsSuccessTotal	Esta métrica indica el número de registros del flujo procesados correctamente.
pipeline-name .rds.streamRecordsFailedTotal	Esta métrica indica el número total de registros del flujo que no se han podido procesar.

Aurora PostgreSQL

Complete los siguientes pasos para configurar una canalización OpenSearch de ingestión con Amazon Aurora para Aurora PostgreSQL.

Temas

- Requisitos previos de Aurora PostgreSQL
- Paso 1: configurar el rol de canalización
- Paso 2: crear la canalización

- Coherencia de datos
- Asignación de tipos de datos
- Limitaciones
- CloudWatch Alarmas recomendadas

Requisitos previos de Aurora PostgreSQL

Antes de crear la canalización de OpenSearch ingestión, lleve a cabo los siguientes pasos:

1. Cree un grupo de parámetros de clúster de base de datos personalizado en Amazon Aurora para configurar la replicación lógica.

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

2. Seleccione o cree un clúster de base de datos Aurora PostgreSQL y asocie el grupo de parámetros creado en el paso 1 al clúster de base de datos.
3. Configure la autenticación con nombre de usuario y contraseña en su clúster de Amazon Aurora mediante la administración de contraseñas con Aurora y AWS Secrets Manager. También puedes crear una username/password combinación creando un secreto de Secrets Manager.
4. Si utiliza la función de instantáneas iniciales completa, cree una función de IAM AWS KMS key y una función de IAM para exportar datos de Amazon Aurora a Amazon S3.

   El rol de IAM debe tener la siguiente política de permisos:

   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/*"
            ]
        }
    ]
}
```

El rol también debe tener las siguientes relaciones de confianza:

JSON

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

5. Seleccione o cree un dominio de OpenSearch servicio o una colección OpenSearch sin servidor. Para obtener más información, consulte Creación de dominios OpenSearch de servicio y Creación de colecciones.
6. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster de base de datos de Amazon Aurora en su dominio o colección.

Paso 1: configurar el rol de canalización

Una vez configurados los requisitos previos de la canalización de Amazon Aurora, configure el rol de la canalización que se usará en la configuración de la canalización. Agregue también los siguientes permisos para el origen de Amazon Aurora al rol:

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:*"
    ]
    }
    ]
    }
```

Paso 2: crear la canalización

Configure una canalización de OpenSearch ingestión como la siguiente, que especifique el clúster de Aurora PostgreSQL como origen.

```
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**
>
> Puede utilizar un esquema de Amazon Aurora preconfigurado para crear esta canalización. Para obtener más información, consulte Uso de esquemas.

A fin de utilizar Amazon Aurora como origen, debe configurar el acceso a la VPC para la canalización. La VPC que elija debe ser la misma que usa su origen de Amazon Aurora. A continuación, elija una o varias subredes y uno o varios grupos de seguridad de VPC. Tenga en cuenta que la canalización necesita acceso de red a una base de datos Aurora MySQL, por lo que también debe comprobar que el clúster de Aurora esté configurado con un grupo de seguridad de

VPC que permita el tráfico entrante desde el grupo de seguridad de VPC de la canalización al puerto de la base de datos. Para obtener más información, consulte Control de acceso con grupos de seguridad.

Si utiliza la Consola de administración de AWS para crear su canalización, también debe adjuntarla a la VPC para poder utilizar Amazon Aurora como fuente. Para ello, busque la sección Configuración de la red, seleccione Adjuntar a la VPC y elija el CIDR de una de las opciones predeterminadas proporcionadas o seleccione una opción propia. Puede usar cualquier CIDR de un espacio de direcciones privado, tal como se define en RFC 1918 Best Current Practice.

Para proporcionar un CIDR personalizado, seleccione Otros en el menú desplegable. Para evitar una colisión en las direcciones IP entre OpenSearch Ingestión y Amazon Aurora, asegúrese de que el CIDR de VPC de Amazon Aurora sea diferente del CIDR de Ingestión. OpenSearch

Para obtener más información, consulte Configuring VPC access for a pipeline.

Coherencia de datos

La canalización garantiza la coherencia de los datos mediante el sondeo continuo o la recepción de cambios del clúster de Amazon Aurora y la actualización de los documentos correspondientes del OpenSearch índice.

OpenSearch La ingestión apoya el end-to-end reconocimiento para garantizar la durabilidad de los datos. Cuando una canalización lee instantáneas o transmisiones, crea particiones de forma dinámica para el procesamiento paralelo. La canalización marca una partición como completa cuando recibe un acuse de recibo después de ingerir todos los registros del dominio o la OpenSearch colección. Si quieres incorporarlos a una colección de búsquedas OpenSearch sin servidor, puedes generar un identificador de documento en la canalización. Si quieres incorporarlo a una colección de series temporales OpenSearch sin servidor, ten en cuenta que la canalización no genera un identificador de documento, por lo que debes omitirlo `document_id: "${getMetadata(\"primary_key\")}"` en la configuración del colector de canalizaciones.

Una canalización OpenSearch de ingestión también asigna las acciones de los eventos entrantes a las correspondientes acciones de indexación masiva para facilitar la ingesta de documentos. Esto mantiene la coherencia de los datos, de modo que cada cambio de datos en Amazon Aurora se concilia con los cambios en OpenSearch los documentos correspondientes.

Asignación de tipos de datos

OpenSearch La canalización de ingestión asigna los tipos de datos de Aurora PostgreSQL a representaciones que son adecuadas OpenSearch para el consumo de los dominios o colecciones

de servicios. Si no hay ninguna plantilla de mapeo definida OpenSearch, determine OpenSearch automáticamente los tipos de campo con un mapeo dinámico basado en el primer documento enviado. También puede definir de forma explícita los tipos de campo que mejor se adapten a sus necesidades OpenSearch mediante una plantilla de mapeo.

La siguiente tabla muestra los tipos de datos de Aurora PostgreSQL y OpenSearch los tipos de campo correspondientes. La columna Tipo de OpenSearch campo predeterminado muestra el tipo de campo correspondiente OpenSearch si no se ha definido un mapeo explícito. En este caso, determina OpenSearch automáticamente los tipos de campo con un mapeo dinámico. La columna Tipo de OpenSearch campo recomendado es el tipo de campo que se recomienda especificar de forma explícita en una plantilla de mapeo. Estos tipos de campos están más alineados con los tipos de datos de Aurora PostgreSQL y, por lo general, pueden habilitar mejores funciones de búsqueda disponibles en. OpenSearch

Tipo de datos de Aurora PostgreSQL	Tipo de campo predeterminado OpenSearch	Tipo OpenSearch de campo recomendado
smallint	long	short
entero	long	entero
bigint	long	long
decimal	text	doble o palabra
numeric[(p, s)]	text	doble o palabra
real	float	float
double precision	float	double
smallserial	long	short

Tipo de datos de Aurora PostgreSQL	Tipo de campo predeterminado OpenSearch	Tipo OpenSearch de campo recomendado
serial	long	entero
bigserial	long	long
money	objeto	objeto
character varying(n)	text	text
varchar(n)	text	text
character (n)	text	text
char(n)	text	text
bpchar (n)	text	text
BPCHAR	text	text
text	text	text
enum	text	text
bytea	text	binario
timestamp [(p)] [sin zona horaria]	largo (en milisegundos de época)	date

Tipo de datos de Aurora PostgreSQL	Tipo de campo predeterminado OpenSearch	Tipo OpenSearch de campo recomendado
timestamp [(p)] con zona horaria	largo (en milisegundos de época)	date
date	largo (en milisegundos de época)	date
time [(p)] [sin zona horaria]	largo (en milisegundos de época)	date
time [(p)] con zona horaria	largo (en milisegundos de época)	date
intervalo [campos] [(p)]	texto (ISO8601 formato)	text
booleano	booleano	booleano
point	texto (en formato WKT)	geo_shape
línea	texto (en formato WKT)	geo_shape
LSEG	texto (en formato WKT)	geo_shape
BOX	texto (en formato WKT)	geo_shape
path	texto (en formato WKT)	geo_shape
polígono	texto (en formato WKT)	geo_shape
CIRCLE	objeto	objeto

Tipo de datos de Aurora PostgreSQL	Tipo de campo predeterminado OpenSearch	Tipo OpenSearch de campo recomendado
cidr	text	text
inet	text	text
macaddr	text	text
macaddr8	text	text
bit(n)	long	byte, corto, entero o largo (según el número de bits)
bit varying(n)	long	byte, corto, entero o largo (según el número de bits)
json	objeto	objeto
jsonb	objeto	objeto
jsonpath	text	text

Te recomendamos que configures la cola de mensajes sin salida (DLQ) en tu canal de ingestión. OpenSearch Si has configurado la cola, el OpenSearch servicio envía a la cola todos los documentos fallidos que no se puedan ingerir debido a errores de mapeo dinámico.

En caso de que las asignaciones automáticos fallen, puede usar `template_type` y `template_content` en su configuración de canalización para definir reglas de asignación explícitas. Como alternativa, puede crear plantillas de asignación directamente en su dominio o colección de búsqueda antes de iniciar la canalización.

Limitaciones

Tenga en cuenta las siguientes limitaciones al configurar una canalización de OpenSearch ingestión para Aurora PostgreSQL:

- La integración solo admite una base de datos PostgreSQL de Aurora por canalización.
- Actualmente, la integración no admite la ingesta de datos entre regiones; el clúster y el OpenSearch dominio de Amazon Aurora deben estar en el mismo lugar. Región de AWS
- Actualmente, la integración no admite la ingesta de datos entre cuentas; el clúster de Amazon Aurora y la canalización de OpenSearch ingestión deben estar en el mismo lugar. Cuenta de AWS
- Asegúrese de que el clúster de Amazon Aurora tenga habilitada la autenticación mediante AWS Secrets Manager, que es el único mecanismo de autenticación admitido.
- La configuración de canalización existente no se puede actualizar para ingerir datos de una base and/or de datos diferente o de una tabla diferente. Para actualizar la base de datos o el nombre de la tabla de una canalización, debe detener la canalización y reiniciarla con una configuración actualizada o crear una canalización nueva.
- Por lo general, las instrucciones en lenguaje de definición de datos (DDL) no son compatibles. La coherencia de datos no se mantendrá si:
  - Se cambian las claves principales (add/delete/rename).
  - Las tablas se eliminan o se truncan.
  - Se cambian los nombres de las columnas o los tipos de datos.
- Si las tablas de PostgreSQL de Aurora que se van a sincronizar no tienen definidas las claves principales, no se garantiza la coherencia de datos. Deberá definir correctamente la `document_id` opción OpenSearch y la configuración del receptor para poder updates/deletes sincronizarla OpenSearch.
- Versiones compatibles: Aurora PostgreSQL 16.4 y versiones posteriores.

CloudWatch Alarmas recomendadas

Se recomiendan las siguientes CloudWatch métricas para supervisar el rendimiento del proceso de ingestión. Estas métricas pueden ayudarlo a identificar la cantidad de datos procesados a partir de exportaciones y flujos, los errores al procesar eventos de transmisión y exportación y el número de documentos escritos en el destino. Puedes configurar CloudWatch alarmas para realizar una acción cuando una de estas métricas supere un valor específico durante un período de tiempo específico.

Métrica	Description (Descripción)
`pipeline-name` .RDS. Se	Esta métrica indica la frecuencia con la que se rotan los secretos. AWS

Métrica	Description (Descripción)
han modificado las credenciales	
pipeline-name .rds. executorRefreshErrors	Esta métrica indica errores al actualizar AWS los secretos.
pipeline-name .rds. exportRecordsTotal	Esta métrica indica el número de registros exportados desde Amazon Aurora.
pipeline-name .rds. exportRecordsProcessed	Esta métrica indica la cantidad de registros procesados por la canalización OpenSearch de ingestión.
pipeline-name .rds. exportRecordProcessingErrores	Esta métrica indica el número de errores de procesamiento en una canalización de OpenSearch ingestión al leer los datos de un clúster de Amazon Aurora.
pipeline-name .rds. exportRecordsSuccessTotal	Esta métrica indica el número total de registros de exportación procesados correctamente.
pipeline-name .rds. exportRecordsFailedTotal	Esta métrica indica el número total de registros que no se han podido procesar.
pipeline-name .RDS. Bytes recibidos	Esta métrica indica el número total de bytes recibidos por una canalización de ingestión. OpenSearch

Métrica	Description (Descripción)
pipeline-name .rds.bytes procesados	Esta métrica indica el número total de bytes procesados por una canalizac ión de ingestión. OpenSearch
pipeline-name .rds. streamRec ordsSuccessTotal	Esta métrica indica el número de registros del flujo procesados correctam ente.
pipeline-name .rds. streamRecordsFaile dTotal	Esta métrica indica el número total de registros del flujo que no se han podido procesar.

Uso de una canalización OpenSearch de ingestión con Amazon DynamoDB

Puede usar el complemento DynamoDB para transmitir eventos de tablas, como creaciones, actualizaciones y eliminaciones, a los dominios de OpenSearch Amazon Service y a las colecciones de Amazon Serverless. OpenSearch La canalización utiliza la captura de datos de cambios (CDC) para la transmisión de alta escala y baja latencia.

Puede procesar datos de DynamoDB con o sin una instantánea inicial completa.

- Con una instantánea completa, DynamoDB point-in-time utiliza la recuperación (PITR) para crear una copia de seguridad y la carga en Amazon S3. OpenSearch Luego, Ingestion indexa la instantánea en uno o varios índices. OpenSearch Para mantener la coherencia, la canalización sincroniza todos los cambios de DynamoDB con. OpenSearch Esta opción requiere que habilite PITR y DynamoDB Streams.
- Sin una instantánea: OpenSearch Inestion solo transmite los eventos nuevos de DynamoDB. Elija esta opción si ya tiene una instantánea o si necesita una transmisión en tiempo real sin datos históricos. Esta opción requiere que habilite únicamente DynamoDB Streams.

Para obtener más información, consulte la integración de DynamoDB Zero-ETL con OpenSearch Amazon Service en la Guía para desarrolladores.Amazon DynamoDB

Temas

- Requisitos previos
- Paso 1: configurar el rol de canalización
- Paso 2: crear la canalización
- Coherencia de datos
- Asignación de tipos de datos
- Limitaciones
- CloudWatch Alarmas recomendadas para DynamoDB

Requisitos previos

Para configurar la canalización, debe tener una tabla de DynamoDB con DynamoDB Streams activado. La transmisión debe usar el tipo de vista de transmisión `NEW_IMAGE`. Sin embargo, las canalizaciones OpenSearch de ingestión también pueden transmitir eventos `NEW_AND_OLD_IMAGES` si este tipo de vista de transmisión se ajusta a su caso de uso.

Si utilizas instantáneas, también debes habilitar la point-in-time recuperación en tu mesa. Para obtener más información, consulte Crear una tabla, Habilitar la point-in-time recuperación y Habilitar una transmisión en la Guía para desarrolladores de Amazon DynamoDB.

Paso 1: configurar el rol de canalización

Una vez configurada la tabla de DynamoDB, configure el rol de canalización que desee usar en la configuración de canalización y añada los siguientes permisos de DynamoDB al rol:

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/*"
            ]
        }
    ]
}
```

También puede usar una clave administrada por el AWS KMS cliente para cifrar los archivos de datos de exportación. Para descifrar los objetos exportados, especifique `s3_sse_kms_key_id` para el ID de clave en la configuración de exportación de la canalización con el siguiente formato: `arn:aws:kms:region:account-id:key/my-key-id`. La siguiente política incluye los permisos necesarios para utilizar una clave administrada por el cliente:

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

Paso 2: crear la canalización

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, que especifica DynamoDB como origen. Este ejemplo de canalización incorpora los datos de la `table-a` con la instantánea de PITR, seguidos de los eventos de DynamoDB Streams. Una posición inicial de `LATEST` indica que la canalización debe leer los datos más recientes de 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"
```

Puede utilizar un esquema preconfigurado de DynamoDB para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Coherencia de datos

OpenSearch La ingestión admite el end-to-end reconocimiento para garantizar la durabilidad de los datos. Cuando una canalización lee instantáneas o transmisiones, crea particiones de forma dinámica para el procesamiento paralelo. La canalización marca una partición como completa cuando recibe un acuse de recibo después de ingerir todos los registros del dominio o la OpenSearch colección.

Si quieres incorporarlos a una colección de búsquedas OpenSearch sin servidor, puedes generar un identificador de documento en la canalización. Si quieres incorporarlo a una colección de series temporales OpenSearch sin servidor, ten en cuenta que la canalización no genera un identificador de documento.

Una canalización OpenSearch de ingestión también asigna las acciones de los eventos entrantes a las correspondientes acciones de indexación masiva para facilitar la ingesta de documentos. Esto mantiene la coherencia de los datos, de modo que cada cambio de datos en DynamoDB se concilia con los cambios correspondientes en el documento. OpenSearch

Asignación de tipos de datos

OpenSearch El servicio asigna dinámicamente los tipos de datos de cada documento entrante al tipo de datos correspondiente en DynamoDB. En la siguiente tabla se muestra cómo OpenSearch Service asigna automáticamente varios tipos de datos.

Tipo de datos:	OpenSearch	DynamoDB
Número	OpenSearch mapea automáticamente los datos numéricos. Si el número es un número entero, lo OpenSearch asigna como un valor largo. Si el número es	DynamoDB admite números.

Tipo de datos:	OpenSearch	DynamoDB
	fraccionario, lo OpenSearch mapea como un valor flotante. OpenSearch mapea dinámicamente varios atributos en función del primer documento enviado. Si tiene una combinación de tipos de datos para el mismo atributo en DynamoDB, como un número entero y un número fraccionario, es posible que se produzca un error en la asignación. Por ejemplo, si el primer documento tiene un atributo que es un número entero y un documento posterior tiene el mismo atributo que un número fracciona rio, OpenSearch no podrá ingerir el segundo documento. En estos casos, debe proporcionar una plantilla de asignación explícita, como la siguiente: `{ "template": { "mappings": { "properties": { "MixedNumberAttribute": { "type": "float" } } } } }` Si necesita una precisión doble, utilice la asignación de campos de tipo cadena. No hay ningún tipo numérico equivalente	

Tipo de datos:	OpenSearch	DynamoDB
	que admita una precisión de 38 dígitos. OpenSearch	
Conjunto de números	OpenSearch asigna automáticamente un conjunto de números a una matriz de valores largos o flotantes. Al igual que con los números escalares, esto depende de si el primer número incorporado es un número entero o un número fraccionario. Puede proporcio nar asignaciones para conjuntos de números de la misma manera que asigna cadenas escalares.	DynamoDB admite tipos que represent an conjuntos de números.
Cadena	OpenSearch mapea automáticamente los valores de cadena como texto. En algunas situaciones, como en el caso de los valores enumerados, puede asignarlos al tipo de palabra clave. En el siguiente ejemplo, se muestra cómo asignar un atributo de DynamoDB `PartType` denominado a una palabra clave. OpenSearch `{ "template": { "mappings": { "properties": { "PartType": { "type": "keyword" } } } } }`	DynamoDB admite cadenas.

Tipo de datos:	OpenSearch	DynamoDB
Conjunto de cadenas	OpenSearch asigna automáticamente un conjunto de cadenas a una matriz de cadenas. Puede proporcionar asignaciones para conjuntos de cadenas de la misma manera que asigna cadenas escalares.	DynamoDB admite tipos que representan conjuntos de cadenas.
Binario	OpenSearch mapea automáticamente los datos binarios como texto. Puede proporcionar un mapeo para escribirlos como campos binarios OpenSearch. El siguiente ejemplo muestra cómo asignar un atributo de DynamoDB `ImageData` denominado a OpenSearch un campo binario. `{ "template": { "mappings": { "properties": { "ImageData": { "type": "binary" } } } } }`	DynamoDB es compatible con los atributos de tipo binarios.
Conjunto binario	OpenSearch asigna automáticamente un conjunto binario a una matriz de datos binarios en forma de texto. Puede proporcionar asignaciones para conjuntos de números de la misma manera que asigna el binario escalar.	DynamoDB admite tipos que representan conjuntos de valores binarios.

Tipo de datos:	OpenSearch	DynamoDB
Booleano	OpenSearch asigna un tipo booleano de DynamoDB a un tipo booleano. OpenSearch	DynamoDB admite atributos de tipo booleano.
Nulo	OpenSearch puede ingerir documento s del tipo null de DynamoDB. Guarda el valor como un valor nulo en el documento. No hay ninguna asignació n para este tipo y este campo no está indexado ni se puede buscar en él. Si se utiliza el mismo nombre de atributo para un tipo nulo y, posteriormente, se cambia a un tipo diferente, como una cadena, OpenSearch crea una asignació n dinámica para el primer valor no nulo. Los valores subsiguientes pueden seguir siendo valores nulos de DynamoDB.	DynamoDB admite atributos de tipo nulo.

Tipo de datos:	OpenSearch	DynamoDB
Asignación	OpenSearch asigna los atributos del mapa de DynamoDB a campos anidados. Las mismas asignaciones se aplican dentro de un campo anidado. El siguiente ejemplo asigna una cadena de un campo anidado a un tipo de palabra clave en: OpenSearch `{ "template": { "mappings": { "properties": { "AdditionalDescriptions": { "properties": { "PartType": { "type": "keyword" } } } } } } }`	DynamoDB admite atributos de tipo asignación.

Tipo de datos:	OpenSearch	DynamoDB
Enumeración	OpenSearch proporciona resultados diferentes para las listas de DynamoDB, según el contenido de la lista. Cuando una lista contiene todos los tipos escalares del mismo tipo (por ejemplo, una lista de todas las cadenas), OpenSearch ingiere la lista como una matriz de ese tipo. Esto funciona para los tipos cadena, número, booleano y nulo. Las restricciones para cada uno de estos tipos son las mismas que las restricciones para un escalar de ese tipo. También puede proporcionar asignaciones para listas de mapas con la misma asignación que usaría para un mapa. No puede proporcionar una lista de tipos mixtos.	DynamoDB admite atributos de tipo lista.

Tipo de datos:	OpenSearch	DynamoDB
Establezca	OpenSearch proporciona resultados diferentes para los conjuntos de DynamoDB en función del contenido del conjunto. Cuando un conjunto contiene todos los tipos escalares del mismo tipo (por ejemplo, un conjunto de todas las cadenas), OpenSearch ingiere el conjunto como una matriz de ese tipo. Esto funciona para los tipos cadena, número, booleano y nulo. Las restricciones para cada uno de estos tipos son las mismas que las restricciones para un escalar de ese tipo. También puede proporcionar asignaciones para conjuntos de mapas con la misma asignación que usaría para un mapa. No puede proporcionar un conjunto de tipos mixtos.	DynamoDB admite tipos que representan conjuntos.

Le recomendamos que configure la cola de mensajes sin salida (DLQ) en su canalización de ingestión. OpenSearch Si has configurado la cola, el OpenSearch servicio envía a la cola todos los documentos fallidos que no se puedan ingerir debido a errores de mapeo dinámico.

En caso de que las asignaciones automáticos fallen, puede usar `template_type` y `template_content` en su configuración de canalización para definir reglas de asignación explícitas. Como alternativa, puede crear plantillas de asignación directamente en su dominio o colección de búsqueda antes de iniciar la canalización.

Limitaciones

Tenga en cuenta las siguientes limitaciones al configurar una canalización de OpenSearch ingestión para DynamoDB:

- La integración OpenSearch de ingestión con DynamoDB actualmente no admite la ingestión entre regiones. La tabla de DynamoDB OpenSearch y la canalización de ingestión deben estar en la misma posición. Región de AWS
- La tabla de DynamoDB OpenSearch y la canalización de ingestión deben estar en la misma posición. Cuenta de AWS
- Una canalización OpenSearch de ingestión solo admite una tabla de DynamoDB como fuente.
- DynamoDB Streams solo almacena datos en un registro durante un máximo de 24 horas. Si la ingesta de una instantánea inicial de una tabla grande demora 24 horas o más, se producirá una pérdida inicial de datos. Para mitigar esta pérdida de datos, calcule el tamaño de la tabla y configure las unidades informáticas adecuadas de las canalizaciones de OpenSearch ingestión.

CloudWatch Alarmas recomendadas para DynamoDB

Se recomiendan las siguientes CloudWatch métricas para supervisar el rendimiento del proceso de ingestión. Estas métricas pueden ayudarle a identificar la cantidad de datos procesados a partir de exportaciones y flujos, los errores al procesar eventos de transmisión y exportación y la cantidad de documentos escritos en el destino. Puedes configurar CloudWatch alarmas para realizar una acción cuando una de estas métricas supere un valor específico durante un período de tiempo específico.

Métrica	Description (Descripción)
`dynamodb-pipeline.BlockingBuffer.bufferUsage.value`	Indica qué cantidad del búfer se está utilizando.
`dynamodb-pipeline.dynamodb.activeExportS3ObjectConsumers.value`	Muestra el número total de objetos de Amazon S3 OCUs que están procesando activamente para la exportación.
`dynamodb-pipeline.dynamodb.bytesProcessed.count`	Recuento de bytes procesados desde el origen de DynamoDB.

Métrica	Description (Descripción)
`dynamodb-pipeline.dynamodb.changeEventsProcessed.count`	Número de eventos de cambio procesados desde el flujo de DynamoDB.
`dynamodb-pipeline.dynamodb.changeEventsProcessingErrors.count`	Número de errores de eventos de cambio procesados desde DynamoDB.
`dynamodb-pipeline.dynamodb.exportJobFailure.count`	Número de intentos de envío de trabajos de exportación que han fallado.
`dynamodb-pipeline.dynamodb.exportJobSuccess.count`	Número de trabajos de exportación que se han enviado correctamente.
`dynamodb-pipeline.dynamodb.exportRecordsProcessed.count`	Número total de registros procesados desde la exportación.
`dynamodb-pipeline.dynamodb.exportRecordsTotal.count`	Número total de registros exportados desde DynamoDB, esencial para el seguimiento de los volúmenes de exportación de datos.
`dynamodb-pipeline.dynamodb.exportS3ObjectsProcessed.count`	Número total de archivos de datos de exportación que se han procesado correctamente desde Amazon S3.
`dynamodb-pipeline.opensearch.bulkBadRequestErrors.count`	Recuento de errores durante las solicitudes masivas debido a un formato incorrecto de la solicitud.
`dynamodb-pipeline.opensearch.bulkRequestLatency.avg`	Latencia media de las solicitudes de escritura masiva realizadas a OpenSearch.
`dynamodb-pipeline.opensearch.bulkRequestNotFoundErrors.count`	Número de solicitudes masivas que fallaron porque no se pudieron encontrar los datos de destino.

Métrica	Description (Descripción)
dynamodb-pipeline.opensearch.bulkRequestNumberOfRetries.count	Número de reintentos realizados por las canalizaciones OpenSearch de ingestión para escribir el clúster. OpenSearch
dynamodb-pipeline.opensearch.bulkRequestSizeBytes.sum	Tamaño total en bytes de todas las solicitudes masivas realizadas a. OpenSearch
dynamodb-pipeline.opensearch.documentErrors.count	Número de errores al enviar documentos a OpenSearch. Los documentos que causan los errores se enviarán a DLQ.
dynamodb-pipeline.opensearch.documentsSuccess.count	Número de documentos escritos correctamente en un OpenSearch clúster o colección.
dynamodb-pipeline.opensearch.documentsSuccessFirstAttempt.count	Número de documentos indexados correctamente OpenSearch en el primer intento.
dynamodb-pipeline.opensearch.documentsVersionConflictErrors.count	Recuento de errores debidos a conflictos de versiones en los documentos durante el procesamiento.
dynamodb-pipeline.opensearch.PipelineLatency.avg	Latencia media de la canalización de OpenSearch ingestión para procesar los datos desde el origen hasta escribirlos en el destino.
dynamodb-pipeline.opensearch.PipelineLatency.max	Latencia máxima de la canalización de OpenSearch ingestión para procesar los datos desde el origen hasta escribir el destino.
dynamodb-pipeline.opensearch.recordsIn.count	Recuento de registros ingresados correctamente. OpenSearch Esta métrica es esencial para realizar un seguimiento del volumen de datos que se procesan y almacenan.
dynamodb-pipeline.opensearch.s3.dlqS3RecordsFailed.count	Número de registros que no se pudieron escribir en DLQ.

Métrica	Description (Descripción)
`dynamodb-pipeline.opensearch.s3.dlqS3RecordsSuccess.count`	Número de registros que se escriben en DLQ.
`dynamodb-pipeline.opensearch.s3.dlqS3RequestLatency.count`	Recuento de las mediciones de latencia de las solicitudes a la cola de cartas muertas de Amazon S3.
`dynamodb-pipeline.opensearch.s3.dlqS3RequestLatency.sum`	Latencia total de todas las solicitudes a la cola de cartas muertas de Amazon S3
`dynamodb-pipeline.opensearch.s3.dlqS3RequestSizeBytes.sum`	Tamaño total en bytes de todas las solicitudes realizadas a la cola de cartas muertas de Amazon S3.
`dynamodb-pipeline.recordsProcessed.count`	Número total de registros procesados en la canalización, una métrica clave para el rendimiento general.
`dynamodb.changeEventsProcessed.count`	No se está recopilando ningún registro de las transmisiones de DynamoDB. Esto puede deberse a que no hay actividad sobre la mesa, a que se esté realizando una exportación o a un problema al acceder a las transmisiones de DynamoDB.
`dynamodb.exportJobFailure.count`	Falló el intento de activar una exportación a S3.
`dynamodb-pipeline.opensearch.bulkRequestInvalidInputErrors.count`	El recuento de errores en las solicitudes masivas OpenSearch se debe a una entrada no válida, algo crucial para supervisar la calidad de los datos y los problemas operativos.

Métrica	Description (Descripción)
`opensearch.EndToEndLatency.avg`	La latencia de extremo a extremo es superior a la deseada para la lectura de transmisiones de DynamoDB. Esto podría deberse a un OpenSearch clúster de menor escala o a una capacidad máxima de OCU de canalización demasiado baja para el rendimiento de la WCU de la tabla de DynamoDB. Esta latencia de extremo a extremo será alta después de una exportación y debería disminuir con el tiempo a medida que se vaya actualizando con las últimas transmisiones de DynamoDB.

Uso de una canalización OpenSearch de ingestión con Amazon DocumentDB

Puedes usar el complemento DocumentDB para transmitir los cambios en los documentos, como las creaciones, actualizaciones y eliminaciones, a Amazon Service. OpenSearch La canalización admite la captura de datos de cambios (CDC), si está disponible, o el sondeo de API para la transmisión a alta escala y baja latencia.

Puede procesar los datos con o sin una instantánea inicial completa. Una instantánea completa captura una colección completa de Amazon DocumentDB y la carga en Amazon S3. A continuación, la canalización envía los datos a uno o más OpenSearch índices. Una vez que ingiere la instantánea, la canalización sincroniza los cambios en curso para mantener la coherencia y, finalmente, recupera las actualizaciones casi en tiempo real.

Si ya tiene una instantánea completa de otro origen o solo necesita procesar eventos nuevos, puede transmitir sin una instantánea. En este caso, la canalización lee directamente los flujos de cambios de Amazon DocumentDB sin una carga masiva inicial.

Debe habilitar una transmisión de cambios en la recopilación de Amazon DocumentDB si habilita una transmisión. Sin embargo, si solo realiza una carga completa o una exportación, no necesitará ningún flujo de cambios.

Requisitos previos

Antes de crear tu canalización OpenSearch de ingestión, lleva a cabo los siguientes pasos:

1. Para crear un clúster de Amazon DocumentDB con permiso para leer datos, siga los pasos que se indican en Creación de un clúster de Amazon DocumentDB en la Guía para desarrolladores de Amazon DocumentDB. Si utiliza la infraestructura de CDC, configure su clúster de Amazon DocumentDB para publicar flujos de cambios.
2. Habilitar TLS en su clúster de Amazon DocumentDB.
3. Configure un CIDR de VPC de un espacio de direcciones privado para usarlo con Ingestión. OpenSearch
4. Configure la autenticación en su clúster de Amazon DocumentDB con. AWS Secrets Manager Para habilitar la rotación de secretos, siga los pasos de Automatically Rotating Passwords for Amazon DocumentDB. Para obtener más información, consulte Acceso a la base de datos mediante el control de acceso basado en roles y Seguridad en Amazon DocumentDB.
5. Si utiliza un flujo de cambios para suscribirse a los cambios de datos de la colección de Amazon DocumentDB, puede ampliar el periodo de retención hasta 7 días con el parámetro `change_stream_log_retention_duration` para evitar la pérdida de datos. Los eventos de flujos de cambios se almacenan durante 3 horas, de forma predeterminada, una vez registrado el evento, lo que no es tiempo suficiente para colecciones de gran tamaño. Para modificar el periodo de retención del flujo de cambios, consulte Modifying the change stream log retention duration.
6. Cree un dominio OpenSearch de servicio o una colección OpenSearch sin servidor. Para obtener más información, consulte the section called “Creación de dominios OpenSearch de servicio” y the section called “Creación de colecciones”.
7. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster de Amazon DocumentDB en su dominio o colección.

   El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización, que creará en el paso siguiente, escriba datos en un dominio. Asegúrese de actualizar el `resource` con su propio 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"
      ]
    }
  ]
}
```

Para crear un rol de IAM con los permisos correctos para acceder a los datos de escritura de la colección o dominio, consulte the section called “Configuración de roles y usuarios”.

Paso 1: configurar el rol de canalización

Una vez configurados los requisitos previos de la canalización de Amazon DocumentDB, configure el rol de la canalización que desee usar en la configuración de la canalización y agregue los siguientes permisos de Amazon DocumentDB al rol:

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"
                }
            }
        }
    ]
}
```

Debe proporcionar los permisos de Amazon EC2 anteriores en la función de IAM que utilice para crear la canalización de OpenSearch ingestión, ya que la canalización utiliza estos permisos para crear y eliminar una interfaz de red en la VPC. La canalización solo puede acceder al clúster de Amazon DocumentDB a través de esta interfaz de red.

Paso 2: crear la canalización

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, que especifica Amazon DocumentDB como origen. Tenga en cuenta que, para rellenar el nombre del índice, la función `getMetadata` utiliza *`documentdb_collection`* como clave de metadatos. Si desea usar un nombre de índice diferente sin el método `getMetadata`, puede usar la configuración `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
```

Puede utilizar un esquema de Amazon DocumentDB preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Si utiliza la Consola de administración de AWS para crear su canalización, también debe adjuntarla a la VPC para poder utilizar Amazon DocumentDB como fuente. Para ello, busque la sección Opciones de la red de origen, seleccione la casilla Attach to VPC (Adjuntar a la VPC) y elija el CIDR de una de las opciones predeterminadas proporcionadas. Puede usar cualquier CIDR de un espacio de direcciones privado, tal como se define en RFC 1918 Best Current Practice.

Para proporcionar un CIDR personalizado, seleccione Otros en el menú desplegable. Para evitar una colisión en las direcciones IP entre OpenSearch Ingestión y Amazon DocumentDB, asegúrese de que el CIDR de la VPC de Amazon DocumentDB sea diferente del CIDR de Ingestión. OpenSearch

Para obtener más información, consulte Configuring VPC access for a pipeline.

Coherencia de datos

La canalización garantiza la coherencia de los datos mediante el sondeo continuo o la recepción de cambios del clúster de Amazon DocumentDB y la actualización de los documentos correspondientes del OpenSearch índice.

OpenSearch La ingestión apoya el end-to-end reconocimiento para garantizar la durabilidad de los datos. Cuando una canalización lee instantáneas o transmisiones, crea particiones de forma dinámica para el procesamiento paralelo. La canalización marca una partición como completa cuando recibe un acuse de recibo después de ingerir todos los registros del dominio o la OpenSearch colección.

Si quieres incorporarlos a una colección de búsquedas OpenSearch sin servidor, puedes generar un identificador de documento en la canalización. Si quieres incorporarlo a una colección de series temporales OpenSearch sin servidor, ten en cuenta que la canalización no genera un identificador de documento, por lo que debes omitirlo `document_id: "${getMetadata(\"primary_key\")}"` en la configuración del colector de canalizaciones.

Una canalización OpenSearch de ingestión también asigna las acciones de los eventos entrantes a las correspondientes acciones de indexación masiva para facilitar la ingesta de documentos. Esto mantiene la coherencia de los datos, de modo que cada cambio de datos en Amazon DocumentDB se concilia con los cambios en los documentos correspondientes. OpenSearch

Asignación de tipos de datos

OpenSearch El servicio asigna dinámicamente los tipos de datos de cada documento entrante al tipo de datos correspondiente en Amazon DocumentDB. La siguiente tabla muestra cómo OpenSearch Service mapea automáticamente varios tipos de datos.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Entero	OpenSearch asigna automáticamente valores enteros de Amazon DocumentDB a OpenSearch enteros. OpenSearch mapea dinámicamente el campo en función del primer documento enviado. Si tiene una combinación de tipos de datos para el mismo atributo en Amazon DocumentDB, es posible que se produzca un error en la asignación automática. Por ejemplo, si el primer documento tiene un atributo largo y un documento posterior tiene el mismo atributo que un número entero, OpenSearch no podrá ingerir el segundo documento. En estos casos, debe proporcionar una plantilla de asignación explícita que elija el tipo de número más flexible, como la siguiente: `{ "template": { "mappings": { "properties": { "MixedNumberField": { "type": "float" } } } } }`	Amazon DocumentDB admite enteros.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Largo	OpenSearch asigna automáticamente los valores largos de Amazon DocumentDB a OpenSearch largos. OpenSearch mapea dinámicamente el campo en función del primer documento enviado. Si tiene una combinación de tipos de datos para el mismo atributo en Amazon DocumentDB, es posible que se produzca un error en la asignación automática. Por ejemplo, si el primer documento tiene un atributo largo y un documento posterior tiene el mismo atributo que un número entero, OpenSearch no podrá ingerir el segundo documento. En estos casos, debe proporcionar una plantilla de asignación explícita que elija el tipo de número más flexible, como la siguiente: `{ "template": { "mappings": { "properties": { "MixedNumberField": { "type": "float" } } } } }`	Amazon DocumentDB admite largos.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Cadena	OpenSearch asigna automáticamente los valores de cadena como texto. En algunas situaciones, como en el caso de los valores enumerados, puede asignarlos al tipo de palabra clave. El siguiente ejemplo muestra cómo asignar un atributo de Amazon DocumentDB denominado `PartType` a una OpenSearch palabra clave. `{ "template": { "mappings": { "properties": { "PartType": { "type": "keyword" } } } } }`	Amazon DocumentDB admite cadenas.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Double	OpenSearch asigna automáticamente los valores dobles de Amazon DocumentDB a OpenSearch dobles. OpenSearch mapea dinámicamente el campo en función del primer documento enviado. Si tiene una combinación de tipos de datos para el mismo atributo en Amazon DocumentDB, es posible que se produzca un error en la asignación automática. Por ejemplo, si el primer documento tiene un atributo largo y un documento posterior tiene el mismo atributo que un número entero, OpenSearch no podrá ingerir el segundo documento. En estos casos, debe proporcionar una plantilla de asignación explícita que elija el tipo de número más flexible, como la siguiente: `{ "template": { "mappings": { "properties": { "MixedNumberField": { "type": "float" } } } } }`	Amazon DocumentDB admite dobles.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Date	De forma predeterminada, la fecha se asigna a una entrada entera. OpenSearch Puede definir una plantilla de mapeo personalizada para asignar una fecha a una OpenSearch fecha. `{ "template": { "mappings": { "properties": { "myDateField": { "type": "date", "format": "epoch_second" } } } } }`	Amazon DocumentDB admite fechas.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Timestamp	De forma predeterminada, la marca de tiempo se asigna a una entrada entera. OpenSearch Puede definir una plantilla de mapeo personalizada para asignar una fecha a una OpenSearch fecha. `{ "template": { "mappings": { "properties": { "myTimestampField": { "type": "date", "format": "epoch_second" } } } } }`	Amazon DocumentDB admite marcas temporales.
Booleano	OpenSearch asigna un tipo booleano de Amazon DocumentDB a un tipo booleano. OpenSearch	Amazon DocumentDB admite atributos de tipo booleano.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Decimal	OpenSearch asigna los atributos del mapa de Amazon DocumentDB a campos anidados. Las mismas asignaciones se aplican dentro de un campo anidado. El siguiente ejemplo asigna una cadena de un campo anidado a un tipo de palabra clave en: OpenSearch `{ "template": { "mappings": { "properties": { "myDecimalField": { "type": "double" } } } } }` Con esta asignación personalizada, puede consultar y agregar el campo con una precisión de doble nivel. El valor original conserva toda la precisión de la `_source` propiedad del OpenSearch documento. Sin este mapeo, OpenSearch usa texto por defecto.	Amazon DocumentDB admite decimales.
Expresión regular	El tipo de expresión regular crea campos anidados. Entre ellos se encuentran *<myFieldName>* `.pattern` y *<myFieldName>* `.options`.	Amazon DocumentDB admite expresiones regulares.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Datos Binary	OpenSearch asigna automáticamente los datos binarios de Amazon DocumentDB a OpenSearch texto. Puede proporcionar un mapeo para escribirlos como campos binarios. OpenSearch El siguiente ejemplo muestra cómo asignar un campo de Amazon DocumentDB denominado `imageData` a un campo OpenSearch binario. `{ "template": { "mappings": { "properties": { "imageData": { "type": "binary" } } } } }`	Amazon DocumentDB admite campos de datos binarios.
ObjectId	Los campos con un tipo de ObjectID se asignan a OpenSearch campos de texto. El valor será la representación en cadena del objectId.	Amazon DocumentDB admite objectId.

Tipo de datos:	OpenSearch	Amazon DocumentDB
Nulo	OpenSearch puede ingerir documentos con el tipo nulo de Amazon DocumentDB. Guarda el valor como un valor nulo en el documento. No hay ninguna asignación para este tipo y este campo no está indexado ni se puede buscar en él. Si se utiliza el mismo nombre de atributo para un tipo nulo y, posteriormente, se cambia a un tipo diferente, como una cadena, OpenSearch crea un mapeo dinámico para el primer valor no nulo. Los valores subsiguientes pueden seguir siendo valores nulos de Amazon DocumentDB.	Amazon DocumentDB admite campos de tipo nulo.
Sin definir	OpenSearch puede ingerir documentos con el tipo indefinido de Amazon DocumentDB. Guarda el valor como un valor nulo en el documento. No hay ninguna asignación para este tipo y este campo no está indexado ni se puede buscar en él. Si se utiliza el mismo nombre de campo para un tipo indefinido y, posteriormente, se cambia a un tipo diferente, como una cadena, OpenSearch crea un mapeo dinámico para el primer valor no indefinido. Los valores subsiguientes pueden seguir siendo valores indefinidos de Amazon DocumentDB.	Amazon DocumentDB admite campos de tipo indefinido.

Tipo de datos:	OpenSearch	Amazon DocumentDB
MinKey	OpenSearch puede ingerir documentos con el tipo MinKey de Amazon DocumentDB. Guarda el valor como un valor nulo en el documento. No hay ninguna asignación para este tipo y este campo no está indexado ni se puede buscar en él. Si se utiliza el mismo nombre de campo para un tipo MinKey y, posteriormente, se cambia a un tipo diferente, como una cadena, OpenSearch crea un mapeo dinámico para el primer valor que no sea Minkey. Los valores subsiguientes pueden seguir siendo valores minKey de Amazon DocumentDB.	Amazon DocumentDB admite campos de tipo minKey.
MaxKey	OpenSearch puede ingerir documentos con el tipo MaxKey de Amazon DocumentDB. Guarda el valor como un valor nulo en el documento. No hay ninguna asignación para este tipo y este campo no está indexado ni se puede buscar en él. Si se utiliza el mismo nombre de campo para un tipo MaxKey y, posteriormente, cambia a un tipo diferente, como una cadena, OpenSearch crea un mapeo dinámico para el primer valor que no sea MaxKey. Los valores subsiguientes pueden seguir siendo valores maxKey de Amazon DocumentDB.	Amazon DocumentDB admite campos de tipo maxKey.

Le recomendamos que configure la cola de mensajes sin salida (DLQ) en su canalización de ingestión. OpenSearch Si has configurado la cola, el OpenSearch servicio envía a la cola todos los documentos fallidos que no se puedan ingerir debido a errores de mapeo dinámico.

En caso de que las asignaciones automáticos fallen, puede usar `template_type` y `template_content` en su configuración de canalización para definir reglas de asignación explícitas. Como alternativa, puede crear plantillas de asignación directamente en su dominio o colección de búsqueda antes de iniciar la canalización.

Limitaciones

Tenga en cuenta las siguientes limitaciones al configurar una canalización de OpenSearch ingestión para Amazon DocumentDB:

- La integración OpenSearch de ingestión con Amazon DocumentDB actualmente no admite la ingesta entre regiones. El clúster de Amazon DocumentDB y la canalización OpenSearch de ingestión deben estar en el mismo lugar. Región de AWS
- La integración OpenSearch de ingestión con Amazon DocumentDB actualmente no admite la ingesta entre cuentas. El clúster de Amazon DocumentDB y la canalización OpenSearch de ingestión deben estar en el mismo lugar. Cuenta de AWS
- Una canalización OpenSearch de ingestión solo admite un clúster de Amazon DocumentDB como origen.
- La integración OpenSearch de Ingestión con Amazon DocumentDB admite específicamente los clústeres basados en instancias de Amazon DocumentDB. No es compatible con los clústeres elásticos de Amazon DocumentDB.
- La integración OpenSearch de ingestión solo es compatible AWS Secrets Manager como mecanismo de autenticación para su clúster de Amazon DocumentDB.
- No puede actualizar la configuración de la canalización existente para ingerir datos de una base de datos o colección diferente. En su lugar, debe crear una canalización nueva.

Alarmas recomendadas CloudWatch

Para obtener el mejor rendimiento, le recomendamos que utilice las siguientes CloudWatch alarmas al crear una canalización de OpenSearch ingestión para acceder a un clúster de Amazon DocumentDB como fuente.

CloudWatch Alarma	Description (Descripción)
<pipeline-name> . Base de datos del documento. Credenciales cambiadas	Esta métrica indica la frecuencia con la que se rotan los secretos. AWS
<pipeline-name> .documentdb. executorRefreshErrors	Esta métrica indica errores al actualizar los secretos de AWS .
<pipeline-name> .documento/base de datos. exportRecordsTotal	Esta métrica indica el número de registros exportados desde Amazon DocumentDB.
<pipeline-name> .documento/base de datos. exportRecordsProcessed	Esta métrica indica la cantidad de registros procesados por la canalización OpenSearch de ingestión.
<pipeline-name> .documentdb. exportRecordProcessingErrores	Esta métrica indica el número de errores de procesamiento en una canalización OpenSearch de ingestión al leer los datos de un clúster de Amazon DocumentDB.
<pipeline-name> .documentdb. exportRecordsSuccessTotal	Esta métrica indica el número total de registros de exportación procesados correctamente.
<pipeline-name> .documentdb. exportRecordsFailedTotal	Esta métrica indica el número total de registros que no se han podido procesar.
<pipeline-name> . Base de datos del documento: bytes recibidos	Esta métrica indica el número total de bytes recibidos por una canalización de ingestión. OpenSearch
<pipeline-name> .DocumentDB. Bytes procesados	Esta métrica indica el número total de bytes procesados por una canalización de ingestión. OpenSearch
<pipeline-name> .documentdb. exportPartitionQueryTotal	Esta métrica indica el total de la partición exportada.
<pipeline-name> .documentdb. streamRecordsSuccessTotal	Esta métrica indica el número de registros del flujo procesados correctamente.

CloudWatch Alarma	Description (Descripción)
<pipeline-name> .documentdb. streamRecordsFailedTotal	Esta métrica indica el número total de registros del flujo que no se han podido procesar.

Uso de una canalización OpenSearch de ingestión con Confluent Cloud Kafka

Puedes usar una canalización de OpenSearch ingestión para transmitir datos desde los clústeres de Kafka de Confluent Cloud a los dominios de Amazon OpenSearch Service y OpenSearch a las colecciones sin servidor. OpenSearch Ingestion admite configuraciones de redes públicas y privadas para la transmisión de datos desde los clústeres Kafka de Confluent Cloud a dominios o colecciones gestionados por Service o Serverless. OpenSearch OpenSearch

Conectividad con clústeres de Confluent Cloud Kafka públicos

Puede usar las canalizaciones de OpenSearch ingestión para migrar datos desde un clúster Kafka de Confluent Cloud con una configuración pública, lo que significa que el nombre DNS del dominio se puede resolver públicamente. Para ello, configure una canalización de OpenSearch ingestión con el clúster público de Kafka de Confluent Cloud como origen y OpenSearch Service o Serverless como destino. OpenSearch Esto procesa los datos de streaming desde un clúster de origen autogestionado hasta un dominio o colección de destino AWS gestionado.

Requisitos previos

Antes de crear tu canalización OpenSearch de ingestión, lleva a cabo los siguientes pasos:

1. Cree un clúster de clústeres de Confluent Cloud Kafka que actúe como origen. El clúster debe contener los datos que desea incorporar a Service. OpenSearch
2. Cree un dominio OpenSearch de servicio o una colección OpenSearch sin servidor a la que desee migrar los datos. Para obtener más información, consulte the section called “Creación de dominios OpenSearch de servicio” y the section called “Creación de colecciones”.
3. Configure la autenticación en el clúster de Confluent Cloud Kafka con AWS Secrets Manager. Siga los pasos de Rotate AWS Secrets Manager secrets para activar la rotación de secretos.
4. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster autogestionado en su dominio o colección.

El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización, que creará en el paso siguiente, escriba datos en un dominio. Asegúrese de actualizar el `resource` con su propio 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"
      ]
    }
  ]
}
```

Para crear un rol de IAM con los permisos correctos para acceder a los datos de escritura de la colección o dominio, consulte the section called “Configuración de roles y usuarios”.

Paso 1: configurar el rol de canalización

Una vez establecidos los requisitos previos de la canalización de clústeres Kafka de Confluent Cloud, configure el rol de canalización que quiere usar en la configuración de la canalización y añada permiso para escribir en un dominio de OpenSearch servicio o en una colección OpenSearch sin servidor, así como permiso para leer los secretos de Secrets Manager.

Se necesita el siguiente permiso para administrar la interfaz de red:

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" }
            }
        }
    ]
}
```

El siguiente es el permiso necesario para leer los secretos del servicio: AWS Secrets Manager

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "SecretsManagerReadAccess",
            "Effect": "Allow",
            "Action": ["secretsmanager:GetSecretValue"],
            "Resource": ["arn:aws:secretsmanager:us-
east-1:111122223333:secret:,secret-name"]
        }
    ]
}
```

Se necesitan los siguientes permisos para escribir en un dominio de Amazon OpenSearch Service:

```
{
  "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/*"
    }
  ]
}
```

Paso 2: crear la canalización

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, que especifica su Confluent Cloud Kafka como fuente.

Puede especificar varios dominios de OpenSearch servicio como destinos de sus datos. Esta capacidad permite el enrutamiento condicional o la replicación de los datos entrantes en varios dominios OpenSearch de servicio.

También puede migrar datos de un clúster de Confluent Kafka de origen a una colección de VPC OpenSearch sin servidor. Asegúrese de proporcionar una política de acceso a la red en la

configuración de la canalización. Puede utilizar un registro de esquemas de Confluent para definir un esquema de 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"
```

Puede utilizar un esquema preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Conectividad con clústeres de Confluent Cloud Kafka en una VPC

También puede usar las canalizaciones OpenSearch de ingestión para migrar datos desde un clúster Kafka de Confluent Cloud que se ejecute en una VPC. Para ello, configure una canalización de OpenSearch ingestión con un clúster Kafka de Confluent Cloud como origen y Service o Serverless como destino. OpenSearch OpenSearch Esto procesa los datos de streaming desde un clúster fuente de Confluent Cloud Kafka a un dominio o colección de destino administrado de AWS.

OpenSearch Ingestion es compatible con los clústeres Kafka de Confluent Cloud configurados en todos los modos de red compatibles en Confluent. Los siguientes modos de configuración de red son compatibles como fuente en Ingestion: OpenSearch

- AWS Emparejamiento de VPC
- AWS PrivateLink para clústeres dedicados
- AWS PrivateLink para clústeres empresariales
- AWS Transit Gateway

Requisitos previos

Antes de crear la canalización OpenSearch de ingestión, lleve a cabo los siguientes pasos:

1. Cree un clúster Kafka de Confluent Cloud con una configuración de red de VPC que contenga los datos que desea incorporar al servicio. OpenSearch
2. Cree un dominio OpenSearch de servicio o una colección OpenSearch sin servidor a la que desee migrar los datos. Para obtener más información, consulte the section called “Creación de dominios OpenSearch de servicio” ythe section called “Creación de colecciones”.
3. Configure la autenticación en el clúster de Confluent Cloud Kafka con AWS Secrets Manager. Siga los pasos de Rotate AWS Secrets Manager secrets para activar la rotación de secretos.
4. Obtenga el ID de la VPC que tiene acceso al clúster de Kafka de Confluent Cloud. Elija el CIDR de VPC que utilizará Ingestion. OpenSearch

> **Note**
>
> Si utilizas el Consola de administración de AWS para crear tu canalización, también debes adjuntar tu canalización de OpenSearch ingestión a tu VPC para poder usar el clúster Kafka de Confluent Cloud. Para ello, busque la sección Configuración de la red, seleccione la casilla Adjuntar a la VPC y elija el CIDR de una de las opciones

predeterminadas proporcionadas o seleccione una opción propia. Puede usar cualquier CIDR de un espacio de direcciones privado, tal como se define en RFC 1918 Best Current Practice.
Para proporcionar un CIDR personalizado, seleccione Otros en el menú desplegable.
Para evitar una colisión en las direcciones IP entre la OpenSearch ingesta y la autogestión OpenSearch, asegúrese de que el CIDR de la OpenSearch VPC autogestionada sea diferente del CIDR de la ingestión. OpenSearch

5. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster autogestionado en su dominio o colección.

> **Note**
>
> Si utiliza AWS PrivateLink para conectar Confluent Cloud Kafka, tendrá que configurar Opciones de DHCP de VPC. Nombres de host de DNS y Resolución de DNS deben estar habilitados.
> En concreto, utilice los siguientes valores de conjunto de opciones:
> Clústeres empresariales:
>
> ```
> domain-name: aws.private.confluent.cloud
> domain-name-servers: AmazonProvidedDNS
> ```
>
> Clústeres dedicados:
>
> ```
> domain-name: aws.confluent.cloud
> domain-name-servers: AmazonProvidedDNS
> ```
>
> Este cambio garantiza que la resolución de DNS para el PrivateLink punto final de Confluent funcione correctamente en la VPC.

El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización, que creará en el paso siguiente, escriba datos en un dominio. Asegúrese de actualizar el `resource` con su propio 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"
      ]
    }
  ]
}
```

Para crear un rol de IAM con los permisos correctos para acceder a los datos de escritura de la colección o dominio, consulte the section called “Configuración de roles y usuarios”.

Paso 1: configurar el rol de canalización

Una vez configurados los requisitos previos de la canalización, configure el rol de la canalización que desee usar en la configuración de la canalización y agregue los siguientes permisos al rol:

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"
                    }
            }
```

```
        }
    ]
}
```

Debe proporcionar los permisos de Amazon EC2 anteriores en la función de IAM que utilice para crear la canalización de OpenSearch ingestión, ya que la canalización utiliza estos permisos para crear y eliminar una interfaz de red en la VPC. La canalización solo puede acceder al clúster de Kafka a través de esta interfaz de red.

Paso 2: crear la canalización

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, en la que se especifica a Kafka como fuente.

Puede especificar varios dominios OpenSearch de servicio como destinos de sus datos. Esta capacidad permite el enrutamiento condicional o la replicación de los datos entrantes en varios dominios OpenSearch de servicio.

También puede migrar datos de un clúster de Confluent Kafka de origen a una colección de VPC OpenSearch sin servidor. Asegúrese de proporcionar una política de acceso a la red en la configuración de la canalización. Puede utilizar un registro de esquemas de Confluent para definir un esquema de 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"
```

Uso de una canalización OpenSearch de ingestión con Amazon Managed Streaming for Apache Kafka

Puede usar el complemento de Kafka para incorporar datos de Amazon Managed Streaming for Apache Kafka (Amazon MSK) OpenSearch a su canal de ingestión. Con Amazon MSK, puede crear y ejecutar aplicaciones que utilizan Apache Kafka para procesar datos de streaming. OpenSearch Ingestion utiliza AWS PrivateLink para conectarse a Amazon MSK. Puede ingerir datos de clústeres de Amazon MSK y Amazon MSK sin servidor. La única diferencia entre los dos procesos son los pasos de requisitos previos que debe seguir antes de configurar su canalización.

Temas

- Requisitos previos de Amazon MSK aprovisionado
- Requisitos previos de Amazon MSK sin servidor
- Paso 1: Configurar un rol de canalización
- Paso 2: crear la canalización
- Paso 3: (opcional) Utilice el registro de esquemas AWS Glue
- Paso 4: (opcional) Configurar las unidades de cómputo recomendadas (OCUs) para la canalización de Amazon MSK

Requisitos previos de Amazon MSK aprovisionado

Antes de crear su canalización OpenSearch de ingestión, lleve a cabo los siguientes pasos:

1. Siga los pasos que se indican en Creating a cluster en la Guía para desarrolladores de Amazon Managed Streaming para Apache Kafka para crear un clúster aprovisionado de Amazon MSK. Para el tipo de agente, elige cualquier opción excepto los `t3` tipos, ya que OpenSearch Ingestion no los admite.
2. Cuando el clúster tenga el estado Activo, siga los pasos que se indican en Activar la conectividad de varias VPC.
3. Siga los pasos que se indican en Adjuntar una política de clúster al clúster de MSK para adjuntar una de las siguientes políticas, en función de si el clúster y la canalización están en la misma Cuenta de AWS. Esta política permite a OpenSearch Ingestion crear una AWS PrivateLink conexión con su clúster de Amazon MSK y leer datos de los temas de Kafka. Asegúrese de actualizar el `resource` con su propio ARN.

   Las siguientes políticas se aplican cuando el clúster y la canalización están en la misma Cuenta de 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"
    }
  ]
}
```

Si tu clúster de Amazon MSK está en un proceso Cuenta de AWS diferente al tuyo, adjunta la siguiente política en su lugar. Tenga en cuenta que el acceso entre cuentas solo es posible con los clústeres de Amazon MSK aprovisionados y no con los clústeres de Amazon MSK sin servidor. El ARN de AWS `principal` debe ser el ARN de la misma función de canalización que proporcionaste a la configuración de la canalización:

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. Cree un tema de Kafka siguiendo los pasos de Creación de un tema. Asegúrese de que *BootstrapServerString* sea uno de los bootstrap del punto final privado (VPC única). URLs El valor de `--replication-factor` debe ser 2 o 3, en función del número de zonas que tenga su clúster de Amazon MSK. El valor de las `--partitions` debe ser como mínimo `10`.
5. Produzca y consuma datos siguiendo los pasos de Producir y consumir datos. De nuevo, asegúrate de que *BootstrapServerString* sea uno de los bootstrap de tu punto final privado (VPC única). URLs

Requisitos previos de Amazon MSK sin servidor

Antes de crear tu canalización de OpenSearch ingestión, lleva a cabo los siguientes pasos:

1. Siga los pasos que se indican en Create an MSK Serverless cluster en la Guía para desarrolladores de Amazon Managed Streaming para Apache Kafka para crear un clúster de Amazon MSK sin servidor.
2. Cuando el clúster tenga el estado Activo, siga los pasos que se indican en Attach a cluster policy to the MSK cluster para adjuntar la siguiente política. Asegúrese de actualizar el `resource` con su propio 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"
    }
  ]
}
```

Esta política permite a OpenSearch Ingestion crear una AWS PrivateLink conexión con su clúster de Amazon MSK Serverless y leer datos de los temas de Kafka. Esta política se aplica cuando el clúster y la canalización se encuentran en el mismo lugar Cuenta de AWS, lo que debe ser cierto, ya que Amazon MSK Serverless no admite el acceso entre cuentas.

3. Cree un tema de Kafka siguiendo los pasos de Creación de un tema. Asegúrese de que *BootstrapServerString* sea uno de sus programas de IAM de nivel simple de autenticación y seguridad (SASL). URLs El valor de `--replication-factor` debe ser 2 o 3, en función

del número de zonas que tenga su clúster de Amazon MSK sin servidor. El valor de las `--partitions` debe ser como mínimo `10`.

4. Produzca y consuma datos siguiendo los pasos de Producir y consumir datos. Una vez más, asegúrate de que *`BootstrapServerString`* sea uno de los bootstrap de IAM de capa simple de autenticación y seguridad (SASL). URLs

Paso 1: Configurar un rol de canalización

Una vez configurado el clúster aprovisionado o sin servidor de Amazon MSK, agregue los siguientes permisos de Kafka al rol de canalización que desee usar en la configuración de la canalización:

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/*"
            ]
        }
    ]
}
```

Paso 2: crear la canalización

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, en la que se especifica a Kafka como fuente:

```
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
```

Puede utilizar un esquema de Amazon MSK preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Paso 3: (opcional) Utilice el registro de esquemas AWS Glue

Al utilizar OpenSearch Ingestion con Amazon MSK, puede utilizar el formato de datos AVRO para los esquemas alojados en el registro de esquemas. AWS Glue Con AWS Glue Schema Registry, puede descubrir, controlar y evolucionar de forma centralizada los esquemas de flujo de datos.

Para usar esta opción, habilite el `type` de esquema en la configuración de su canalización:

```
schema:
  type: "aws_glue"
```

También debe proporcionar AWS Glue permisos de acceso de lectura en su rol de canalización. Puedes usar la política AWS gestionada llamada AWSGlueSchemaRegistryReadonlyAccess. Además, tu registro debe estar en la misma región Cuenta de AWS y en la misma región que tu canal OpenSearch de ingestión.

Paso 4: (opcional) Configurar las unidades de cómputo recomendadas (OCUs) para la canalización de Amazon MSK

Cada unidad de cómputo tiene un consumidor por tema. Los agentes equilibran las particiones entre estos consumidores para un tema determinado. Sin embargo, cuando el número de particiones es superior al número de consumidores, Amazon MSK aloja varias particiones en cada consumidor. OpenSearch Ingestion tiene un escalado automático incorporado para escalar hacia arriba o hacia abajo según el uso de la CPU o la cantidad de registros pendientes en proceso.

Para un rendimiento óptimo, distribuya las particiones entre varias unidades de cómputo para el procesamiento paralelo. Si los temas tienen un gran número de particiones (por ejemplo, más de 96, que es el máximo OCUs por canalización), se recomienda configurar una canalización con 1 a OCUs 96. Esto se debe a que se escalará automáticamente según sea necesario. Si un tema tiene un número reducido de particiones (por ejemplo, menos de 96), mantenga la unidad de cómputo máxima igual a la cantidad de particiones.

Cuando una canalización tenga más de un tema, seleccione el tema con el mayor número de particiones como referencia para configurar las unidades de cómputo máximas. Si añades otra

canalización con un nuevo conjunto de OCUs temas y grupos de consumidores, puedes escalar el rendimiento de forma casi lineal.

Uso de una canalización OpenSearch de ingestión con Amazon RDS

Puede utilizar una canalización de OpenSearch ingestión con Amazon RDS para exportar los datos existentes y transmitir los cambios (como crear, actualizar y eliminar) a los dominios y colecciones de Amazon OpenSearch Service. La canalización de OpenSearch ingestión incorpora una infraestructura de captura de datos de cambios (CDC) para proporcionar una forma de alta escala y baja latencia de transmitir datos de forma continua desde Amazon RDS. Se admiten RDS para MySQL y RDS para PostgreSQL.

Existen dos formas de utilizar Amazon RDS como origen para procesar datos: con o sin una instantánea inicial completa. Una instantánea inicial completa es una instantánea de las tablas especificadas y esta instantánea se exporta a Amazon S3. A partir de ahí, una canalización de OpenSearch ingestión los envía a un índice de un dominio o los divide en varios índices de un dominio. Para mantener la OpenSearch coherencia de los datos en Amazon RDS, la canalización sincroniza todos los eventos de creación, actualización y eliminación de las tablas de las instancias de Amazon RDS con los documentos guardados en el OpenSearch índice o los índices.

Cuando utiliza una instantánea inicial completa, su canalización de OpenSearch ingestión primero ingiere la instantánea y, a continuación, comienza a leer los datos de las transmisiones de cambios de Amazon RDS. Con el tiempo, recupera y mantiene la coherencia de los datos casi en tiempo real entre Amazon RDS y OpenSearch.

También puede utilizar la integración de OpenSearch Ingestion con Amazon RDS para realizar un seguimiento de los cambios, capturar datos e ingerir todas las actualizaciones de Aurora. OpenSearch Elija esta opción si ya tiene una instantánea completa de algún otro mecanismo o si solo quiere capturar todos los cambios en los datos de una instancia de Amazon RDS.

Si elige esta opción, debe configurar el registro binario de Amazon RDS para MySQL o configurar la replicación lógica para la instancia de base de datos Amazon RDS para PostgresSQL.

Temas

- RDS para MySQL
- RDS para PostgreSQL

RDS para MySQL

Complete los siguientes pasos para configurar una canalización OpenSearch de ingestión con Amazon RDS para RDS for MySQL.

Temas

- Requisitos previos de RDS para MySQL
- Paso 1: configurar el rol de canalización
- Paso 2: crear la canalización
- Coherencia de datos
- Asignación de tipos de datos
- Limitaciones
- Alarmas recomendadas CloudWatch

Requisitos previos de RDS para MySQL

Antes de crear su canalización de OpenSearch ingestión, lleve a cabo los siguientes pasos:

1. Cree un grupo de parámetros de base de datos personalizado en Amazon RDS para configurar el registro binario y establecer los siguientes parámetros.

   ```
   binlog_format=ROW
   binlog_row_image=full
   binlog_row_metadata=FULL
   ```

   Compruebe también que el parámetro `binlog_row_value_options` no esté establecido en `PARTIAL_JSON`.

   Para obtener más información, consulte Configuración del registro binario de RDS para MySQL.
2. Seleccione o cree una instancia de base de datos RDS para MySQL y asocie el grupo de parámetros creado en el paso anterior a la instancia de base de datos.
3. Compruebe que las copias de seguridad automáticas están activadas en la base de datos. Para obtener más información, consulte Habilitación de copias de seguridad automatizadas.
4. Configure la retención de registros binarios con tiempo suficiente para que se produzca la replicación, por ejemplo, 24 horas. Para obtener más información, consulte Configuración y visualización de la configuración de registros binarios en la Guía del usuario de Amazon RDS.

5. Configure la autenticación de nombre de usuario y contraseña en su instancia de Amazon RDS mediante la administración de contraseñas con Amazon RDS y AWS Secrets Manager. También puedes crear una username/password combinación creando un secreto de Secrets Manager.
6. Si utiliza la función de instantáneas iniciales completa, cree un rol de IAM AWS KMS key y un rol de IAM para exportar datos de Amazon RDS a Amazon S3.

   El rol de IAM debe tener la siguiente política de permisos:

   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/*"
            ]
        }
    ]
}
```

   El rol también debe tener las siguientes relaciones de confianza:

   JSON

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

```
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

7. Seleccione o cree un dominio de OpenSearch servicio o una colección OpenSearch sin servidor. Para obtener más información, consulte Creación de dominios OpenSearch de servicio y Creación de colecciones.
8. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su instancia de base de datos de Amazon RDS en su dominio o colección.

Paso 1: configurar el rol de canalización

Una vez configurados los requisitos previos de la canalización de Amazon RDS, configure el rol de la canalización que se usará en la configuración de la canalización. Agregue también los siguientes permisos para el origen de Amazon RDS al rol:

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:*"
    ]
    }
    ]
    }
```

Paso 2: crear la canalización

Configure una canalización OpenSearch de ingestión similar a la siguiente. La canalización de ejemplo especifica una instancia de Amazon RDS como origen.

```
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
```

Puede utilizar un esquema de Amazon RDS preconfigurado para crear esta canalización. Para obtener más información, consulte Uso de esquemas.

A fin de utilizar Amazon Aurora como origen, debe configurar el acceso a la VPC para la canalización. La VPC que elija debe ser la misma que usa su origen de Amazon Aurora. A continuación, elija una o varias subredes y uno o varios grupos de seguridad de VPC. Tenga en cuenta que la canalización necesita acceso de red a una base de datos Aurora MySQL, por lo que también debe comprobar que el clúster de Aurora esté configurado con un grupo de seguridad de VPC que permita el tráfico entrante desde el grupo de seguridad de VPC de la canalización al puerto de la base de datos. Para obtener más información, consulte Control de acceso con grupos de seguridad.

Si utiliza la Consola de administración de AWS para crear su canalización, también debe adjuntarla a la VPC para poder utilizar Amazon Aurora como fuente. Para ello, busque la sección Configuración de la red, seleccione Adjuntar a la VPC y elija el CIDR de una de las opciones predeterminadas proporcionadas o seleccione una opción propia. Puede usar cualquier CIDR de un espacio de direcciones privado, tal como se define en RFC 1918 Best Current Practice.

Para proporcionar un CIDR personalizado, seleccione Otros en el menú desplegable. Para evitar una colisión de direcciones IP entre OpenSearch Ingestión y Amazon RDS, asegúrese de que el CIDR de la VPC de Amazon RDS sea diferente del CIDR de Ingestión. OpenSearch

Para obtener más información, consulte Configuring VPC access for a pipeline.

Coherencia de datos

La canalización garantiza la coherencia de los datos al sondear o recibir cambios de forma continua desde la instancia de Amazon RDS y actualizar los documentos correspondientes del OpenSearch índice.

OpenSearch La ingestión apoya el end-to-end reconocimiento para garantizar la durabilidad de los datos. Cuando una canalización lee instantáneas o transmisiones, crea particiones de forma dinámica para el procesamiento paralelo. La canalización marca una partición como completa cuando recibe un acuse de recibo después de ingerir todos los registros del dominio o la OpenSearch colección. Si quieres incorporarlos a una colección de búsquedas OpenSearch sin servidor, puedes generar un identificador de documento en la canalización. Si quieres incorporarlo a una colección de series temporales OpenSearch sin servidor, ten en cuenta que la canalización no genera un identificador de documento, por lo que debes omitirlo `document_id: "${getMetadata(\"primary_key\")}"` en la configuración del colector de canalizaciones.

Una canalización OpenSearch de ingestión también asigna las acciones de los eventos entrantes a las correspondientes acciones de indexación masiva para facilitar la ingesta de documentos.

Esto mantiene la coherencia de los datos, de modo que cada cambio de datos en Amazon RDS se concilia con los cambios en los documentos correspondientes. OpenSearch

Asignación de tipos de datos

OpenSearch La canalización de ingestión asigna los tipos de datos de MySQL a representaciones que son adecuadas para el consumo de los dominios o colecciones de OpenSearch servicios. Si no hay ninguna plantilla de mapeo definida OpenSearch, determina OpenSearch automáticamente los tipos de campo con un mapeo dinámico en función del primer documento enviado. También puede definir de forma explícita los tipos de campo que mejor se adapten a sus necesidades OpenSearch mediante una plantilla de mapeo.

La siguiente tabla muestra los tipos de datos de MySQL y los tipos de OpenSearch campo correspondientes. La columna Tipo de OpenSearch campo predeterminado muestra el tipo de campo correspondiente OpenSearch si no se ha definido un mapeo explícito. En este caso, determina OpenSearch automáticamente los tipos de campo con un mapeo dinámico. La columna Tipo de OpenSearch campo recomendado es el tipo de campo correspondiente que se recomienda especificar de forma explícita en una plantilla de mapeo. Estos tipos de campos están más alineados con los tipos de datos de MySQL y, por lo general, pueden habilitar mejores funciones de búsqueda disponibles en OpenSearch.

Tipo de datos de MySQL	Tipo OpenSearch de campo predeterminado	Tipo OpenSearch de campo recomendado
BIGINT	long	long
BIGINT UNSIGNED	long	largo sin signo
BIT	long	byte, corto, entero o largo según el número de bits
DECIMAL	text	doble o palabra
DOUBLE	float	double
FLOAT	float	float
INT	long	entero

Tipo de datos de MySQL	Tipo OpenSearch de campo predeterminado	Tipo OpenSearch de campo recomendado
INT UNSIGNED	long	long
MEDIUMIN	long	entero
MEDIUMIN UNSIGNED	long	entero
NUMERIC	text	doble o palabra
SMALLINT	long	short
SMALLINT UNSIGNED	long	entero
TINYINT	long	byte
TINYINT UNSIGNED	long	short
BINARIO	text	binario
BLOB	text	binario
CHAR	text	text
ENUM	text	palabra clave
LONGBLOE	text	binario
LONGTEXT	text	text
MEDIUMBL B	text	binario
MEDIUMTE T	text	text

Tipo de datos de MySQL	Tipo OpenSearch de campo predeterminado	Tipo OpenSearch de campo recomendado
SET	text	palabra clave
TEXT	text	text
TINYBLOB	text	binario
TINYTEXT	text	text
VARBINAR	text	binario
VARCHAR	text	text
DATE	largo (en milisegundos de época)	date
DATETIME	largo (en milisegundos de época)	date
TIME	largo (en milisegundos de época)	date
TIMESTAM	largo (en milisegundos de época)	date
YEAR	largo (en milisegundos de época)	date
GEOMETR	texto (en formato WKT)	geo_shape
GEOMETR OLLECTIO	texto (en formato WKT)	geo_shape
LINESTRIN G	texto (en formato WKT)	geo_shape
MULTILINE STRING	texto (en formato WKT)	geo_shape
MULTIPOIN T	texto (en formato WKT)	geo_shape

Tipo de datos de MySQL	Tipo OpenSearch de campo predeterminado	Tipo OpenSearch de campo recomendado
MULTIPOLYGON	texto (en formato WKT)	geo_shape
POINT	texto (en formato WKT)	geo_point o geo_shape
POLYGON	texto (en formato WKT)	geo_shape
JSON	text	objeto

Te recomendamos que configures la cola de mensajes sin salida (DLQ) en tu canal de ingestión. OpenSearch Si has configurado la cola, el OpenSearch servicio envía a la cola todos los documentos fallidos que no se puedan ingerir debido a errores de mapeo dinámico.

Si las asignaciones automáticas fallan, puede usar `template_type` y `template_content` en su configuración de canalización para definir reglas de asignación explícitas. Como alternativa, puede crear plantillas de asignación directamente en su dominio o colección de búsqueda antes de iniciar la canalización.

Limitaciones

Tenga en cuenta las siguientes limitaciones al configurar una canalización de OpenSearch ingestión para RDS for MySQL:

- La integración solo admite una base de datos MySQL por canalización.
- Actualmente, la integración no admite la ingesta de datos entre regiones; la instancia y el OpenSearch dominio de Amazon RDS deben estar en el mismo lugar. Región de AWS
- Actualmente, la integración no admite la ingesta de datos entre cuentas; la instancia de Amazon RDS y la canalización de OpenSearch ingestión deben estar en la misma ubicación. Cuenta de AWS
- Asegúrese de que la instancia de Amazon RDS tenga habilitada la autenticación mediante Secrets Manager, que es el único mecanismo de autenticación admitido.
- La configuración de canalización existente no se puede actualizar para ingerir datos de una base de datos diferente o de una tabla diferente. and/or Para actualizar la base de datos o el nombre de la tabla de una canalización, debe crear una canalización nueva.

- Por lo general, las instrucciones en lenguaje de definición de datos (DDL) no son compatibles. La coherencia de datos no se mantendrá si:
  - Se cambian las claves principales (add/delete/rename).
  - Las tablas se eliminan o se truncan.
  - Se cambian los nombres de las columnas o los tipos de datos.
- Si las tablas MySQL que se van a sincronizar no tienen definidas las claves principales, no se garantiza la coherencia de datos. Deberá definir correctamente la `document_id` opción personalizada en la configuración del OpenSearch sumidero para poder updates/deletes sincronizarla OpenSearch.
- No se admiten las referencias a claves externas con acciones de eliminación en cascada y pueden provocar inconsistencias en los datos entre RDS for MySQL y. OpenSearch
- No se admiten los clústeres de bases de datos de zonas de disponibilidad múltiple de Amazon RDS.
- Versiones compatibles: MySQL 8.0 y versiones posteriores.

Alarmas recomendadas CloudWatch

Se recomiendan las siguientes CloudWatch métricas para supervisar el rendimiento del proceso de ingestión. Estas métricas pueden ayudarlo a identificar la cantidad de datos procesados a partir de exportaciones y flujos, los errores al procesar eventos de transmisión y exportación y el número de documentos escritos en el destino. Puedes configurar CloudWatch alarmas para realizar una acción cuando una de estas métricas supere un valor específico durante un período de tiempo específico.

Métrica	Description (Descripción)
pipeline-name .RDS. Se han modificado las credenciales	Esta métrica indica la frecuencia con la que se rotan los secretos. AWS
pipeline-name .rds. executorRefreshErrors	Esta métrica indica errores al actualizar AWS los secretos.

Métrica	Description (Descripción)
pipeline-name .rds. exportRecordsTotal	Esta métrica indica el número de registros exportados desde Amazon Aurora.
pipeline-name .rds. exportRecordsProcessed	Esta métrica indica la cantidad de registros procesados por la canalización OpenSearch de ingestión.
pipeline-name .rds. exportRecordProcessingErrores	Esta métrica indica el número de errores de procesamiento en una canalización de OpenSearch ingestión al leer los datos de un clúster de Amazon Aurora.
pipeline-name .rds. exportRecordsSuccessTotal	Esta métrica indica el número total de registros de exportación procesados correctamente.
pipeline-name .rds. exportRecordsFailedTotal	Esta métrica indica el número total de registros que no se han podido procesar.
pipeline-name .RDS. Bytes recibidos	Esta métrica indica el número total de bytes recibidos por una canalización de ingestión. OpenSearch
pipeline-name .rds.bytes procesados	Esta métrica indica el número total de bytes procesados por una canalización de ingestión. OpenSearch

Métrica	Description (Descripción)
pipeline-name .rds.streamRecordsSuccessTotal	Esta métrica indica el número de registros del flujo procesados correctamente.
pipeline-name .rds.streamRecordsFailedTotal	Esta métrica indica el número total de registros del flujo que no se han podido procesar.

RDS para PostgreSQL

Complete los siguientes pasos para configurar una canalización de OpenSearch ingestión con Amazon RDS para RDS for PostgreSQL.

Temas

- Requisitos previos de PostgreSQL
- Paso 1: configurar el rol de canalización
- Paso 2: crear la canalización
- Coherencia de datos
- Asignación de tipos de datos
- Limitaciones
- CloudWatch Alarmas recomendadas

Requisitos previos de PostgreSQL

Antes de crear la canalización de OpenSearch ingestión, lleve a cabo los siguientes pasos:

1. Cree un grupo de parámetros de base de datos personalizado en Amazon RDS para configurar la replicación lógica.

   ```
   rds.logical_replication=1
   ```

   Para obtener más información, consulte Replicación lógica de Amazon RDS para PostgreSQL.

2. Seleccione o cree una instancia de base de datos de RDS para PostgreSQL y asocie el grupo de parámetros creado en el paso 1 a la instancia de base de datos.
3. Configure la autenticación de nombre de usuario y contraseña en su instancia de Amazon RDS mediante la administración de contraseñas con Aurora y AWS Secrets Manager. También puedes crear una username/password combinación creando un secreto de Secrets Manager.
4. Si utiliza la función de instantáneas iniciales completa, cree un rol de IAM AWS KMS key y un rol de IAM para exportar datos de Amazon RDS a Amazon S3.

   El rol de IAM debe tener la siguiente política de permisos:

   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/*"
            ]
        }
    ]
}
```

   El rol también debe tener las siguientes relaciones de confianza:

   JSON

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

```
            "Principal": {
                "Service": "export.rds.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

5. Seleccione o cree un dominio de OpenSearch servicio o una colección OpenSearch sin servidor. Para obtener más información, consulte Creación de dominios OpenSearch de servicio y Creación de colecciones.
6. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su instancia de base de datos de Amazon RDS en su dominio o colección.

Paso 1: configurar el rol de canalización

Una vez configurados los requisitos previos de la canalización de Amazon RDS, configure el rol de la canalización que se usará en la configuración de la canalización. Agregue también los siguientes permisos para el origen de Amazon RDS al rol:

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:*"
    ]
    }
    ]
    }
```

Paso 2: crear la canalización

Configure una canalización de OpenSearch ingestión como la siguiente, que especifique una instancia de RDS para PostgreSQL como origen.

```
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**
>
> Puede utilizar un esquema de Amazon RDS preconfigurado para crear esta canalización. Para obtener más información, consulte Uso de esquemas.

A fin de utilizar Amazon Aurora como origen, debe configurar el acceso a la VPC para la canalización. La VPC que elija debe ser la misma que usa su origen de Amazon Aurora. A continuación, elija una o varias subredes y uno o varios grupos de seguridad de VPC. Tenga en cuenta que la canalización necesita acceso de red a una base de datos Aurora MySQL, por lo que también debe comprobar que el clúster de Aurora esté configurado con un grupo de seguridad de VPC que permita el tráfico entrante desde el grupo de seguridad de VPC de la canalización al puerto de la base de datos. Para obtener más información, consulte Control de acceso con grupos de seguridad.

Si utiliza la Consola de administración de AWS para crear su canalización, también debe adjuntarla a la VPC para poder utilizar Amazon Aurora como fuente. Para ello, busque la sección Configuración de la red, seleccione Adjuntar a la VPC y elija el CIDR de una de las opciones predeterminadas proporcionadas o seleccione una opción propia. Puede usar cualquier CIDR de un espacio de direcciones privado, tal como se define en RFC 1918 Best Current Practice.

Para proporcionar un CIDR personalizado, seleccione Otros en el menú desplegable. Para evitar una colisión de direcciones IP entre OpenSearch Ingestión y Amazon RDS, asegúrese de que el CIDR de VPC de Amazon Aurora sea diferente del CIDR de Ingestión. OpenSearch

Para obtener más información, consulte Configuring VPC access for a pipeline.

Coherencia de datos

La canalización garantiza la coherencia de los datos al sondear o recibir cambios de forma continua desde la instancia de Amazon RDS y actualizar los documentos correspondientes del OpenSearch índice.

OpenSearch La ingestión apoya el end-to-end reconocimiento para garantizar la durabilidad de los datos. Cuando una canalización lee instantáneas o transmisiones, crea particiones de forma dinámica para el procesamiento paralelo. La canalización marca una partición como completa cuando recibe un acuse de recibo después de ingerir todos los registros del dominio o la OpenSearch colección. Si quieres incorporarlos a una colección de búsquedas OpenSearch sin servidor, puedes generar un identificador de documento en la canalización. Si quieres

incorporarlo a una colección de series temporales OpenSearch sin servidor, ten en cuenta que la canalización no genera un identificador de documento, por lo que debes omitirlo `document_id: "${getMetadata(\"primary_key\")}"` en la configuración del colector de canalizaciones.

Una canalización OpenSearch de ingestión también asigna las acciones de los eventos entrantes a las correspondientes acciones de indexación masiva para facilitar la ingesta de documentos. Esto mantiene la coherencia de los datos, de modo que cada cambio de datos en Amazon RDS se concilia con los cambios en los documentos correspondientes. OpenSearch

Asignación de tipos de datos

OpenSearch La canalización de ingestión asigna los tipos de datos de PostgreSQL a representaciones que son adecuadas OpenSearch para el consumo de los dominios o colecciones de servicios. Si no hay ninguna plantilla de mapeo definida OpenSearch, determine OpenSearch automáticamente los tipos de campo con un mapeo dinámico basado en el primer documento enviado. También puede definir de forma explícita los tipos de campo que mejor se adapten a sus necesidades OpenSearch mediante una plantilla de mapeo.

En la siguiente tabla se enumeran los tipos de datos de RDS para PostgreSQL y los tipos de campo correspondientes. OpenSearch La columna Tipo de OpenSearch campo predeterminado muestra el tipo de campo correspondiente OpenSearch si no se ha definido un mapeo explícito. En este caso, determina OpenSearch automáticamente los tipos de campo con un mapeo dinámico. La columna Tipo de OpenSearch campo recomendado es el tipo de campo que se recomienda especificar de forma explícita en una plantilla de mapeo. Estos tipos de campos están más alineados con los tipos de datos de RDS para PostgreSQL y, por lo general, pueden habilitar mejores funciones de búsqueda disponibles en. OpenSearch

Tipo de datos de RDS para PostgreSQL	Tipo de campo predeterminado OpenSearch	Tipo OpenSearch de campo recomendado
smallint	long	short
entero	long	entero
bigint	long	long

Tipo de datos de RDS para PostgreSQL	Tipo de campo predeterminado OpenSearch	Tipo OpenSearch de campo recomendado
decimal	text	doble o palabra
numeric[(p, s)]	text	doble o palabra
real	float	float
double precision	float	double
smallserial	long	short
serial	long	entero
bigserial	long	long
money	objeto	objeto
character varying(n)	text	text
varchar(n)	text	text
character (n)	text	text
char(n)	text	text
bpchar (n)	text	text
BPCHAR	text	text
text	text	text

Tipo de datos de RDS para PostgreSQL	Tipo de campo predeterminado OpenSearch	Tipo OpenSearch de campo recomendado
enum	text	text
bytea	text	binario
timestamp [(p)] [sin zona horaria]	largo (en milisegundos de época)	date
timestamp [(p)] con zona horaria	largo (en milisegundos de época)	date
date	largo (en milisegundos de época)	date
time [(p)] [sin zona horaria]	largo (en milisegundos de época)	date
time [(p)] con zona horaria	largo (en milisegundos de época)	date
intervalo [campos] [(p)]	texto (ISO8601 formato)	text
booleano	booleano	booleano
point	texto (en formato WKT)	geo_shape
línea	texto (en formato WKT)	geo_shape

Tipo de datos de RDS para PostgreSQL	Tipo de campo predeterminado OpenSearch	Tipo OpenSearch de campo recomendado
LSEG	texto (en formato WKT)	geo_shape
BOX	texto (en formato WKT)	geo_shape
path	texto (en formato WKT)	geo_shape
polígono	texto (en formato WKT)	geo_shape
CIRCLE	objeto	objeto
cidr	text	text
inet	text	text
macaddr	text	text
macaddr8	text	text
bit(n)	long	byte, corto, entero o largo (según el número de bits)
bit varying(n)	long	byte, corto, entero o largo (según el número de bits)
json	objeto	objeto
jsonb	objeto	objeto
jsonpath	text	text

Te recomendamos que configures la cola de mensajes sin salida (DLQ) en tu canal de ingestión. OpenSearch Si has configurado la cola, el OpenSearch servicio envía a la cola todos los documentos fallidos que no se puedan ingerir debido a errores de mapeo dinámico.

En caso de que las asignaciones automáticos fallen, puede usar `template_type` y `template_content` en su configuración de canalización para definir reglas de asignación explícitas. Como alternativa, puede crear plantillas de asignación directamente en su dominio o colección de búsqueda antes de iniciar la canalización.

Limitaciones

Tenga en cuenta las siguientes limitaciones al configurar una canalización de OpenSearch ingestión de RDS para PostgreSQL:

- La integración solo admite una base de datos PostgreSQL por canalización.
- Actualmente, la integración no admite la ingesta de datos entre regiones; la instancia y el OpenSearch dominio de Amazon RDS deben estar en el mismo lugar. Región de AWS
- Actualmente, la integración no admite la ingesta de datos entre cuentas; la instancia de Amazon RDS y la canalización de OpenSearch ingestión deben estar en la misma ubicación. Cuenta de AWS
- Asegúrese de que la instancia de Amazon RDS tenga habilitada la autenticación mediante AWS Secrets Manager, que es el único mecanismo de autenticación compatible.
- La configuración de canalización existente no se puede actualizar para ingerir datos de una base de datos diferente o de and/or una tabla diferente. Para actualizar la base de datos o el nombre de la tabla de una canalización, debe detener la canalización y reiniciarla con una configuración actualizada o crear una canalización nueva.
- Por lo general, las instrucciones en lenguaje de definición de datos (DDL) no son compatibles. La coherencia de datos no se mantendrá si:
  - Se cambian las claves principales (add/delete/rename).
  - Las tablas se eliminan o se truncan.
  - Se cambian los nombres de las columnas o los tipos de datos.
- Si las tablas de PostgreSQL que se van a sincronizar no tienen definidas las claves principales, no se garantiza la coherencia de datos. Deberá definir correctamente la `document_id` opción OpenSearch y la configuración del receptor para poder updates/deletes sincronizarla OpenSearch.
- No se admiten los clústeres de bases de datos Multi-AZ de RDS.
- Versiones compatibles: PostgreSQL 16 y versiones posteriores.

CloudWatch Alarmas recomendadas

Se recomiendan las siguientes CloudWatch métricas para supervisar el rendimiento del proceso de ingestión. Estas métricas pueden ayudarlo a identificar la cantidad de datos procesados a partir de exportaciones y flujos, los errores al procesar eventos de transmisión y exportación y el número de documentos escritos en el destino. Puedes configurar CloudWatch alarmas para realizar una acción cuando una de estas métricas supere un valor específico durante un período de tiempo específico.

Métrica	Description (Descripción)
pipeline-name .RDS. Se han modificado las credenciales	Esta métrica indica la frecuencia con la que se rotan los secretos. AWS
pipeline-name .rds. executorRefreshErrors	Esta métrica indica errores al actualizar AWS los secretos.
pipeline-name .rds. exportRecordsTotal	Esta métrica indica el número de registros exportados desde Amazon Aurora.
pipeline-name .rds. exportRecordsProcessed	Esta métrica indica la cantidad de registros procesados por la canalización OpenSearch de ingestión.
pipeline-name .rds. exportRecordProcessingErrores	Esta métrica indica el número de errores de procesamiento en una canalización de OpenSearch ingestión al leer los datos de un clúster de Amazon Aurora.
pipeline-name .rds.	Esta métrica indica el número total de registros de exportación procesados correctamente.

Métrica	Description (Descripción)
exportRecordsSuccessTotal	
pipeline-name .rds.exportRecordsFailedTotal	Esta métrica indica el número total de registros que no se han podido procesar.
pipeline-name .RDS. Bytes recibidos	Esta métrica indica el número total de bytes recibidos por una canalización de ingestión. OpenSearch
pipeline-name .rds.bytes procesados	Esta métrica indica el número total de bytes procesados por una canalización de ingestión. OpenSearch
pipeline-name .rds.streamRecordsSuccessTotal	Esta métrica indica el número de registros del flujo procesados correctamente.
pipeline-name .rds.streamRecordsFailedTotal	Esta métrica indica el número total de registros del flujo que no se han podido procesar.

Uso de una canalización OpenSearch de ingestión con Amazon S3

Con OpenSearch Ingestion, puede utilizar Amazon S3 como origen o destino. Cuando utiliza Amazon S3 como fuente, envía datos a una canalización de OpenSearch ingestión. Cuando utiliza Amazon S3 como destino, escribe datos de una canalización de OpenSearch ingestión en uno o más buckets de S3.

Temas

- Amazon S3 como origen
- Amazon S3 como destino

- Varias cuentas de Amazon S3 como origen

Amazon S3 como origen

Existen dos formas de utilizar Amazon S3 como origen para procesar datos: con el procesamiento S3-SQS y con análisis programados.

Utilice el procesamiento S3-SQS cuando necesite escanear los archivos casi en tiempo real una vez que se hayan escrito en S3. Puede configurar los buckets de Amazon S3 para que generen un evento cada vez que se almacene o modifique un objeto en el bucket. Utilice un escaneo programado único o recurrente para procesar por lotes los datos de un bucket de S3.

Temas

- Requisitos previos
- Paso 1: configurar el rol de canalización
- Paso 2: crear la canalización

Requisitos previos

Para usar Amazon S3 como fuente de una canalización de OpenSearch ingestión tanto para un escaneo programado como para un procesamiento S3-SQS, cree primero un bucket de S3.

> **Note**
>
> Si el bucket de S3 utilizado como fuente en la canalización de OpenSearch ingestión se encuentra en otro Cuenta de AWS, también debe habilitar los permisos de lectura multicuenta en el bucket. Esto permite que la canalización lea y procese los datos. Para habilitar los permisos entre cuentas, consulte Propietario del bucket que concede permisos de bucket entre cuentas en la Guía del usuario de Amazon S3.
> Si sus buckets de S3 están en varias cuentas, utilice una asignación `bucket_owners`. Para ver un ejemplo, consulta el acceso a S3 entre cuentas en la documentación. OpenSearch

Para configurar el procesamiento S3-SQS, también debe realizar los siguientes pasos:

1. Crear una cola de Amazon SQS.
2. Habilitar las notificaciones de eventos en el bucket de S3 con la cola SQS como destino.

Paso 1: configurar el rol de canalización

A diferencia de otros complementos de origen que envían datos a una canalización, el complemento de fuente de S3 tiene una arquitectura basada en lectura en la que la canalización extrae datos de la fuente.

Por lo tanto, para que una canalización pueda leer desde S3, debe especificar un rol dentro de la configuración de fuente de S3 de la canalización que tenga acceso tanto al bucket de S3 como a la cola de Amazon SQS. La canalización asumirá este rol para leer los datos de la cola.

> **Note**
>
> El rol que especifique en la configuración de fuente de S3 debe ser el rol de canalización. Por lo tanto, su rol de canalización debe contener dos políticas de permisos independientes: una para escribir en un receptor y otra para extraer de la fuente de S3. Debe usar el mismo `sts_role_arn` en todos los componentes de la canalización.

El siguiente ejemplo de política muestra los permisos necesarios para usar S3 como fuente:

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"
    }
  ]
}
```

Debe adjuntar estos permisos al rol de IAM que especifique en la opción `sts_role_arn` dentro de la configuración del complemento de fuente de S3:

```
version: "2"
source:
  s3:
    ...
    aws:
      ...
processor:
  ...
sink:
  - opensearch:
      ...
```

Paso 2: crear la canalización

Una vez configurados los permisos, puede configurar una canalización de OpenSearch ingestión en función del caso de uso de Amazon S3.

Procesamiento S3-SQS

Para configurar el procesamiento S3-SQS, configure su canalización para especificar S3 como origen y configure las notificaciones de 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"
```

Si observa un bajo uso de la CPU al procesar archivos pequeños en Amazon S3, considere la posibilidad de aumentar el rendimiento mediante la modificación del valor de la opción `workers`. Para obtener más información, consulte S3 plugin configuration options.

Escaneo programado

Para configurar un escaneo programado, configure su canalización con un cronograma a nivel de escaneo que se aplique a todos los buckets de S3 o a nivel de bucket. Una programación a nivel de bucket o una configuración de intervalos de escaneo siempre sobrescribe una configuración a nivel de escaneo.

Puede configurar los escaneos programados con un escaneo único, que es ideal para la migración de datos, o un escaneo periódico, que es ideal para el procesamiento por lotes.

Para configurar su canalización para que lea desde Amazon S3, utilice los esquemas de Amazon S3 preconfigurados. Puede editar la parte `scan` de la configuración de la canalización para adaptarla a sus necesidades de programación. Para más información, consulte the section called “Uso de esquemas”.

Escaneo único

Un escaneo único programado se ejecuta una sola vez. En su configuración de la canalización, puede usar una `start_time` y `end_time` para especificar cuándo quiere que se escaneen los objetos del bucket. O bien, puede usar un `range` para especificar el intervalo de tiempo en relación con la hora actual en el que desea que se escaneen los objetos del bucket.

Por ejemplo, un rango configurado para PT4H escanea todos los archivos creados en las últimas cuatro horas. Para configurar un escaneo único para que se ejecute por segunda vez, debe detener y reiniciar la canalización. Si no tiene un rango configurado, también debe actualizar las horas de inicio y finalización.

La siguiente configuración establece un escaneo único de todos los buckets y todos los objetos de esos buckets:

```
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 siguiente configuración establece un escaneo único de todos los buckets durante un período de tiempo específico. Esto significa que S3 procesa solo los objetos con tiempos de creación que se encuentran dentro de este período.

```
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 siguiente configuración establece un escaneo único tanto a nivel de escaneo como a nivel del bucket. Las horas de inicio y finalización a nivel del bucket anulan las horas de inicio y finalización a nivel de escaneo.

```
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
```

Al detener una canalización, se elimina cualquier referencia preexistente de los objetos que la canalización analizó antes de la detención. Si se detiene una sola canalización de análisis, volverá a analizar todos los objetos una vez iniciada, incluso si ya se analizaron. Si necesita detener una sola canalización de análisis, se recomienda cambiar el intervalo de tiempo antes de volver a iniciar la canalización.

Si necesita filtrar los objetos por hora de inicio y hora de finalización, la única opción es detener e iniciar la canalización. Si no necesita filtrar por hora de inicio y hora de finalización, puede filtrar los objetos por nombre. No es necesario detener e iniciar la canalización para filtrar por nombre. Para ello, use `include_prefix` y `exclude_suffix`.

Escaneo periódico

Un escaneo periódico programado ejecuta un escaneo de los buckets de S3 especificados a intervalos regulares y programados. Solo puede configurar estos intervalos a nivel de escaneo, ya que no se admiten configuraciones individuales a nivel de bucket.

En la configuración de la canalización, el `interval` especifica la frecuencia del escaneo periódico, que puede ser entre 30 segundos y 365 días. El primero de estos escaneos siempre se produce al crear la canalización. El `count` define el número total de instancias de escaneo.

La siguiente configuración establece un escaneo periódico, con un retraso de 12 horas entre los escaneos:

```
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 como destino

Para escribir datos de una canalización de OpenSearch ingestión en un bucket de S3, utilice el esquema preconfigurado de S3 para crear una canalización con un receptor de S3. Esta canalización dirige los datos selectivos a un OpenSearch sumidero y, simultáneamente, envía todos los datos para archivarlos en S3. Para obtener más información, consulte the section called “Uso de esquemas”.

Al crear el receptor de S3, puede especificar el formato que prefiera entre una variedad de códecs de receptor. Por ejemplo, si desea escribir datos en formato de columnas, seleccione el códec Parquet o Avro. Si prefiere un formato basado en filas, seleccione JSON o NDJSON. Para escribir datos en S3 en un esquema específico, también puede definir un esquema insertado dentro de los códecs de receptor usando el formato Avro.

El siguiente ejemplo define un esquema insertado en un receptor de 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"}
           ]
         }
```

Cuando defina este esquema, especifique un superconjunto de todas las claves que pueden estar presentes en los distintos tipos de eventos que su canalización envía a un receptor.

Por ejemplo, si en un evento existe la posibilidad de que falte una clave, añada esa clave al esquema con un valor `null`. Las declaraciones de valores nulos permiten que el esquema procese datos no uniformes (algunos eventos tienen estas claves y otros no). Cuando los eventos entrantes tienen estas claves presentes, sus valores se escriben en los receptores.

Esta definición de esquema actúa como un filtro que solo permite enviar las claves definidas a los receptores y elimina las claves indefinidas de los eventos entrantes.

También puede usar `include_keys` y `exclude_keys` en su receptor para filtrar los datos que se dirigen a otros receptores. Estos dos filtros se excluyen mutuamente, por lo que solo puede usar uno a la vez en su esquema. Además, no puede utilizarlos en esquemas definidos por el usuario.

Para crear canalizaciones con dichos filtros, utilice el esquema de filtro de receptor preconfigurado. Para obtener más información, consulte the section called “Uso de esquemas”.

Varias cuentas de Amazon S3 como origen

Puede conceder acceso a todas las cuentas con Amazon S3 para que las canalizaciones de OpenSearch ingestión puedan acceder a los buckets de S3 de otra cuenta como fuente. Para habilitar los accesos entre cuentas, consulte Propietario del bucket que concede permisos de bucket entre cuentas en la Guía del usuario de Amazon S3. Una vez que haya concedido el acceso, asegúrese de que su rol en la canalización cuente con los permisos necesarios.

A continuación, puede crear una canalización con `bucket_owners` para habilitar el acceso entre cuentas a un bucket de Amazon S3 como origen:

```
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"
```

Uso de una canalización de OpenSearch ingestión con Amazon Security Lake

Puede usar el complemento fuente de S3 para incorporar datos de Amazon Security Lake a su canalización de OpenSearch ingestión. Security Lake centraliza automáticamente los datos de seguridad de los AWS entornos, los entornos locales y los proveedores de SaaS en un lago de datos

diseñado específicamente. Puede crear una suscripción que replique los datos de Security Lake en su canal de OpenSearch ingestion y, a continuación, los guarde en su dominio de servicio o en su colección Serverless. OpenSearch OpenSearch

Para configurar su canalización para que lea desde Security Lake, utilice el esquema de Security Lake preconfigurado. El esquema incluye una configuración predeterminada para incorporar los archivos Parquet de Open Cybersecurity Schema Framework (OCSF) de Security Lake. Para obtener más información, consulte the section called “Uso de esquemas”.

Temas

- Uso de una canalización de OpenSearch ingestión con Amazon Security Lake como fuente
- Uso de una canalización de OpenSearch ingestión con Amazon Security Lake como sumidero

Uso de una canalización de OpenSearch ingestión con Amazon Security Lake como fuente

Puede usar el complemento fuente de Amazon S3 en su canalización OpenSearch de ingestión para ingerir datos de Amazon Security Lake. Security Lake centraliza automáticamente los datos de seguridad de los entornos de AWS , los sistemas en las instalaciones y los proveedores de SaaS en un lago de datos diseñado específicamente.

Amazon Security Lake tiene los siguientes atributos de metadatos dentro de una canalización:

- `bucket_name`: nombre del bucket de Amazon S3 creado por Security Lake para almacenar datos de seguridad.
- `path_prefix`: el nombre del origen personalizado definido en la política de roles de IAM de Security Lake.
- `region`: El Región de AWS lugar donde se encuentra el depósito S3 de Security Lake.
- `accountID`: El Cuenta de AWS ID en el que está activado Security Lake.
- `sts_role_arn`: el ARN del rol de IAM diseñado para usarse con Security Lake.

Requisitos previos

Antes de crear la canalización OpenSearch de ingestión, lleve a cabo los siguientes pasos:

- Habilitar Security Hub.
- Crear un suscriptor en Security Lake.

- Seleccione las fuentes que quiere incorporar a su canalización.
- En el caso de las credenciales de suscriptor, añade el ID del Cuenta de AWS lugar en el que deseas crear la canalización. Para el ID externo, especifique `OpenSearchIngestion-{accountid}`.
- Para el Método de acceso a los datos, seleccione S3.
- Para los Detalles de notificación, seleccione Cola de SQS.

Al crear un suscriptor, Security Lake crea automáticamente dos políticas de permisos insertadas: una para S3 y otra para SQS. Las políticas adoptan el siguiente formato: `AmazonSecurityLake-amzn-s3-demo-bucket-S3` y `AmazonSecurityLake-AWS Demo-SQS`. Para permitir que su canalización acceda a las fuentes del suscriptor, debe asociar los permisos necesarios a su rol de canalización.

Configuración del rol de la canalización

Cree una nueva política de permisos en IAM que combine solo los permisos necesarios de las dos políticas que Security Lake creó automáticamente. El siguiente ejemplo de política muestra los privilegios mínimos necesarios para que una canalización de OpenSearch ingestion lea datos de varias fuentes de 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 no administra la política de roles de canalización por usted. Si agrega o elimina fuentes de su suscripción de Security Lake, debe actualizar la política manualmente. Security Lake crea particiones para cada fuente de registro, por lo que debe agregar o eliminar manualmente los permisos en el rol de canalización.

Debe adjuntar estos permisos al rol de IAM que especifique en la opción `sts_role_arn` incluida en la configuración del complemento fuente de S3, en `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:
      ...
```

Creación de la canalización

Después de agregar los permisos al rol de la canalización, utilice el esquema de Security Lake preconfigurado para crear la canalización. Para obtener más información, consulte the section called "Uso de esquemas".

Debe especificar la opción `queue_url` en la configuración fuente de `s3`, que es la URL de la cola de Amazon SQS desde la que debe leerse. Para formatear la URL, localice el Punto de conexión de la suscripción en la configuración del suscriptor y cambie `arn:aws:` a `https://`. Por ejemplo, `https://sqs.`*`us-east-1`*`amazonaws.com/`*`account-id`*`/AmazonSecurityLake-`*`AWS Demo`*`-Main-Queue`.

El `sts_role_arn` que especifica en la configuración de fuente de S3 debe ser el ARN del rol de canalización.

Uso de una canalización de OpenSearch ingestión con Amazon Security Lake como sumidero

Utilice el complemento Amazon S3 sink en OpenSearch Ingestion para enviar datos desde cualquier fuente compatible a Amazon Security Lake. Security Lake recopila y almacena datos de seguridad de AWS entornos locales y proveedores de SaaS en un lago de datos dedicado.

Para configurar su canalización para escribir datos de registro en Security Lake, utilice el esquema de registros de tráfico de firewall preconfigurado. El esquema incluye una configuración predeterminada para recuperar registros de seguridad sin procesar u otros datos almacenados en un bucket de Amazon S3, procesar los registros y normalizarlos. A continuación, asigna los datos al marco Open Cybersecurity Schema Framework (OCSF) y envía los datos transformados compatibles con OCSF a Security Lake.

La canalización tiene los siguientes atributos de metadatos:

- `bucket_name`: nombre del bucket de Amazon S3 creado por Security Lake para almacenar datos de seguridad.
- `path_prefix`: el nombre del origen personalizado definido en la política de roles de IAM de Security Lake.
- `region`: El Región de AWS lugar donde se encuentra el depósito S3 de Security Lake.

- `accountID`: El Cuenta de AWS ID en el que está activado Security Lake.
- `sts_role_arn`: el ARN del rol de IAM diseñado para usarse con Security Lake.

Requisitos previos

Antes de crear una canalización para enviar datos a Security Lake, realice los siguientes pasos:

- Habilitar y configurar Amazon Security Lake: configure Amazon Security Lake para centralizar los datos de seguridad de diversos orígenes. Para obtener instrucciones, consulte Habilitar Security Lake mediante la consola.

  Al seleccionar un origen, elija Ingesta de orígenes de AWS específicos y seleccione uno o más orígenes de registros y eventos que desee ingerir.
- Configure los permisos: configure la función de canalización con los permisos necesarios para escribir datos en Security Lake. Para más información, consulte Rol de canalización.

Creación de la canalización

Utilice el esquema de Security Lake preconfigurado para crear la canalización. Para obtener más información, consulte Uso de esquemas para crear una canalización.

Uso de una canalización OpenSearch de ingestión con Fluent Bit

Este ejemplo de archivo de configuración de Fluent Bit envía los datos de registro de Fluent Bit a una canalización OpenSearch de ingestión. Para más información sobre la ingesta de datos de registro, consulte Log Analytics en la documentación de Data Prepper.

Tenga en cuenta lo siguiente:

- El valor `host` debe ser el punto de conexión de su canalización. Por ejemplo, *`pipeline-endpoint.us-east-1`*`osis.amazonaws.com`.
- El valor `aws_service` debe ser `osis`.
- El `aws_role_arn` valor es el ARN de la función de AWS IAM que el cliente debe asumir y utilizar para la autenticación de la versión 4 de Signature.

```
[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
```

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, que tiene HTTP como origen:

```
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"
```

Uso de una canalización OpenSearch de ingestión con Fluentd

Fluentd es un ecosistema de recopilación de datos de código abierto que proporciona SDKs diferentes lenguajes y subproyectos, como Fluent Bit. Este ejemplo de archivo de configuración de Fluentd envía los datos de registro de Fluentd a una canalización de ingestion. OpenSearch Para más información sobre la ingesta de datos de registro, consulte Log Analytics en la documentación de Data Prepper.

Tenga en cuenta lo siguiente:

- El valor `endpoint` debe ser el punto de conexión de su canalización. Por ejemplo, *`pipeline-endpoint`*`.us-east-1`osis.amazonaws.com/apache-log-pipeline/logs.
- El valor `aws_service` debe ser `osis`.
- El `aws_role_arn` valor es el ARN de la función de AWS IAM que el cliente debe asumir y utilizar para la autenticación de la versión 4 de Signature.

```
<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>
```

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, que tiene HTTP como origen:

```
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
```

Uso de una canalización OpenSearch de ingestión con aprendizaje automático e inferencia por lotes fuera de línea

Las canalizaciones de OpenSearch Amazon Ingestion (OSI) admiten el procesamiento de inferencias por lotes fuera de línea con aprendizaje automático (ML) para enriquecer de manera eficiente grandes volúmenes de datos a bajo costo. Utilice la inferencia en lotes sin conexión siempre que tenga conjuntos de datos grandes que puedan procesarse de forma asíncrona. La inferencia de lotes sin conexión funciona con Amazon Bedrock y SageMaker sus modelos. Esta función está disponible en todos los dominios Regiones de AWS compatibles con OpenSearch Ingestion with OpenSearch Service 2.17 o más.

> **Note**
>
> Para el procesamiento de inferencias en tiempo real, utilice Conectores Amazon OpenSearch Service ML para plataformas de terceros.

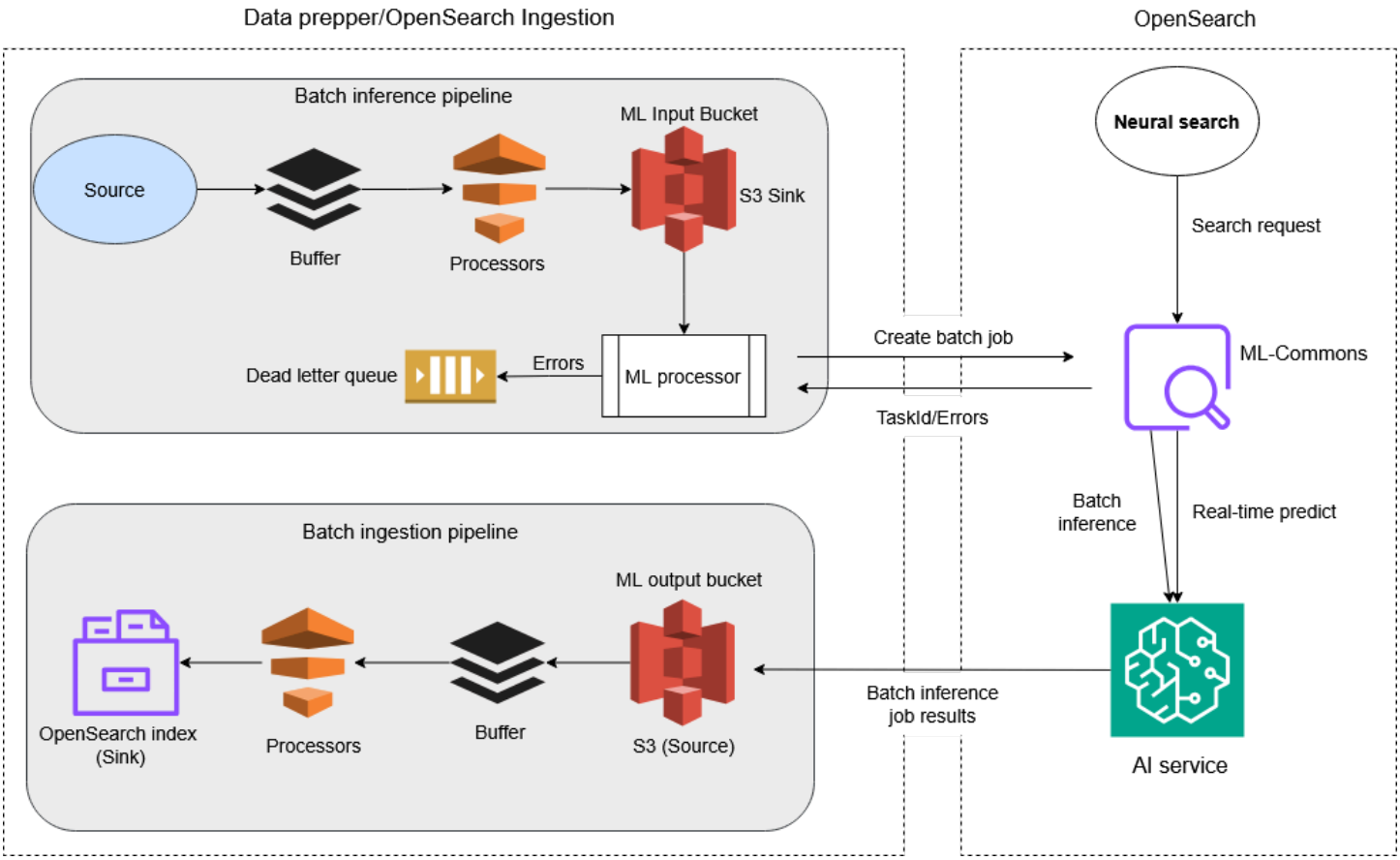
El procesamiento de inferencias por lotes fuera de línea aprovecha una función llamada ML Commons. OpenSearch ML Commons proporciona algoritmos de ML mediante llamadas a la API de REST y de transporte. Esas llamadas eligen los nodos y los recursos correctos para cada solicitud de ML y supervisan las tareas de ML para garantizar el tiempo de actividad. De esta manera, ML Commons le permite aprovechar los algoritmos de ML de código abierto existentes y reducir el esfuerzo necesario para desarrollar nuevas características de ML. Para obtener más información sobre ML Commons, consulte Aprendizaje automático en la documentación de OpenSearch .org.

Funcionamiento

Puede crear una canalización de inferencias por lotes sin conexión en OpenSearch Ingestion añadiendo un procesador de inferencias de aprendizaje automático a una canalización. Este procesador permite que tu canalización se conecte a servicios de IA, por ejemplo, SageMaker para ejecutar trabajos de inferencia por lotes. Puede configurar su procesador para que se conecte al

servicio de IA que desee a través de los conectores de IA (compatibles con batch_predict) que se ejecutan en su dominio de destino.

OpenSearch Ingestion utiliza el `ml_inference` procesador con ML Commons para crear trabajos de inferencia por lotes sin conexión a Internet. A continuación, ML Commons utiliza la API batch_predict, que realiza inferencias sobre conjuntos de datos de gran tamaño en un modo asíncrono sin conexión mediante un modelo implementado en servidores de modelos externos en Amazon Bedrock, Amazon, Cohere y OpenAI. SageMaker El siguiente diagrama muestra una canalización de OpenSearch ingestión que organiza varios componentes para llevar a cabo este proceso de principio a fin:



Los componentes de la canalización funcionan de la siguiente manera:

Canalización 1 (Preparación y transformación de datos)*:

- Fuente: los datos se escanean desde una fuente externa compatible con Ingestion. OpenSearch
- Procesadores de datos: los datos sin procesar se procesan y transforman al formato correcto para su inferencia en lotes en el servicio de IA integrado.

- S3 (receptor): los datos procesados se almacenan en un bucket de Amazon S3 y están listos para servir como entrada para ejecutar tareas de inferencia en lotes en el servicio de IA integrado.

Canalización 2 (Activar batch_inference de ML):

- Origen: detección automática de eventos de S3 de nuevos archivos creados por la salida de la Canalización 1.
- Procesador ml_inference: procesador que genera inferencias de ML mediante una tarea en lotes asíncrona. Se conecta a los servicios de IA a través del conector de IA configurado que se ejecuta en el dominio de destino.
- ID de tarea: cada tarea en lotes está asociada a un ID de tarea en ml-commons para su seguimiento y administración.
- OpenSearch ML Commons: ML Commons, que aloja el modelo de búsqueda neuronal en tiempo real, gestiona los conectores a los servidores remotos de IA y sirve APIs para la inferencia de lotes y la gestión de trabajos.
- Servicios de IA: OpenSearch ML Commons interactúa con servicios de IA como Amazon Bedrock y Amazon SageMaker para realizar inferencias por lotes de los datos, produciendo predicciones o información. Los resultados se guardan de forma asíncrona en un archivo S3 independiente.

Canalización 3 (Ingesta masiva):

- S3 (origen): los resultados de las tareas en lotes se almacenan en S3, que es el origen de esta canalización.
- Procesadores de transformación de datos: el procesamiento y la transformación adicionales se aplican al resultado de la inferencia en lotes antes de la ingesta. Esto garantiza que los datos estén mapeados correctamente en el índice. OpenSearch
- OpenSearch índice (sumidero): los resultados procesados se indexan OpenSearch para almacenarlos, buscarlos y analizarlos más a fondo.

> **Note**
>
> * El proceso descrito en Canalización 1 es opcional. Si lo prefiere, puede omitir ese proceso y solo cargar los datos preparados en el receptor de S3 para crear tareas en lotes.

Acerca del procesador ml_inference

OpenSearch Ingestion utiliza una integración especializada entre la fuente de S3 Scan y el procesador de inferencias ML para el procesamiento por lotes. El S3 Scan funciona solo en modo de metadatos para recopilar de manera eficiente la información de los archivos S3 sin leer el contenido real del archivo. El `ml_inference` procesador utiliza el archivo S3 URLs para coordinarse con ML Commons para el procesamiento por lotes. Este diseño optimiza el flujo de trabajo de inferencia en lotes al minimizar la transferencia de datos innecesaria durante la fase de digitalización. El procesador `ml_inference` se define mediante parámetros. A continuación se muestra un ejemplo:

```
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:
            # Región de 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"
```

Mejoras en el rendimiento de la ingesta mediante el procesador ml_inference

El `ml_inference` procesador OpenSearch de ingestión mejora considerablemente el rendimiento de la ingesta de datos para las búsquedas habilitadas para ML. El procesador es ideal para casos de uso que requieren datos generados por modelos de machine learning, como la búsqueda semántica, la búsqueda multimodal, el enriquecimiento de documentos y la comprensión de consultas. En la búsqueda semántica, el procesador puede acelerar la creación e ingesta de vectores de gran volumen y alta dimensión en un orden de magnitud.

La capacidad de inferencia en lotes sin conexión del procesador ofrece claras ventajas en comparación con la invocación de modelos en tiempo real. Si bien el procesamiento en tiempo real requiere un servidor modelo en vivo con limitaciones de capacidad, la inferencia en lotes escala de forma dinámica los recursos de cómputo bajo demanda y procesa los datos en paralelo. Por ejemplo, cuando la canalización de OpenSearch ingestión recibe mil millones de solicitudes de datos de origen, crea 100 archivos S3 para la entrada de inferencias por lotes de aprendizaje automático. A continuación, el `ml_inference` procesador inicia un trabajo SageMaker por lotes con 100 instancias de `ml.m4.xlarge` Amazon Elastic Compute Cloud (Amazon EC2) y completa la vectorización de mil millones de solicitudes en 14 horas, una tarea que sería prácticamente imposible de realizar en tiempo real.

Configurar el procesador ml_inference de modo que ingiera las solicitudes de datos para una búsqueda semántica

Los siguientes procedimientos explican el proceso de instalación y configuración del `ml_inference` procesador de OpenSearch ingestión para procesar mil millones de solicitudes de datos con fines de búsqueda semántica mediante un modelo de incrustación de texto.

Temas

- Paso 1: Cree conectores y registre modelos en OpenSearch
- Paso 2: Cree una canalización de OpenSearch ingestión para la inferencia de lotes fuera de línea de ML
- Paso 3: preparar los datos para la ingesta
- Paso 4: Supervisar la tarea de inferencia en lotes
- Paso 5: ejecute la búsqueda

Paso 1: Cree conectores y registre modelos en OpenSearch

Para el siguiente procedimiento, utilice el batch_inference_sagemaker_connector_blueprint de ML Commons para crear un conector y un modelo en Amazon. SageMaker Si prefiere usar plantillas de integración, consulte más adelante en esta sección. OpenSearch CloudFormation (Procedimiento alternativo) Paso 1: Cree conectores y modelos mediante una plantilla de CloudFormation integración

Para crear conectores y registrar modelos en OpenSearch

1. Cree un modelo ML de Deep Java Library (DJL) SageMaker para la transformación por lotes. Para ver otros modelos de DJL, consulte Semantic_Search_With_CFN_Template_for_SageMaker en: 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. Cree un conector con `batch_predict` como el nuevo tipo `action` en el 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. Utilice el ID de conector devuelto para registrar el modelo: SageMaker

```
POST /_plugins/_ml/models/_register
{
    "name": "SageMaker model for batch",
    "function_name": "remote",
    "description": "test model",
    "connector_id": "example123456789-abcde"
}
```

4. Invoque el modelo con el tipo de acción `batch_predict`:

```
POST /_plugins/_ml/models/teHr3JABBiEvs-eod7sn/_batch_predict
{
  "parameters": {
    "TransformJobName": "SM-offline-batch-transform"
  }
}
```

La respuesta contiene un identificador de tarea para la tarea en lotes:

```
{
 "task_id": "exampleIDabdcefd_1234567",
 "status": "CREATED"
}
```

5. Compruebe el estado de la tarea en lotes llamando a la API Get Task con el ID de tarea:

```
GET /_plugins/_ml/tasks/exampleIDabdcefd_1234567
```

La respuesta contiene el estado de la tarea:

```
{
  "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": "$.content",
      "JoinSource": "Input",
      "OutputFilter": "$"
    }
  }
}
```

(Procedimiento alternativo) Paso 1: Cree conectores y modelos mediante una plantilla de CloudFormation integración

Si lo prefiere, puede utilizarlos AWS CloudFormation para crear automáticamente todos los SageMaker conectores y modelos de Amazon necesarios para la inferencia de aprendizaje automático. Este enfoque simplifica la configuración mediante el uso de una plantilla preconfigurada disponible en la consola de Amazon OpenSearch Service. Para obtener más información, consulte Se utiliza CloudFormation para configurar la inferencia remota para la búsqueda semántica.

Para implementar una CloudFormation pila que cree todos los SageMaker conectores y modelos necesarios

1. Abre la consola OpenSearch de Amazon Service.
2. En el panel de navegación, elija Integraciones.
3. En el campo de búsqueda, introduce y**`SageMaker`**, a continuación, selecciona Integración con modelos de incrustación de texto a través de Amazon SageMaker.
4. Seleccione Configurar dominio y, a continuación, seleccione Configurar dominio de VPC o Configurar dominio público.
5. Ingrese la información en los campos de la plantilla. En Habilitar la inferencia en lotes sin conexión, elija true para aprovisionar recursos para el procesamiento en lotes sin conexión.
6. Selecciona Crear para crear la CloudFormation pila.
7. Una vez creada la pila, abra la pestaña Salidas en la CloudFormation consola y busque el connector_id y el model_id. Necesitará estos valores más adelante cuando configure la canalización.

Paso 2: Cree una canalización de OpenSearch ingestión para la inferencia de lotes fuera de línea de ML

Utilice el siguiente ejemplo para crear una canalización de OpenSearch ingestión para la inferencia de lotes de ML sin conexión. Para obtener más información sobre la creación de una canalización para OpenSearch Ingestion, consulte. Creación de canalizaciones OpenSearch de Amazon Ingestion

Antes de empezar

En el siguiente ejemplo, se especifica un ARN de rol de IAM para el parámetro `sts_role_arn`. Utilice el siguiente procedimiento para comprobar que este rol está asignado al rol de backend que tiene acceso a ml-commons en. OpenSearch

1. Navegue hasta el complemento OpenSearch Dashboards de su dominio de servicio. OpenSearch Puedes encontrar el punto de conexión del panel de control en el panel de control de tu dominio, en la consola de OpenSearch servicio.
2. En el menú principal, seleccione Seguridad, Roles y seleccione el rol ml_full_access.
3. Seleccione Usuarios asignados, Administrar mapeo.
4. En Roles de backend, escriba el ARN del rol de Lambda que necesita permiso para llamar a su dominio. Este es un ejemplo: arn:aws:iam: ::role/ *111122223333 lambda-role*
5. Seleccione Asignar y confirme que el usuario o el rol aparecen en Usuarios asignados.

Ejemplo de creación de una canalización de ingestión para la inferencia de lotes offline de 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
```

Paso 3: preparar los datos para la ingesta

Para preparar los datos para el procesamiento de inferencias por lotes de aprendizaje automático fuera de línea, prepare los datos usted mismo con sus propias herramientas o procesos o utilice el OpenSearch Data Prepper. Compruebe que los datos estén organizados en el formato correcto, ya sea mediante una canalización para consumir los datos del origen de datos o con la creación de un conjunto de datos de machine learning.

El siguiente ejemplo usa el conjunto de datos MS MARCO, que incluye una colección de consultas de usuarios reales para tareas de procesamiento del lenguaje natural. El conjunto de datos está estructurado en formato JSONL, donde cada línea representa una solicitud enviada al modelo de incrustación de 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"}}
...
```

Para realizar una prueba con el conjunto de datos MS MARCO, imagine un escenario en el que cree mil millones de solicitudes de entrada distribuidas en 100 archivos, cada uno con 10 millones de solicitudes. Los archivos se almacenarían en Amazon S3 con el prefijo s3://offlinebatch/sagemaker/sagemaker_djl_batch_input/. El proceso OpenSearch de ingestión escanearía estos 100 archivos simultáneamente e iniciaría un trabajo SageMaker por lotes con 100 trabajadores para su procesamiento en paralelo, lo que permitiría la vectorización y la ingestión eficientes de los mil millones de documentos en los que se encuentran. OpenSearch

En los entornos de producción, puede utilizar una canalización de OpenSearch ingestión para generar archivos S3 para la entrada de datos por lotes. La canalización admite varios orígenes de datos y funciona según un cronograma para transformar de manera continua los datos de origen

en archivos de S3. Luego, los servidores de IA procesan automáticamente estos archivos mediante tareas en lotes programadas sin conexión, lo que garantiza el procesamiento y la ingesta continuos de datos.

Paso 4: Supervisar la tarea de inferencia en lotes

Puede supervisar los trabajos de inferencia por lotes mediante la SageMaker consola o el. AWS CLI También puede utilizar la API Get Task para supervisar tareas en lotes:

```
GET /_plugins/_ml/tasks/_search
{
  "query": {
    "bool": {
      "filter": [
        {
          "term": {
            "state": "RUNNING"
          }
        }
      ]
    }
  },
  "_source": ["model_id", "state", "task_type", "create_time", "last_update_time"]
}
```

La API devuelve una lista de tareas en lotes activas:

```
{
  "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
        }
      }
    ]
  }
}
```

Paso 5: ejecute la búsqueda

Tras supervisar la tarea de inferencia en lotes y confirmar que se ha completado, puede ejecutar varios tipos de búsquedas de IA, incluidas las semánticas, las híbridas, las conversacionales (con RAG), las dispersas neuronales y las multimodales. Para obtener más información sobre las búsquedas de IA compatibles con el OpenSearch Servicio, consulte Búsqueda de IA.

Para buscar vectores sin procesar, utilice el tipo de consulta `knn`, proporcione la matriz `vector` como entrada y especifique el `k` número de resultados devueltos:

```
GET /my-raw-vector-index/_search
{
  "query": {
    "knn": {
      "my_vector": {
        "vector": [0.1, 0.2, 0.3],
        "k": 2
      }
    }
  }
}
```

Para ejecutar una búsqueda basada en la IA, utilice el tipo de consulta `neural`. Especifique la `query_text` entrada, el modelo `model_id` de incrustación que configuró en el proceso de OpenSearch ingestión y el `k` número de resultados devueltos. Para excluir las incrustaciones de los resultados de búsqueda, especifique el nombre del campo de incrustación en el parámetro: `_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
      }
```

```
    }
  }
}
```

Uso de una canalización OpenSearch de ingestión con Collector OpenTelemetry

Puede usar el OpenTelemetry recopilador para incorporar registros, trazas y métricas en los canales de OpenSearch ingestión. Se puede usar una única canalización para incorporar todos los registros, trazas y métricas a distintos índices de un dominio o colección. También puedes usar canalizaciones para ingerir solo registros, trazas o métricas de forma individual.

Temas

- Requisitos previos
- Paso 1: configurar el rol de canalización
- Paso 2: crear la canalización
- Conectividad entre cuentas
- Limitaciones
- CloudWatch Alarmas recomendadas para las fuentes OpenTelemetry

Requisitos previos

Al configurar el archivo de OpenTelemetry configuración, debe configurar lo siguiente para que se produzca la ingestión:

- La función de ingestión necesita el `osis:Ingest` permiso para interactuar con la canalización. Para obtener más información, consulta la función de ingestión.
- El valor del punto final debe incluir el punto final de la canalización. Por ejemplo, `https://pipeline-endpoint.us-east-1.osis.amazonaws.com.`
- El valor del servicio debe ser`osis`.
- La opción de compresión del OTLP/HTTP exportador debe coincidir con la opción de compresión de la fuente seleccionada de la canalización.

```
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]
```

Paso 1: configurar el rol de canalización

Tras configurar el OpenTelemetry recopilador, configura el rol de canalización que deseas usar en la configuración de la canalización. La función de canalización no necesita permisos específicos para la fuente de OTLP, solo permisos para conceder a las canalizaciones el acceso al OpenSearch dominio o la colección.

Paso 2: crear la canalización

A continuación, puedes configurar una canalización OpenSearch de ingestión como la siguiente, que especifica OTLP como fuente. También puedes configurar OpenTelemetry los registros, las métricas y los seguimientos como fuentes individuales.

Configuración de la canalización de origen 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 configuración de la canalización:

```
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 Configuración de la canalización de métricas:

```
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 Configuración de la canalización de rastreos:

```
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"
```

Puede utilizar un esquema preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Conectividad entre cuentas

OpenSearch Las canalizaciones de ingesta con OpenTelemetry fuentes tienen la capacidad de ingesta entre cuentas. Amazon OpenSearch Ingestion le permite compartir canalizaciones Cuentas de AWS desde una nube privada virtual (VPC) a un punto final de canalización en una VPC independiente. Para obtener más información, consulte the section called “Configuración de canalizaciones para la ingesta entre cuentas”.

Limitaciones

La canalización OpenSearch de ingestión no puede recibir solicitudes de más de 20 MB. El usuario configura este valor en la `max_request_length` opción. El valor predeterminado de esta opción es 10 MB.

CloudWatch Alarmas recomendadas para las fuentes OpenTelemetry

Se recomiendan las siguientes CloudWatch métricas para supervisar el rendimiento del proceso de ingestión. Estas métricas pueden ayudarle a identificar la cantidad de datos procesados a partir de exportaciones y flujos, los errores al procesar eventos de transmisión y exportación y la cantidad de documentos escritos en el destino. Puedes configurar CloudWatch alarmas para realizar una acción cuando una de estas métricas supere un valor específico durante un período de tiempo específico.

Las CloudWatch métricas de la fuente OTLP tienen el formato siguiente. `{pipeline-name}.otlp.{logs | traces | metrics}.{metric-name}` Por ejemplo, `otel-pipeline.otlp.metrics.requestTimeouts.count`.

En el caso de utilizar una OpenTelemetry fuente individual, las métricas se formatearán como. `{pipeline-name}.{source-name}.{metric-name}` Por ejemplo, `trace-pipeline.otel_trace_source.requestTimeouts.count`.

OpenTelemetry Los tres tipos de datos tendrán las mismas métricas, pero por motivos de brevedad, las métricas solo se enumerarán en la siguiente tabla para los datos de tipo registro de origen OTLP.

Métrica	Description (Descripción)
`otel-pipeline.BlockingBuffer.bufferUsage.value`	Indica qué cantidad del búfer se está utilizando.
`otel-pipeline.otlp.logs.requestTimeouts.count`	El número de solicitudes cuyo tiempo de espera se ha agotado.
`otel-pipeline.otlp.logs.requestsReceived.count`	El número de solicitudes recibidas por el OpenTelemetry recopilador.
`otel-pipeline.otlp.logs.badRequests.count`	El número de solicitudes con formato incorrecto recibidas por el OpenTelemetry recopilador.
`otel-pipeline.otlp.logs.requestsTooLarge.count`	El número de solicitudes que supera el máximo de 20 MB recibidas por el OpenTelemetry recopilador.
`otel-pipeline.otlp.logs.internalServerError.count`	El número de errores HTTP 500 recibidos del OpenTelemetry recopilador.
`otel-pipeline.opensearch.bulkBadRequestErrors.count`	Recuento de errores durante las solicitudes masivas debido a un formato incorrecto de la solicitud.
`otel-pipeline.opensearch.bulkRequestLatency.avg`	Latencia media de las solicitudes de escritura masiva realizadas a OpenSearch.

Métrica	Description (Descripción)
otel-pipeline.opensearch.bulkRequestNotFoundErrors.count	Número de solicitudes masivas que fallaron porque no se pudieron encontrar los datos de destino.
otel-pipeline.opensearch.bulkRequestNumberOfRetries.count	Número de reintentos realizados por las canalizaciones OpenSearch de ingestión para escribir el clúster. OpenSearch
otel-pipeline.opensearch.bulkRequestSizeBytes.sum	Tamaño total en bytes de todas las solicitudes masivas realizadas a. OpenSearch
otel-pipeline.opensearch.documentErrors.count	Número de errores al enviar documentos a OpenSearch. Los documentos que causan los errores se enviarán a DLQ.
otel-pipeline.opensearch.documentsSuccess.count	Número de documentos escritos correctamente en un OpenSearch clúster o colección.
otel-pipeline.opensearch.documentsSuccessFirstAttempt.count	Número de documentos indexados correctamente OpenSearch en el primer intento.
otel-pipeline.opensearch.documentsVersionConflictErrors.count	Recuento de errores debidos a conflictos de versiones en los documentos durante el procesamiento.
otel-pipeline.opensearch.PipelineLatency.avg	Latencia media de la canalización de OpenSearch ingestión para procesar los datos desde la fuente hasta escribirlos en el destino.
otel-pipeline.opensearch.PipelineLatency.max	Latencia máxima de la canalización de OpenSearch ingestión para procesar los datos desde el origen hasta escribir el destino.

Métrica	Description (Descripción)
`otel-pipeline.opensearch.recordsIn.count`	Recuento de registros ingresados correctamente. OpenSearch Esta métrica es esencial para realizar un seguimiento del volumen de datos que se procesan y almacenan.
`otel-pipeline.opensearch.s3.dlqS3RecordsFailed.count`	Número de registros que no se pudieron escribir en DLQ.
`otel-pipeline.opensearch.s3.dlqS3RecordsSuccess.count`	Número de registros que se escriben en DLQ.
`otel-pipeline.opensearch.s3.dlqS3RequestLatency.count`	Recuento de las mediciones de latencia de las solicitudes a la cola de cartas muertas de Amazon S3.
`otel-pipeline.opensearch.s3.dlqS3RequestLatency.sum`	Latencia total de todas las solicitudes a la cola de cartas muertas de Amazon S3
`otel-pipeline.opensearch.s3.dlqS3RequestSizeBytes.sum`	Tamaño total en bytes de todas las solicitudes realizadas a la cola de cartas muertas de Amazon S3.
`otel-pipeline.recordsProcessed.count`	Número total de registros procesados en la canalización, una métrica clave para el rendimiento general.
`otel-pipeline.opensearch.bulkRequestInvalidInputErrors.count`	Cuenta los errores en las solicitudes masivas OpenSearch debidos a una entrada no válida, algo crucial para supervisar la calidad de los datos y los problemas operativos.

Uso de una canalización OpenSearch de ingestión con Amazon Managed Service para Prometheus

Puedes usar Amazon Managed Service for Prometheus como destino de OpenSearch tu proceso de ingestión para almacenar las métricas en formato de series temporales. El receptor de Prometheus

te permite OpenTelemetry enviar métricas u otros datos de series temporales de tu canalización a un espacio de trabajo de Amazon Managed Service for Prometheus para su supervisión, alerta y análisis.

El complemento `prometheus` sink permite a las canalizaciones de OpenSearch ingestión escribir datos de métricas en los espacios de trabajo de Amazon Managed Service for Prometheus mediante el protocolo de escritura remota Prometheus. Esta integración le permite:

- Almacena datos de métricas de series temporales en Amazon Managed Service para Prometheus
- Supervisa las métricas y envía alertas con Amazon Managed Service for Prometheus y Amazon Managed Grafana
- Enrute las métricas a varios destinos simultáneamente (por ejemplo, OpenSearch y Amazon Managed Service para Prometheus)
- Procesa OpenTelemetry métricas de agentes externos o genera métricas dentro del proceso

Temas

- Requisitos previos
- Paso 1: configurar el rol de canalización
- Paso 2: crear la canalización
- Supervisión y solución de problemas
- Limitaciones
- Prácticas recomendadas

Requisitos previos

Antes de configurar el sumidero Prometheus, asegúrese de disponer de lo siguiente:

- Amazon Managed Service para el espacio de trabajo de Prometheus: cree un espacio de trabajo en el mismo canal de ingestión Región de AWS y en el Cuenta de AWS mismo. OpenSearch Para obtener instrucciones, consulte Creación de un espacio de trabajo en la Guía del usuario de Amazon Managed Service for Prometheus.
- Permisos de IAM: configure un rol de IAM con permisos para escribir en Amazon Managed Service for Prometheus. Para obtener más información, consulte the section called “Paso 1: configurar el rol de canalización”.

> **Note**
>
> Los espacios de trabajo de Amazon Managed Service para Prometheus deben utilizar claves gestionadas por el servicio. AWS AWS KMS Actualmente, AWS KMS las claves administradas por el cliente no son compatibles con los receptores de Amazon Managed Service for Prometheus en las canalizaciones de ingestión. OpenSearch

Paso 1: configurar el rol de canalización

El receptor de Prometheus hereda automáticamente los permisos de IAM del rol de canalización para la autenticación, por lo que no se requiere ninguna configuración de rol adicional (`sts_role_arn`como) en la configuración del receptor.

En el siguiente ejemplo de política se muestran los permisos necesarios para utilizar Amazon Managed Service for Prometheus como receptor:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AMPRemoteWrite",
      "Effect": "Allow",
      "Action": [
        "aps:RemoteWrite"
      ],
      "Resource": "arn:aws:aps:region:account-id:workspace/workspace-id"
    }
  ]
}
```

Reemplace los siguientes marcadores de posición:

- *`region`*: Tu Región de AWS (por ejemplo,) `us-east-1`
- *`account-id`*: Tu Cuenta de AWS carné de identidad
- *`workspace-id`*: tu ID de espacio de trabajo de Amazon Managed Service para Prometheus

Debe adjuntar estos permisos a su función de canalización.

Asegúrate de que tu rol en canalización tenga una relación de confianza que permita a OpenSearch Ingestion asumirlo:

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

Paso 2: crear la canalización

Una vez configurados los permisos, puede configurar una canalización de OpenSearch ingestión para utilizar Amazon Managed Service for Prometheus como receptor.

Configuración básica

El siguiente ejemplo muestra una configuración mínima del sumidero de Prometheus:

```
version: "2"
sink:
  - prometheus:
      url: "https://aps-workspaces.region.amazonaws.com/workspaces/workspace-id/api/v1/
remote_write"
      aws:
        region: "region"
```

Debe especificar la `url` opción en la configuración del receptor`prometheus`, que es el punto final de escritura remota de Amazon Managed Service for Prometheus. Para formatear la URL, localice el ID de su espacio de trabajo en la consola de Amazon Managed Service for Prometheus y cree la URL de la siguiente manera:. `https://aps-workspaces.region.amazonaws.com/workspaces/workspace-id/api/v1/remote_write`

Opciones de configuración

Utilice las siguientes opciones para configurar el comportamiento de procesamiento por lotes y descarga del sumidero Prometheus:

Opciones de configuración del fregadero Prometheus

Opción	Obligatorio	Tipo	Description (Descripción)
`max_events`	No	Entero	El número máximo de eventos que se pueden acumular antes de que lleguen a Prometheus. El valor predeterminado es `1000`.
`max_request_size`	No	Recuento de bytes	El tamaño máximo de la carga útil de la solicitud antes de vaciarla. El valor predeterminado es `1mb`.
`flush_interval`	No	Duración	El tiempo máximo de espera antes de que se produzcan eventos de descarga. El valor predeterminado es `10s`. El valor máximo permitido es. `60s`

Ejemplo de canalizaciones

Ejemplo 1: OpenTelemetry métricas de Amazon Managed Service para Prometheus

Esta canalización recibe OpenTelemetry las métricas de un agente externo y las escribe en 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"
```

Ejemplo 2: Doble receptor OpenSearch y Amazon Managed Service para Prometheus

Esta canalización redirige las métricas a OpenSearch Amazon Managed Service for Prometheus:

```
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"
```

Ejemplo 3: Métricas con filtrado

Esta canalización filtra las métricas antes de enviarlas a Amazon Managed Service para 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"
```

Puedes usar un plano preconfigurado de Amazon Managed Service for Prometheus para crear estas canalizaciones. Para obtener más información, consulte the section called “Uso de esquemas”.

Creación de una canalización con Amazon Managed Service para Prometheus sink

Uso de la consola AWS

1. Navegue hasta la consola OpenSearch de servicio.
2. Selecciona Pipelines en Ingestión.
3. Seleccione Create pipeline.
4. Seleccione Crear mediante un plano y elija OpenTelemetry las métricas para el plano de Amazon Prometheus.
5. Configure la canalización:
   - Introduce tu ID de espacio de trabajo de Amazon Managed Service para Prometheus
   - Especifique el ARN de la función de canalización
   - Configure los ajustes de origen y procesador según sea necesario
6. Revise y cree la canalización.

Uso de la AWS CLI

Cree un archivo de configuración de canalización (por ejemplo`amp-pipeline.yaml`) con la configuración que desee y, a continuación, ejecute:

```
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"
```

Supervisión y solución de problemas

CloudWatch métricas

Supervisa el rendimiento de tu canalización mediante CloudWatch métricas:

- `DocumentsWritten`: Número de métricas escritas correctamente en Amazon Managed Service para Prometheus
- `DocumentsWriteFailed`: Número de métricas que no se pudieron escribir
- `RequestLatency`: Latencia de las solicitudes de escritura remota

Problemas comunes

Problema: Pipeline no escribe a Amazon Managed Service para Prometheus

Soluciones:

- Comprueba que el ID del espacio de trabajo y la región de la URL sean correctos
- Asegúrese de que el rol de canalización tenga `aps:RemoteWrite` permiso
- Comprueba que el espacio de trabajo utilice claves administradas por el servicio AWS KMS
- Compruebe que la canalización y el espacio de trabajo estén en el mismo lugar Cuenta de AWS

Problema: errores de autenticación

Soluciones:

- Compruebe que la relación de confianza le `osis-pipelines.amazonaws.com` permita asumir la función de canalización
- Asegúrese de que el rol de canalización tenga el `aps:RemoteWrite` permiso necesario

Problema: alta latencia o limitación

Soluciones:

- Aumente las unidades de capacidad de la tubería
- Implemente el procesamiento por lotes en el procesador
- Consulta las cuotas de servicio de Amazon Managed Service para ver las cuotas de servicio de Prometheus

Limitaciones

Tenga en cuenta las siguientes limitaciones al configurar una canalización de OpenSearch ingestión para Amazon Managed Service for Prometheus:

- Los espacios de trabajo de Amazon Managed Service para Prometheus deben utilizar claves gestionadas por el servicio. AWS AWS KMS En la actualidad, no se admiten las claves gestionadas por el cliente AWS KMS .
- La canalización y el espacio de trabajo de Amazon Managed Service for Prometheus deben estar en el mismo lugar. Cuenta de AWS

Prácticas recomendadas

- Utilice la misma función de IAM: el sumidero de Prometheus utiliza automáticamente la función de canalización. Si se utilizan otros sumideros, asegúrese de que la función de `sts_role_arn` canalización sea la misma
- Supervise las métricas: configure CloudWatch alarmas para las escrituras fallidas y la alta latencia
- Implemente el filtrado: utilice procesadores para filtrar las métricas innecesarias antes de enviarlas a Amazon Managed Service for Prometheus

- Capacidad del tamaño adecuado: comience con una capacidad mínima y escale en función del volumen de las métricas
- Planos de uso: aproveche los planos preconfigurados para casos de uso comunes

Uso de una canalización OpenSearch de ingestión con Kafka

Puedes usar el complemento de Kafka para transmitir datos desde clústeres de Kafka autogestionados a dominios de Amazon OpenSearch Service y OpenSearch colecciones sin servidor. OpenSearch La ingestión admite conexiones desde clústeres de Kafka configurados con redes públicas o privadas (VPC). En este tema, se describen los requisitos previos y los pasos para configurar una canalización de ingesta, incluida la configuración de los ajustes de red y los métodos de autenticación, como el TLS mutuo (mTLS), SASL/SCRAM o IAM.

Migración de datos desde clústeres públicos de Kafka

Puede usar las canalizaciones OpenSearch de ingestión para migrar datos desde un clúster de Kafka público y autogestionado, lo que significa que el nombre DNS del dominio se puede resolver públicamente. Para ello, configura una canalización de OpenSearch ingestión con Kafka autogestionado como origen y OpenSearch Service o Serverless como destino. OpenSearch Esto procesa los datos de streaming desde un clúster de origen autogestionado hasta un dominio o colección de destino AWS gestionado.

Requisitos previos

Antes de crear tu canalización OpenSearch de ingestión, lleva a cabo los siguientes pasos:

1. Cree un clúster de Kafka autogestionado con una configuración de red pública. El clúster debe contener los datos que desea incorporar a Service. OpenSearch
2. Cree un dominio OpenSearch de servicio o una colección OpenSearch sin servidor a la que desee migrar los datos. Para obtener más información, consulte the section called “Creación de dominios OpenSearch de servicio” y the section called “Creación de colecciones”.
3. Configure la autenticación en su clúster autogestionado con. AWS Secrets Manager Siga los pasos de Rotate AWS Secrets Manager secrets para activar la rotación de secretos.
4. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster autogestionado en su dominio o colección.

El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización, que creará en el paso siguiente, escriba datos en un dominio. Asegúrese de actualizar el `resource` con su propio 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"
      ]
    }
  ]
}
```

Para crear un rol de IAM con los permisos correctos para acceder a los datos de escritura de la colección o dominio, consulte the section called “Configuración de roles y usuarios”.

Paso 1: configurar el rol de canalización

Una vez establecidos los requisitos previos de la canalización de Kafka, configura el rol de canalización que quieres usar en la configuración de la canalización y añade permiso para escribir en un dominio de OpenSearch servicio o en una colección OpenSearch sin servidor, así como permiso para leer los secretos de Secrets Manager.

Paso 2: crear la canalización

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, en la que se especifica a Kafka como fuente.

Puede especificar varios dominios OpenSearch de servicio como destinos de sus datos. Esta capacidad permite el enrutamiento condicional o la replicación de los datos entrantes en varios dominios OpenSearch de servicio.

También puede migrar datos de un clúster de Confluent Kafka de origen a una colección de VPC OpenSearch sin servidor. Asegúrese de proporcionar una política de acceso a la red en la configuración de la canalización. Puede utilizar un registro de esquemas de Confluent para definir un esquema de 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"
```

Puede utilizar un esquema preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Migración de datos desde clústeres de Kafka en una VPC

También puede usar las canalizaciones OpenSearch de ingestión para migrar datos desde un clúster Kafka autogestionado que se ejecute en una VPC. Para ello, configura una canalización de OpenSearch ingestión con Kafka autogestionado como origen y Service o Serverless como destino. OpenSearch OpenSearch Esto procesa los datos de streaming desde un clúster de origen autogestionado hasta un dominio o colección de destino AWS gestionado.

Requisitos previos

Antes de crear tu canalización OpenSearch de ingestión, lleva a cabo los siguientes pasos:

1. Cree un clúster Kafka autogestionado con una configuración de red de VPC que contenga los datos que desee incorporar al servicio. OpenSearch
2. Cree un dominio OpenSearch de servicio o una colección OpenSearch sin servidor a la que desee migrar los datos. Para obtener más información, consulte Creación de dominios OpenSearch de servicio y Creación de colecciones.
3. Configure la autenticación en su clúster autogestionado con AWS Secrets Manager. Siga los pasos de Rotate AWS Secrets Manager secrets para activar la rotación de secretos.
4. Obtenga el ID de la VPC que tiene acceso a Kafka autogestionado. Elija el CIDR de VPC que utilizará Ingestion. OpenSearch

> **Note**
>
> Si utilizas el Consola de administración de AWS para crear tu canalización, también debes adjuntar tu canalización de OpenSearch ingestión a tu VPC para poder usar Kafka autogestionado. Para ello, busque la sección Configuración de la red, seleccione la casilla Adjuntar a la VPC y elija el CIDR de una de las opciones predeterminadas proporcionadas o seleccione una opción propia. Puede usar cualquier CIDR de un espacio de direcciones privado, tal como se define en RFC 1918 Best Current Practice.

Para proporcionar un CIDR personalizado, seleccione Otros en el menú desplegable. Para evitar una colisión en las direcciones IP entre la OpenSearch ingesta y la autogestión OpenSearch, asegúrese de que el CIDR de la OpenSearch VPC autogestionada sea diferente del CIDR de la ingestión. OpenSearch

5. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster autogestionado en su dominio o colección.

   El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización, que creará en el paso siguiente, escriba datos en un dominio. Asegúrese de actualizar el `resource` con su propio 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"
         ]
       }
     ]
   }
   ```

   Para crear un rol de IAM con los permisos correctos para acceder a los datos de escritura de la colección o dominio, consulte the section called “Configuración de roles y usuarios”.

Paso 1: configurar el rol de canalización

Una vez configurados los requisitos previos de la canalización, configure el rol de la canalización que desee usar en la configuración de la canalización y agregue los siguientes permisos al rol:

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"
                    }
            }
        }
    ]
}
```

Debe proporcionar los permisos de Amazon EC2 anteriores en la función de IAM que utilice para crear la canalización de OpenSearch ingestión, ya que la canalización utiliza estos permisos para crear y eliminar una interfaz de red en la VPC. La canalización solo puede acceder al clúster de Kafka a través de esta interfaz de red.

Paso 2: crear la canalización

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, en la que se especifica a Kafka como fuente.

Puede especificar varios dominios OpenSearch de servicio como destinos de sus datos. Esta capacidad permite el enrutamiento condicional o la replicación de los datos entrantes en varios dominios OpenSearch de servicio.

También puede migrar datos de un clúster de Confluent Kafka de origen a una colección de VPC OpenSearch sin servidor. Asegúrese de proporcionar una política de acceso a la red en la configuración de la canalización. Puede utilizar un registro de esquemas de Confluent para definir un esquema de 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"
```

Puede utilizar un esquema preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Migración de datos desde OpenSearch clústeres autogestionados mediante Amazon Ingestion OpenSearch

Puede utilizar una canalización de Amazon OpenSearch Ingestion autogestionada OpenSearch o con Elasticsearch para migrar los datos a los dominios de Amazon OpenSearch Service y a las colecciones sin servidor. OpenSearch OpenSearch Ingestion admite configuraciones de redes públicas y privadas para la migración de datos de Elasticsearch y de Automanagement. OpenSearch

Migración desde clústeres públicos OpenSearch

Puedes usar las canalizaciones de OpenSearch ingestión para migrar datos desde un clúster autogestionado OpenSearch o de Elasticsearch con una configuración pública, lo que significa que el nombre DNS del dominio se puede resolver públicamente. Para ello, configura una canalización de OpenSearch ingestión con la autogestión OpenSearch o Elasticsearch como origen y Service o Serverless como destino. OpenSearch OpenSearch Esto permite migrar de forma eficaz los datos de un clúster de origen autogestionado a un dominio o colección de destino gestionado. AWS

Requisitos previos

Antes de crear su canalización de OpenSearch ingestión, lleve a cabo los siguientes pasos:

1. Cree un clúster autogestionado OpenSearch o de Elastisearch que contenga los datos que desee migrar y configure un nombre de DNS público. Para obtener instrucciones, consulte Crear un clúster en la OpenSearch documentación.
2. Cree un dominio OpenSearch de servicio o una colección OpenSearch sin servidor a la que desee migrar los datos. Para obtener más información, consulte the section called “Creación de dominios OpenSearch de servicio” y the section called “Creación de colecciones”.
3. Configure la autenticación en su clúster autogestionado con. AWS Secrets Manager Siga los pasos de Rotate AWS Secrets Manager secrets para activar la rotación de secretos.
4. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster autogestionado en su dominio o colección.

   El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización, que creará en el paso siguiente, escriba datos en un dominio. Asegúrese de actualizar el `resource` con su propio 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"
      ]
    }
  ]
}
```

Para crear un rol de IAM con los permisos correctos para acceder a los datos de escritura de la colección o dominio, consulte the section called “Configuración de roles y usuarios”.

Paso 1: configurar el rol de canalización

Una vez establecidos los requisitos previos de la OpenSearch canalización, configura el rol de canalización que deseas usar en la configuración de la canalización y añade permiso para escribir en un dominio de OpenSearch servicio o en una colección OpenSearch sin servidor, así como permiso para leer los secretos de Secrets Manager.

Paso 2: crear la canalización

A continuación, puede configurar una canalización OpenSearch de ingestión como la siguiente, que se especifica OpenSearch como fuente.

Puede especificar varios dominios OpenSearch de servicio como destinos de sus datos. Esta capacidad permite el enrutamiento condicional o la replicación de los datos entrantes en varios dominios OpenSearch de servicio.

También puedes migrar datos de una fuente OpenSearch o un clúster de Elasticsearch a una colección de VPC OpenSearch sin servidor. Asegúrese de proporcionar una política de acceso a la red en la configuración de la canalización.

```
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
```

Puede utilizar un esquema preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Migración de datos desde OpenSearch clústeres en una VPC

También puede usar las canalizaciones de OpenSearch ingestión para migrar datos desde un clúster autogestionado OpenSearch o de Elasticsearch que se ejecute en una VPC. Para ello, configura una canalización de OpenSearch ingestión con la autogestión OpenSearch o Elasticsearch como origen y con Service o Serverless como destino. OpenSearch OpenSearch Esto permite migrar de forma eficaz los datos de un clúster de origen autogestionado a un dominio o colección de destino gestionado. AWS

Requisitos previos

Antes de crear su canalización de OpenSearch ingestión, lleve a cabo los siguientes pasos:

1. Cree un clúster de Elastisearch OpenSearch o autogestionado con una configuración de red de VPC que contenga los datos que desee migrar.
2. Cree un dominio OpenSearch de servicio o una colección OpenSearch sin servidor a la que desee migrar los datos. Para obtener más información, consulte Creación de dominios OpenSearch de servicio y Creación de colecciones.
3. Configure la autenticación en su clúster autogestionado con AWS Secrets Manager. Siga los pasos de Rotate AWS Secrets Manager secrets para activar la rotación de secretos.
4. Obtenga el ID de la VPC que tiene acceso a la búsqueda autogestionada OpenSearch o a Elasticsearch. Elija el CIDR de VPC que utilizará Ingestion. OpenSearch

> **Note**
>
> Si utilizas la Consola de administración de AWS para crear tu canalización, también debes adjuntar la canalización de OpenSearch ingestión a tu VPC para poder usar Elasticsearch o OpenSearch autogestionada. Para ello, busque la sección Opciones de la red de origen, seleccione la casilla Attach to VPC (Adjuntar a la VPC) y elija el CIDR de una de las opciones predeterminadas proporcionadas. Puede usar cualquier CIDR de un espacio de direcciones privado, tal como se define en RFC 1918 Best Current Practice.

Para proporcionar un CIDR personalizado, seleccione Otros en el menú desplegable. Para evitar una colisión en las direcciones IP entre la OpenSearch ingesta y la autogestión OpenSearch, asegúrese de que el CIDR de la OpenSearch VPC autogestionada sea diferente del CIDR de la ingestión. OpenSearch

5. Adjunte una política basada en recursos al dominio o una política de acceso a datos a la colección. Estas políticas de acceso permiten a OpenSearch Ingestion escribir datos de su clúster autogestionado en su dominio o colección.

   El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización, que creará en el paso siguiente, escriba datos en un dominio. Asegúrese de actualizar el `resource` con su propio 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"
         ]
       }
     ]
   }
   ```

   Para crear un rol de IAM con los permisos correctos para acceder a los datos de escritura de la colección o dominio, consulte the section called “Configuración de roles y usuarios”.

Paso 1: configurar el rol de canalización

Una vez configurados los requisitos previos de la canalización, configure el rol de la canalización que desee usar en la configuración de la canalización y agregue los siguientes permisos al rol:

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"
                    }
            }
        }
    ]
}
```

Debe proporcionar los permisos de Amazon EC2 anteriores en la función de IAM que utilice para crear la canalización de OpenSearch ingestión, ya que la canalización utiliza estos permisos para crear y eliminar una interfaz de red en la VPC. La canalización solo puede acceder al OpenSearch clúster a través de esta interfaz de red.

Paso 2: crear la canalización

A continuación, puede configurar una canalización de OpenSearch ingestión como la siguiente, que se especifica OpenSearch como fuente.

Puede especificar varios dominios OpenSearch de servicio como destinos de sus datos. Esta capacidad permite el enrutamiento condicional o la replicación de los datos entrantes en varios dominios OpenSearch de servicio.

También puedes migrar datos de una fuente OpenSearch o un clúster de Elasticsearch a una colección de VPC OpenSearch sin servidor. Asegúrese de proporcionar una política de acceso a la red en la configuración de la canalización.

```
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
```

Puede utilizar un esquema preconfigurado para crear esta canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

Utilice una canalización OpenSearch de ingestión con Amazon Kinesis Data Streams

Utilice una canalización de OpenSearch ingestión con Amazon Kinesis Data Streams para incorporar datos de registros de transmisiones de varias transmisiones a dominios y colecciones de OpenSearch Amazon Service. La canalización de OpenSearch ingestión incorpora la infraestructura de ingesta de streaming para proporcionar una forma a gran escala y baja latencia de ingerir continuamente los registros de streaming de Kinesis.

Temas

- Amazon Kinesis Data Streams como origen
- Amazon Kinesis Data Streams como origen

Amazon Kinesis Data Streams como origen

Con el siguiente procedimiento, aprenderá a configurar una canalización de OpenSearch ingestión que utilice Amazon Kinesis Data Streams como fuente de datos. En esta sección se describen los requisitos previos necesarios, como la creación de un dominio de OpenSearch servicio o una colección OpenSearch sin servidor, y se explican los pasos para configurar la función de canalización y crear la canalización.

Requisitos previos

Para configurar su canalización, necesitará uno o varios Kinesis Data Streams activos. Estas transmisiones deben recibir registros o estar listas para recibir registros de otros orígenes. Para obtener más información, consulte Descripción general de la OpenSearch ingestión.

Configuración de su canalización

1. Cree un dominio OpenSearch de servicio o una colección OpenSearch sin servidor

   Para crear un dominio o una colección, consulta Cómo empezar a usar OpenSearch Ingestion.

   Para crear un rol de IAM con los permisos correctos para acceder a los datos de escritura de la colección o dominio, consulte Políticas basadas en recursos.

2. Configure el rol de canalización con permisos

   Configure el rol de canalización que desee usar en la configuración de canalización y agregue los siguientes permisos. Sustituya *placeholder values* por su propia información.

   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"
            ]
        }
    ]
}
```

   Si el cifrado del lado del servidor está habilitado en las transmisiones, la siguiente política de AWS KMS permite descifrar los registros. Sustituya *placeholder values* por su propia información.

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"
        }
    ]
}
```

Para que una canalización escriba datos en un dominio, el dominio debe tener una política de acceso a nivel de dominio que permita al rol de canalización sts_role_arn acceder a ellos.

El siguiente ejemplo de política de acceso al dominio permite que el rol de canalización que creó en el paso anterior, (`pipeline-role`), escriba datos en el dominio `ingestion-domain`. Sustituya *`placeholder values`* por su propia información.

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::your-account-id:role/pipeline-role"
      },
      "Action": ["es:DescribeDomain", "es:ESHttp*"],
      "Resource": "arn:aws:es:Región de AWS:account-id:domain/domain-name/*"
    }
  ]
}
```

3. Creación de la canalización

Configure una canalización OpenSearch de ingestión especificando K inesis-data-streams como fuente. Puede encontrar un plano listo para usar en la consola de OpenSearch ingestión para crear dicha canalización. (Opcional) Para crear la canalización mediante el AWS CLI, puede utilizar un plano denominado "». **`AWS-KinesisDataStreamsPipeline`** Sustituya *`placeholder values`* por su propia información.

```
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 Región de 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 Región de 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"
```

Opciones de configuración

Para ver las opciones de configuración de Kinesis, consulte las opciones de configuración en la OpenSearchdocumentación.

Atributos de metadatos disponibles

- stream_name: nombre de las transmisiones de datos de Kinesis desde las que se ingirió el registro
- partition_key: clave de partición del registro de Kinesis Data Streams que se está ingiriendo
- sequence_number: número de secuencia del registro de Kinesis Data Streams que se está ingiriendo
- sub_sequence_number: número de subsecuencia del registro de Kinesis Data Streams que se está ingiriendo

4. (Opcional) Configure las unidades de cómputo recomendadas (OCUs) para la canalización de Kinesis Data Streams

   También se puede configurar una canalización de origen de OpenSearch Kinesis Data Streams para ingerir registros de transmisión de más de una transmisión. Para una ingesta más rápida, le recomendamos que agregue una unidad de cómputo adicional por cada nueva transmisión que se agregue.

Coherencia de datos

OpenSearch La ingestión permite el end-to-end reconocimiento para garantizar la durabilidad de los datos. Cuando la canalización lee los registros de transmisión de Kinesis, distribuye dinámicamente el trabajo de lectura de los registros de transmisión en función de los particiones asociadas a las transmisiones. Pipeline comprobará automáticamente las transmisiones de puntos cuando reciba un acuse de recibo tras haber ingerido todos los registros del dominio o la colección. OpenSearch Esto evitará el procesamiento duplicado de los registros de las transmisiones.

Para crear el índice en función del nombre del flujo, defina el índice en la sección del receptor de opensearch como "index_${getMetadata(\"stream_name\")}".

Amazon Kinesis Data Streams como origen

Puede conceder acceso a todas las cuentas con Amazon Kinesis Data Streams para OpenSearch que las canalizaciones de ingestión puedan acceder a Kinesis Data Streams de otra cuenta como fuente. Complete los siguientes pasos para habilitar el acceso entre cuentas:

Configuración del acceso entre cuentas

1. Establezca la política de recursos en la cuenta que tiene la transmisión de Kinesis

   Sustituya *placeholder values* por su propia información.

   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. (Opcional) Configure una política para consumidores y recursos para consumidores

   Este paso es opcional y solo será necesario si planea utilizar la estrategia Enhanced Fanout Consumer para leer los registros de las transmisiones. Para más información, consulte Develop enhanced fan-out consumers with dedicated throughput.

   a. Configurar consumidor

      Para reutilizar un consumidor existente, puede omitir este paso. Para obtener más información, consulte la RegisterStreamConsumerreferencia de la API de Amazon Kinesis Data Streams.

      En el siguiente ejemplo de comando CLI, sustituya el *placeholder values* por su propia información.

Example Ejemplo de comando de la CLI:

```
aws kinesis register-stream-consumer \
--stream-arn "arn:aws:kinesis:Región de AWS:account-id:stream/stream-name" \
--consumer-name consumer-name
```

b. Configure la política de recursos para el consumidor

En la siguiente declaración, *placeholder values* sustitúyala por su propia información.

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. Configuración de canalización

Para la ingesta entre cuentas, agregue los siguientes atributos en `kinesis_data_streams` para cada transmisión:

- `stream_arn`: el ARN de la transmisión que pertenece a la cuenta en la que existe la transmisión.
- `consumer_arn`: se trata de un atributo opcional y debe especificarse si se opta por la estrategia de consumo de distribución ramificada mejorada de forma predeterminada. Especifique el nombre real del consumidor para este campo. Sustituya *`placeholder values`* por su propia información.

```
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 Región de 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 Región de 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 Región de 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. OSI Pipeline Role Kinesis Data Streams

   a. Política de IAM

      Agregue la siguiente política al rol de canalización. Sustituya *placeholder values* por su propia información.

      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. Política de confianza

Para poder ingerir datos de la cuenta de transmisión, tendrá que establecer una relación de confianza entre la función de ingesta de canalización y la cuenta de transmisión. Agregue lo siguiente al rol de canalización. Sustituya *placeholder values* por su propia información.

JSON

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

Siguientes pasos

Tras exportar los datos a una canalización, puede consultarlos desde el dominio de OpenSearch servicio que está configurado como receptor de la canalización. Los siguientes recursos pueden ayudarle a comenzar:

- Observabilidad
- the section called “Rastros”
- Observabilidad

Uso de una canalización OpenSearch de ingestión con AWS Lambda

Utilice el AWS Lambda procesador para enriquecer los datos de cualquier fuente o destino compatible con OpenSearch Ingestion mediante un código personalizado. Con el procesador de Lambda, puede aplicar sus propias transformaciones o enriquecimientos de datos y, a continuación, devolver los eventos procesados a la canalización para su posterior procesamiento. Este procesador permite un procesamiento de datos personalizado y le brinda un control total sobre la forma en que se manipulan los datos antes de que se procesen.

> **Note**
>
> El límite de tamaño de carga útil para un solo evento procesado por un procesador de Lambda es de 5 MB. Además, el procesador de Lambda solo admite respuestas en formato de matriz JSON.

Requisitos previos

Antes de crear una canalización con un procesador de Lambda, cree los siguientes recursos:

- Una AWS Lambda función que enriquece y transforma los datos de origen. Para obtener instrucciones, consulte Crear la función de Lambda.
- Un dominio OpenSearch de servicio o una colección OpenSearch sin servidor que servirá de colector. Para obtener más información, consulte the section called “Creación de dominios OpenSearch de servicio” y the section called “Creación de colecciones”.
- Un rol de canalización que incluye permisos para escribir en el dominio o receptor de la colección. Para obtener más información, consulte the section called “Rol de canalización”.

  El rol de canalización también necesita una política de permisos adjunta que le permita invocar la función de Lambda especificada en la configuración de canalización. Por ejemplo:

  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"

        }
    ]
}
```

Creación de una canalización

Para usarlo AWS Lambda como procesador, configure una canalización de OpenSearch ingestión y especifíquela `aws_lambda` como procesador. También puede usar el esquema de enriquecimiento de AWS Lambda personalizado para crear la canalización. Para obtener más información, consulte the section called “Uso de esquemas”.

El siguiente ejemplo de canalización recibe datos de una fuente HTTP, los enriquece con un procesador de datos y el AWS Lambda procesador e ingiere los datos procesados en 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 siguiente AWS Lambda función de ejemplo transforma los datos entrantes añadiendo un nuevo par clave-valor (`"transformed": "true"`) a cada elemento de la matriz de eventos proporcionada y, a continuación, devuelve la versión modificada.

```
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
```

Agrupación en lotes

Las canalizaciones envían eventos por lotes al procesador de Lambda y ajustan dinámicamente el tamaño del lote para garantizar que se mantenga por debajo del límite de 5 MB.

El siguiente es un ejemplo de lote de canalización:

```
batch:
    key_name: "events"

input_arrary = event.get('events', [])
```

> **Note**
>
> Al crear una canalización, asegúrese de que la opción `key_name` de la configuración del procesador de Lambda coincida con la clave de evento del controlador de Lambda.

Filtrado condicional

El filtrado condicional le permite controlar cuándo el AWS Lambda procesador invoca la función Lambda en función de condiciones específicas de los datos de eventos. Esto resulta especialmente útil cuando desea procesar de forma selectiva ciertos tipos de eventos e ignorar otros.

El siguiente ejemplo de configuración utiliza el filtrado condicional:

```
processors:
  - aws_lambda:
      function_name: "my-lambda-function"
      aws:
        region: "region"
      lambda_when: "/sourceIp == 10.10.10.10"
```

Migración de datos entre dominios y colecciones mediante Amazon OpenSearch Ingestion

Puedes usar las canalizaciones OpenSearch de ingestión para migrar datos entre dominios de Amazon OpenSearch Service o colecciones de VPC OpenSearch sin servidor. Para ello, se configura una canalización en la que se configura un dominio o colección como origen y otro dominio o colección como receptor. Esto migra eficazmente los datos de un dominio o colección a otro.

Para migrar los datos, debe tener los siguientes recursos:

- Un dominio de OpenSearch servicio de origen o una colección de VPC OpenSearch sin servidor. Este dominio o colección contiene los datos que quiere migrar. Si utilizas un dominio, debe ejecutar la OpenSearch versión 1.0 o posterior, o Elasticsearch, la versión 7.4 o posterior. El dominio también debe tener una política de acceso que conceda los permisos adecuados a su rol de canalización.
- Un dominio o colección de VPC independiente al que desee migrar sus datos. Este dominio o colección actuará como receptor de la canalización.
- Un rol de canalización que OpenSearch Ingestion utilizará para leer y escribir en tu colección o dominio. Debe incluir el Nombre de recurso de Amazon (ARN) de este rol en la configuración de la canalización. Para obtener más información, consulte los siguientes recursos:
  - the section called “Concesión de acceso a dominios para canalizaciones”
  - the section called “Concesión de acceso a colecciones para canalizaciones”

Temas

- Limitaciones
- OpenSearch El servicio como fuente
- Especificar varios sumideros OpenSearch de dominio de servicio
- Migración de datos a una colección de OpenSearch VPC sin servidor

Limitaciones

Al designar dominios de OpenSearch servicio o colecciones OpenSearch sin servidor como receptores, se aplican las siguientes limitaciones:

- Una canalización no puede escribir en más de un dominio de VPC.
- Solo puede migrar datos hacia o desde colecciones OpenSearch sin servidor que utilicen el acceso a la VPC. No se admiten las colecciones públicas.
- No puede especificar una combinación de dominios públicos y VPC en una única configuración de la canalización.
- Puede incluir un máximo de 20 receptores que no sean canalizaciones en una sola configuración de canalización.
- Puedes especificar sumideros de un máximo de tres diferentes Regiones de AWS en una única configuración de canalización.
- Si alguno de los receptores está inactivo durante demasiado tiempo o no cuenta con la capacidad suficiente para recibir los datos entrantes, una canalización con varios receptores puede experimentar una reducción en la velocidad de procesamiento con el tiempo.

OpenSearch El servicio como fuente

El dominio o la colección que especifique como origen es desde donde se migran los datos.

Creación de un rol de canalización en IAM

Para crear tu canalización de OpenSearch ingestión, primero debes crear una función de canalización que conceda acceso de lectura y escritura entre dominios o colecciones. Para ello, siga estos pasos:

1. Cree una nueva política de permisos en IAM para adjuntarla al rol de la canalización. Asegúrese de conceder los permisos de lectura desde el origen y de escritura en el receptor. Para obtener

más información sobre cómo configurar los permisos de canalización de IAM para los dominios OpenSearch de servicio, consulte the section called “Concesión de acceso a dominios para canalizaciones” y. the section called “Concesión de acceso a colecciones para canalizaciones”

2. Especifique los siguientes permisos en el rol de la canalización para leer desde el origen:

   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"
         ]
      }
```

```
    ]
}
```

Creación de una canalización

Después de asociar la política a la función de canalización, utilice el plan de AWSOpenSearchDataMigrationPipelinemigración para crear la canalización. Este esquema incluye una configuración predeterminada para migrar datos entre dominios o colecciones OpenSearch de servicios. Para obtener más información, consulte the section called “Uso de esquemas”.

> **Note**
>
> OpenSearch Ingestion utiliza la versión y la distribución del dominio de origen para determinar qué mecanismo utilizar para la migración. Algunas versiones admiten `point_in_time` esta opción. OpenSearch Serverless usa la `search_after` opción porque no admite `point_in_time` o`scroll`.

Es posible que se estén creando nuevos índices durante el proceso de migración o que los documentos se estén actualizando mientras la migración está en curso. Por este motivo, es posible que deba escanear uno o varios escaneos de los datos del índice de su dominio para recoger datos nuevos o actualizados.

Especifique el número de escaneos que se van a ejecutar el configurar `index_read_count` y `interval` en la canalización. El siguiente de ejemplo muestra cómo realizar varios escaneos:

```
scheduling:
    interval: "PT2H"
    index_read_count: 3
    start_time: "2023-06-02T22:01:30.00Z"
```

OpenSearch Ingestion usa la siguiente configuración para garantizar que los datos se escriban en el mismo índice y mantengan el mismo ID de documento:

```
index: "${getMetadata(\"opensearch-index\")}"
document_id: "${getMetadata(\"opensearch-document_id\")}"
```

Especificar varios sumideros OpenSearch de dominio de servicio

Puede especificar varios dominios de OpenSearch servicio público como destinos de sus datos. Puede utilizar esta capacidad para realizar un enrutamiento condicional o replicar los datos entrantes en varios dominios OpenSearch de servicio. Puede especificar hasta 10 dominios de OpenSearch servicio público diferentes como receptores.

En el siguiente ejemplo, los datos entrantes se enrutan condicionalmente a diferentes OpenSearch dominios de servicio:

```
...
  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
```

Migración de datos a una colección de OpenSearch VPC sin servidor

Puede usar OpenSearch Ingestión para migrar datos de un dominio de OpenSearch servicio de origen o de una colección OpenSearch sin servidor a un receptor de recopilación de VPC. Debe proporcionar una política de acceso a la red en la configuración de la canalización. Para obtener más información sobre la ingesta de datos en colecciones de OpenSearch VPC sin servidor, consulte. the section called “Tutorial: incorporar datos en una colección”

Migración de datos a una colección de VPC

1. Cree una colección sin servidor OpenSearch . Para obtener instrucciones, consulte the section called “Tutorial: incorporar datos en una colección”.
2. Cree una política de red para la colección que especifique el acceso a la VPC tanto para el punto de conexión de la colección como para el punto de conexión de Dashboards. Para obtener instrucciones, consulte the section called “Acceso a la red”.
3. Cree el rol de canalización si todavía no tiene uno. Para obtener instrucciones, consulte the section called “Rol de canalización”.
4. Cree la canalización. Para obtener instrucciones, consulte the section called “Uso de esquemas”.

Uso del AWS SDKs para interactuar con Amazon OpenSearch Ingestion

En esta sección se incluye un ejemplo de cómo utilizarla AWS SDKs para interactuar con Amazon OpenSearch Ingestion. El ejemplo de código muestra cómo crear un dominio y una canalización y, a continuación, incorporar datos a la canalización.

Temas

- Python

Python

En el siguiente script de ejemplo, se utiliza AWS SDK para Python (Boto3) para crear un rol de canalización de IAM, un dominio en el que escribir datos y una canalización a través de la cual se incorporan datos. A continuación, ingrese un archivo de registro de muestra en la canalización mediante la biblioteca HTTP `requests`.

Ejecute los siguientes comandos para asegurarse de que dispone de todas las dependencias necesarias:

```
pip install boto3
pip install botocore
pip install requests
pip install requests-auth-aws-sigv4
```

Dentro del script, reemplace todas las instancias de *account-id* por su ID de Cuenta de 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()
```

Seguridad en Amazon OpenSearch Ingestion

La seguridad en la nube AWS es la máxima prioridad. Como AWS cliente, usted se beneficia de una arquitectura de centro de datos y red diseñada para cumplir con los requisitos de las organizaciones más sensibles a la seguridad.

La seguridad es una responsabilidad compartida entre usted AWS y usted. El modelo de responsabilidad compartida la describe como seguridad de la nube y seguridad en la nube:

- Seguridad de la nube: AWS es responsable de proteger la infraestructura que ejecuta AWS los servicios en la AWS nube. AWS también le proporciona servicios que puede utilizar de forma segura. Auditores independientes prueban y verifican periódicamente la eficacia de nuestra seguridad en el marco de los programas de conformidad de AWS.
- Seguridad en la nube: su responsabilidad viene determinada por el AWS servicio que utilice. También es responsable de otros factores, incluida la confidencialidad de los datos, los requisitos de la empresa y la legislación y la normativa aplicables.

Esta documentación le ayuda a entender cómo aplicar el modelo de responsabilidad compartida al utilizar OpenSearch Ingestion. Los siguientes temas muestran cómo configurar OpenSearch Ingestion para cumplir sus objetivos de seguridad y conformidad. También aprenderá a utilizar otros AWS servicios que le ayudan a supervisar y proteger sus recursos de OpenSearch Ingestion.

Temas

- Configuración del acceso a la VPC para las canalizaciones de Amazon Ingestion OpenSearch
- Configuración de canalizaciones OpenSearch de ingestión para la ingestión entre cuentas
- Identity and Access Management para Amazon OpenSearch Ingestion
- Registro de llamadas a la API OpenSearch Amazon Ingestion mediante AWS CloudTrail
- API OpenSearch de Amazon Ingestion y puntos de enlace de interfaz ()AWS PrivateLink

Configuración del acceso a la VPC para las canalizaciones de Amazon Ingestion OpenSearch

Puede acceder a sus canalizaciones de Amazon OpenSearch Ingestion mediante un punto de enlace de VPC de interfaz. Una VPC es una red virtual dedicada a usted. Cuenta de AWS Está aislada de forma lógica de otras redes virtuales de la AWS nube. El acceso a una canalización a través de un punto final de la VPC permite una comunicación segura entre OpenSearch Ingestion y otros servicios de la VPC sin necesidad de una puerta de enlace a Internet, un dispositivo NAT o una conexión VPN. Todo el tráfico permanece seguro en la nube. AWS

OpenSearch Ingestion establece esta conexión privada mediante la creación de un punto final de interfaz, con la tecnología de AWS PrivateLink. Crearemos una interfaz de red de punto de conexión en cada subred que especifique durante la creación de la canalización. Se trata de interfaces de red administradas por el solicitante que sirven como punto de entrada para el tráfico destinado a

la canalización de ingestión. OpenSearch También puede elegir crear y administrar los puntos de conexión de la interfaz.

El uso de una VPC le permite hacer cumplir el flujo de datos a través de sus canalizaciones de OpenSearch ingestión dentro de los límites de la VPC, en lugar de hacerlo a través de la Internet pública. Las canalizaciones que no están dentro de una VPC envían y reciben datos a través de puntos de conexión públicos e Internet.

Una canalización con acceso a VPC puede escribir en dominios públicos o de OpenSearch servicio de VPC y en colecciones públicas o sin servidor de VPC. OpenSearch

Temas

- Consideraciones
- Limitaciones
- Requisitos previos
- Configuración del acceso mediante VPC para una canalización
- Puntos de conexión de VPC autogestionados
- Roles vinculados a servicios para el acceso mediante VPC

Consideraciones

Tenga en cuenta lo siguiente cuando configure el acceso mediante VPC para una canalización.

- No es necesario que una canalización esté en la misma VPC que su receptor. Tampoco es necesario establecer una conexión entre ambos. VPCs OpenSearch Ingestion se encarga de conectarlos por ti.
- Solo puede especificar una VPC para la canalización.
- A diferencia de las canalizaciones públicas, una canalización de VPC debe estar en la misma Región de AWS que el receptor de dominios o colecciones en el que se escribe. Puede configurar una fuente S3 para la canalización a fin de escribir entre regiones.
- Puede elegir implementar una canalización en una, dos o tres subredes de la VPC. Las subredes se distribuyen en las mismas zonas de disponibilidad en las que están desplegadas las unidades de OpenSearch cómputo de ingesta (OCUs).
- Si solo implementa una canalización en una subred y la zona de disponibilidad deja de funcionar, no podrá incorporar datos. Para garantizar una alta disponibilidad, le recomendamos que configure las canalizaciones con dos o tres subredes.

- La especificación de grupos de seguridad es opcional. Si no proporciona un grupo de seguridad, OpenSearch Ingestion utiliza el grupo de seguridad predeterminado que se especifica en la VPC.

Limitaciones

Las canalizaciones con acceso a la VPC presentan las siguientes limitaciones.

- No puede cambiar la configuración de red de una canalización después de crearla. Si lanza una canalización dentro de una VPC, no podrá cambiarla posteriormente a un punto de conexión público y viceversa.
- Puede lanzar la canalización dentro de un punto de conexión de VPC de la interfaz o un punto de conexión público, pero no es posible hacer las dos cosas. Se debe elegir una de ellas al crear una canalización.
- Después de aprovisionar una canalización con acceso a la VPC, no se puede mover a otra VPC, y no puede cambiar la configuración de sus subredes y grupos de seguridad.
- Si su canalización escribe en un receptor de dominios o colecciones que usa un acceso de VPC, no podrá volver atrás más adelante y cambiar el receptor (de VPC o público) una vez creada la canalización. Debe eliminar y volver a crear manualmente la canalización con el nuevo receptor. Aún puede cambiar un receptor de público a un receptor con acceso de VPC.
- No puede proporcionar acceso de incorporación entre cuentas a las canalizaciones de VPC.

Requisitos previos

Antes de poder aprovisionar una canalización con acceso a la VPC, debe hacer lo siguiente:

- Creación de una VPC

  Para crear su VPC, puede utilizar la consola de Amazon VPC, la AWS CLI o una de las. AWS SDKs Para obtener más información, consulte Trabajar con VPCs en la Guía del usuario de Amazon VPC. Si ya tiene una VPC, puede omitir este paso.

- Reservar direcciones IP

  OpenSearch La ingestión coloca una interface de red elástica en cada subred que especifique durante la creación de la canalización. Cada interfaz de red tiene asociada una dirección IP. Debe reservar una dirección IP por subred para las interfaces de red.

Configuración del acceso mediante VPC para una canalización

Puede habilitar el acceso a la VPC para una canalización desde la consola OpenSearch de servicio o mediante. AWS CLI

Consola

El acceso a la VPC se configura durante la creación de la canalización. Bajo Opciones de red de origen, elija Acceso a VPC y configure los siguientes ajustes:

Opción	Description (Descripción)
Administración de puntos de conexión	Elija si quiere crear sus puntos de conexión de VPC usted mismo o dejar que OpenSearch Ingestion los cree por usted.
VPC	Seleccione el ID de la nube privada virtual (VPC) que desee utilizar. La VPC y la canalización deben estar en la misma Región de AWS.
Subredes	Elija una o más subredes. OpenSearch El servicio coloca un punto final de VPC e interfaces de red elásticas en las subredes.
Grupos de seguridad	Elija uno o más grupos de seguridad de VPC que permitan que la aplicació n requerida llegue a la canalización de OpenSearch ingestión en los puertos (80 o 443) y protocolos (HTTP o HTTPs) expuestos por la canalización.
Opciones de conexión de VPC	Si su fuente requiere una comunicación entre VPC, como Amazon DocumentDB, OpenSearch self-managed o Confluent Kafka OpenSearc h , Ingestion crea interfaces de red elásticas ENIs () en las subredes que especifique para conectarse a estas fuentes. OpenSearch Ingestion utiliza en cada zona de disponibilidad para llegar a las fuentes ENIs especific adas. La opción Adjuntar a la VPC conecta la VPC del plano de datos de OpenSearch ingestión a la VPC especificada. Seleccione una reserva de CIDR para que la VPC administrada implemente la interfaz de red.

CLI

Para configurar el acceso a la VPC mediante AWS CLI, especifique el `--vpc-options` parámetro:

```
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"
```

Puntos de conexión de VPC autogestionados

Al crear una canalización, puede utilizar la administración de puntos de conexión para crear una canalización con puntos de conexión autogestionados o puntos de conexión administrados por el servicio. La administración de puntos finales es opcional y, de forma predeterminada, son los puntos finales administrados por Ingestion. OpenSearch

Para crear una canalización con un punto final de VPC autogestionado en, consulte Crear canalizaciones con Consola de administración de AWS OpenSearch la consola de servicio. Para crear una canalización con un punto final de VPC autogestionado en, puedes usar AWS CLI el parámetro `--vpc-options` del comando create-pipeline:

```
--vpc-options SubnetIds=subnet-abcdef01234567890,VpcEndpointManagement=CUSTOMER
```

Puede crear un punto de conexión para la canalización al especificar el servicio del punto de conexión. Para encontrar el servicio del punto de conexión, use el comando get-pipeline, que devuelve una respuesta similar a la siguiente:

```
"vpcEndpointService" : "com.amazonaws.osis.us-east-1.pipeline-
id-1234567890abcdef1234567890",
"vpcEndpoints" : [
  {
    "vpcId" : "vpc-1234567890abcdef0",
    "vpcOptions" : {
      "subnetIds" : [ "subnet-abcdef01234567890", "subnet-021345abcdef6789" ],
      "vpcEndpointManagement" : "CUSTOMER"
    }
  }
```

Utilice la respuesta `vpcEndpointService` de la respuesta para crear un punto final de VPC con o. Consola de administración de AWS AWS CLI

Si utiliza puntos de conexión de VPC autogestionados, debe habilitar los atributos de DNS `enableDnsSupport` y `enableDnsHostnames` en su VPC. Tenga en cuenta que, si tiene una canalización con un punto de conexión autoadministrado que debe detener y reiniciar, debe volver a crear el punto de conexión de VPC en su cuenta.

Roles vinculados a servicios para el acceso mediante VPC

Un rol vinculado a un servicio es un tipo único de rol de IAM; que delega permisos en un servicio para que pueda crear y administrar recursos en nombre del usuario. Si elige un punto de enlace de VPC gestionado por un servicio, OpenSearch Ingestion requiere un rol vinculado al servicio llamado para AWSServiceRoleForAmazonOpenSearchIngestionServiceacceder a su VPC, crear el punto de enlace de canalización y colocar las interfaces de red en una subred de su VPC.

Si eliges un endpoint de VPC autogestionado OpenSearch , Ingestion requiere un rol vinculado a un servicio denominado. AWSServiceRoleForOpensearchIngestionSelfManagedVpce Para obtener más información sobre estos roles, sus permisos y cómo se eliminan, consulte the section called “Rol de creación de canalizaciones”.

OpenSearch Ingestion crea el rol automáticamente al crear una canalización de ingestión. Para que esta creación automática se realice correctamente, el usuario que cree la primera canalización en una cuenta debe tener permisos para realizar la acción `iam:CreateServiceLinkedRole`. Para obtener más información, consulte Permisos de roles vinculados a servicios en la Guía del usuario de IAM. Puede ver el rol en la consola AWS Identity and Access Management (IAM) una vez creado.

Configuración de canalizaciones OpenSearch de ingestión para la ingestión entre cuentas

En el caso de fuentes basadas en push, como HTTP y OTel, Amazon OpenSearch Ingestion le permite compartir canalizaciones Cuentas de AWS desde una nube privada virtual (VPC) a un punto final de canalización en una VPC independiente. Los equipos que comparten análisis con otros equipos de su organización utilizan esta característica como medio más ágil de, por ejemplo, compartir los análisis de registros.

En esta sección, se usan los siguientes términos:

- Propietario de la canalización: la cuenta que posee y administra la canalización de ingestion. OpenSearch Solo una cuenta puede ser propietaria de una canalización.

- Cuenta de conexión: una cuenta que se conecta a una canalización compartida y la utiliza. Se pueden conectar varias cuentas a la misma canalización.

VPCs Para configurar el modo de compartir las canalizaciones de OpenSearch ingestión Cuentas de AWS, complete las siguientes tareas, tal y como se describe aquí:

- (Propietario de la canalización) Conceder a las cuentas conectadas acceso a una canalización
- (Cuenta de conexión) Cree una conexión de punto de conexión de canalización para cada VPC que se conecte

Antes de empezar

Antes de configurar el modo VPCs de compartir los canales OpenSearch de ingestión Cuentas de AWS, complete las siguientes tareas:

Tarea	Details
Cree una o más OpenSearch canalizaciones de ingestión	Establezca las unidades de OpenSearch cómputo mínimas (OSUs) en 2 o más. Para obtener más información, consulte Creación de canalizaciones OpenSearch de Amazon Ingestion. Para obtener información acerca de cómo actualizar una canalización, consulte Actualización de las canalizaciones OpenSearch de Amazon Ingestion.
Cree una o más VPCs para la OpenSearch ingestión	Para habilitar el uso compartido de canalizaciones entre cuentas, cualquier VPC implicada en la canalización y los puntos de conexión de la canalización debe configurarse con los siguientes valores de DNS: • `enableDnsSupport=true` • `enableDnsHostnames=true`

Tarea	Details
	Para obtener más información, consulte Atributos de DNS para su VPC en la Guía del usuario de Amazon VPC.

Conceder a las cuentas conectadas acceso a una canalización

Los procedimientos de esta sección describen cómo utilizar la consola de OpenSearch servicio y configurar el acceso AWS CLI a las canalizaciones multicuenta mediante la creación de una política de recursos. Una política de recursos permite al propietario de una canalización especificar otras cuentas que pueden acceder a una canalización. Una vez creadas, las políticas de canalización existen mientras exista la canalización o hasta que se elimine la política.

> **Note**
>
> Las políticas de recursos no sustituyen a la autorización OpenSearch de ingesta estándar mediante permisos de IAM. Las políticas de recursos son un mecanismo de autorización adicional para permitir el acceso a las canalizaciones entre cuentas.

Temas

- Conceder a las cuentas conectadas acceso a una canalización (consola)
- Conceder a las cuentas conectadas acceso a una canalización (CLI)

Conceder a las cuentas conectadas acceso a una canalización (consola)

Usa el siguiente procedimiento para conceder a las cuentas conectadas acceso a una canalización mediante la consola de Amazon OpenSearch Service.

Para crear una conexión de punto de conexión en la canalización

1. En la consola de Amazon OpenSearch Service, en el panel de navegación, expande Ingestión y, a continuación, selecciona Pipelines.
2. En la sección Canalizaciones, elija el nombre de la canalización a la que desee conceder acceso a una cuenta conectada.
3. Elija la pestaña de puntos de conexión de VPC.

4. En la sección Entidades principales autorizadas, seleccione Autorizar cuenta.
5. En el campo ID de Cuenta de AWS , escriba el número de 12 dígitos (ID) de la cuenta y, a continuación, seleccione Autorizar.

Conceder a las cuentas conectadas acceso a una canalización (CLI)

Utilice el siguiente procedimiento para conceder acceso a las cuentas conectadas a una canalización mediante la AWS CLI.

Para conceder a las cuentas conectadas acceso a la canalización

1. Actualice a la última versión de AWS CLI (versión 2.0). Para obtener más información, consulte Instalación o actualización de la versión más reciente de la AWS CLI.
2. Abre la CLI en la cuenta y Región de AWS con la canalización que quieras compartir.
3. Ejecute el siguiente comando para crear una política de recursos para la canalización. Esta política otorga el permiso `osis:CreatePipelineEndpoint` a la canalización. La política incluye un parámetro en el que puede Cuenta de AWS IDs enumerar las opciones de autorización.

> **Note**
>
> En el siguiente comando, debe utilizar la forma abreviada del identificador de cuenta, proporcionando únicamente el identificador de cuenta de doce dígitos. El uso de un ARN no funcionará. También debe proporcionar el nombre de recurso de Amazon (ARN) de la canalización en el parámetro de la CLI para `resource-arn` y en la política JSON en `Resource`, como se muestra.

```
aws --region region osis put-resource-policy \
  --resource-arn arn:aws:osis:region:pipeline-owner-account-ID:pipeline/pipeline-name
  --policy 'IAM-policy'
```

Utilice una política como la siguiente para *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"
  }
  ]
 }
```

Cree una conexión de punto de conexión de canalización para cada VPC que se conecte

Después de que el propietario de la canalización conceda acceso a una canalización en su VPC mediante el procedimiento anterior, un usuario de la cuenta de conexión crea un punto de conexión de la canalización en su VPC. En esta sección se incluyen los procedimientos para crear puntos finales mediante la consola OpenSearch de servicio y el AWS CLI. Al crear un punto final, OpenSearch Ingestion realiza las siguientes acciones:

- Crea el rol AWSServiceRoleForAmazonOpenSearchIngestionServicevinculado al servicio en tu cuenta, si aún no existe. Este rol otorga al usuario de la cuenta que se conecta permiso para llamar a la acción de la CreatePipelineEndpointAPI.
- Crea el punto de conexión de la canalización.
- Configura el punto de conexión de la canalización para ingerir datos de la canalización compartida en la VPC propietaria de la canalización.

Temas

- Creación de una conexión de punto de conexión para la canalización (consola)
- Creación de una conexión de punto de conexión para la canalización (CLI)

Creación de una conexión de punto de conexión para la canalización (consola)

Utilice el siguiente procedimiento para crear una conexión de punto final de canalización mediante la consola OpenSearch de servicio.

Para crear una conexión de punto de conexión en la canalización

1. En la consola de Amazon OpenSearch Service, en el panel de navegación, expanda Ingestión y, a continuación, seleccione puntos de enlace de VPC.
2. En la página Puntos de conexión de VPC, elija Crear.
3. En la ubicación de la canalización, elija una opción. Si elige Cuenta corriente, seleccione la canalización de la lista. Si elige Cuenta cruzada, especifique el ARN de la canalización en el campo. El propietario de la canalización debe haber concedido el acceso a la canalización, tal y como se describe en Conceder a las cuentas conectadas acceso a una canalización.
4. En la sección de configuración de VPC, para VPC, elija una VPC de la lista.
5. En Subnets (Subredes), elija una subred.
6. En Grupos de seguridad, elija un grupo.
7. Seleccione Crear punto de conexión.

Espere a que se produzca la transición del estado del punto de conexión que creó a `ACTIVE`. Una vez que la canalización esté `ACTIVE`, verá un nuevo campo con el nombre `ingestEndpointUrl`. Utilice este punto final para acceder a la canalización e ingerir datos mediante un cliente similar. FluentBit Para obtener más información sobre cómo utilizar FluentBit para ingerir datos, consulte. Uso de una canalización OpenSearch de ingestión con Fluent Bit

> **Note**
>
> `ingestEndpointUrl` es la misma URL para todas las cuentas conectadas.

Creación de una conexión de punto de conexión para la canalización (CLI)

Utilice el siguiente procedimiento para crear una conexión en el punto de conexión de canalización mediante AWS CLI.

Para crear una conexión de punto de conexión en la canalización

1. Si aún no lo ha hecho, actualice a la última versión de AWS CLI (versión 2.0). Para obtener más información, consulte Instalación o actualización de la versión más reciente de la AWS CLI.
2. Abra la CLI en la cuenta de conexión en Región de AWS la canalización compartida.
3. Ejecute el siguiente comando para crear una canalización en el punto de conexión.

> **Note**
>
> Proporcione al menos una subred y un grupo de seguridad para la VPC de la cuenta conectada. El grupo de seguridad debe incluir el puerto 443 y los clientes de soporte en la VPC de la cuenta que se conecta.

```
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. Ejecute el siguiente comando para enumerar los puntos de conexión en la región especificada en el comando anterior:

```
aws osis --region region list-pipeline-endpoints
```

Espere a que se produzca la transición del estado del punto de conexión que creó a `ACTIVE`. Una vez que la canalización esté `ACTIVE`, verá un nuevo campo con el nombre `ingestEndpointUrl`. Utilice este punto final para acceder a la canalización e ingerir datos mediante un cliente similar FluentBit. Para obtener más información sobre cómo utilizar FluentBit para ingerir datos, consulte. Uso de una canalización OpenSearch de ingestión con Fluent Bit

> **Note**
>
> `ingestEndpointUrl` es la misma URL para todas las cuentas conectadas.

Eliminar los punto de conexión de la canalización

Si ya no quiere proporcionar acceso a una canalización compartida, puede eliminar un punto de conexión de la canalización mediante uno de los siguientes métodos:

- Elimine el punto de conexión de la canalización (cuenta de conexión).
- Revoque el punto de conexión de la canalización (propietario de la canalización).

Siga este procedimiento para eliminar un punto de conexión de la canalización en una cuenta de conexión.

Para eliminar un punto de conexión de la canalización (cuenta de conexión)

1. Abra la CLI en la cuenta de conexión en Región de AWS la canalización compartida.
2. Ejecute el siguiente comando para enumerar un punto de conexión de las canalizaciones en la región:

   ```
   aws osis --region region list-pipeline-endpoints
   ```

   Anote el ID de canalización que desea eliminar.
3. Ejecute el siguiente comando para eliminar el punto de conexión de la canalización:

   ```
   aws osis --region region delete-pipeline-endpoint \
     --endpoint-id 'ID'
   ```

Como propietario de la canalización compartida, utilice el siguiente procedimiento para revocar un punto de conexión de la canalización.

Para revocar un punto de conexión de la canalización (propietario de la canalización)

1. Abra la CLI en la cuenta de conexión en Región de AWS la canalización compartida.

2. Ejecute el siguiente comando para enumerar las conexiones de punto de conexión de las canalizaciones en la región:

```
aws osis --region region list-pipeline-endpoint-connections
```

Anote el ID de canalización que desea eliminar.

3. Ejecute el siguiente comando para eliminar el punto de conexión de la canalización:

```
aws osis --region region revoke-pipeline-endpoint-connections \
  --pipeline-arn pipeline-arn --endpoint-ids ID
```

El comando permite especificar solo un ID de punto de conexión.

Identity and Access Management para Amazon OpenSearch Ingestion

AWS Identity and Access Management (IAM) es una herramienta Servicio de AWS que ayuda al administrador a controlar de forma segura el acceso a los recursos. AWS Los administradores de IAM controlan quién puede autenticarse (iniciar sesión) y quién puede autorizarse (tener permisos) para usar OpenSearch los recursos de Ingestion. La IAM es una opción Servicio de AWS que puede utilizar sin coste adicional.

Temas

- Políticas de ingestión basadas en la identidad OpenSearch
- Acciones políticas para la ingestión OpenSearch
- Recursos de políticas para la ingestión OpenSearch
- Claves de condición de la política para Amazon OpenSearch Ingestion
- ABAC con ingestión OpenSearch
- Uso de credenciales temporales con Ingestion OpenSearch
- Funciones vinculadas al servicio para Ingestion OpenSearch
- Ejemplos de políticas de ingestión basadas en la identidad OpenSearch

Políticas de ingestión basadas en la identidad OpenSearch

Compatibilidad con las políticas basadas en identidad: sí

Las políticas basadas en identidad son documentos de políticas de permisos JSON que puede asociar a una identidad, como un usuario de IAM, un grupo de usuarios o un rol. Estas políticas controlan qué acciones pueden realizar los usuarios y los roles, en qué recursos y en qué condiciones. Para obtener más información sobre cómo crear una política basada en la identidad, consulte Definición de permisos de IAM personalizados con políticas administradas por el cliente en la Guía del usuario de IAM.

Con las políticas basadas en identidades de IAM, puede especificar las acciones y los recursos permitidos o denegados, así como las condiciones en las que se permiten o deniegan las acciones. Para obtener más información sobre los elementos que puede utilizar en una política de JSON, consulte Referencia de los elementos de la política de JSON de IAM en la Guía del usuario de IAM.

Ejemplos de políticas de ingestión basadas en la identidad OpenSearch

Para ver ejemplos de políticas de OpenSearch ingestión basadas en la identidad, consulte. the section called “Ejemplos de políticas basadas en identidades”

Acciones políticas para la ingestión OpenSearch

Compatibilidad con las acciones de políticas: sí

El elemento `Action` de una política JSON describe las acciones que puede utilizar para conceder o denegar el acceso en una política. Las acciones políticas suelen tener el mismo nombre que la operación de AWS API asociada. Hay algunas excepciones, como acciones de solo permiso que no tienen una operación de API coincidente. También hay algunas operaciones que requieren varias acciones en una política. Estas acciones adicionales se denominan acciones dependientes.

Incluya acciones en una política para conceder permisos y así llevar a cabo la operación asociada.

En las acciones de política OpenSearch de Ingestion, se utiliza el siguiente prefijo antes de la acción:

```
osis
```

Para especificar varias acciones en una única instrucción, sepárelas con comas.

```
"Action": [
      "osis:action1",
      "osis:action2"
         ]
```

Puede especificar varias acciones utilizando caracteres comodín (*). Por ejemplo, para especificar todas las acciones que comiencen con la palabra `List`, incluya la siguiente acción:

```
"Action": "osis:List*"
```

Para ver ejemplos de políticas de OpenSearch ingestión basadas en la identidad, consulte. Ejemplos de políticas basadas en la identidad para Serverless OpenSearch

Recursos de políticas para la ingestión OpenSearch

Compatibilidad con los recursos de políticas: sí

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Resource` de la política JSON especifica el objeto u objetos a los que se aplica la acción. Como práctica recomendada, especifique un recurso utilizando el Nombre de recurso de Amazon (ARN). En el caso de las acciones que no admiten permisos por recurso, utilice un carácter comodín (*) para indicar que la instrucción se aplica a todos los recursos.

```
"Resource": "*"
```

Claves de condición de la política para Amazon OpenSearch Ingestion

Compatibilidad con claves de condición de políticas específicas del servicio: no

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Condition` especifica cuándo se ejecutan las instrucciones en función de criterios definidos. Puede crear expresiones condicionales que utilizan operadores de condición, tales como igual o menor que, para que la condición de la política coincida con los valores de la solicitud. Para ver todas las claves de condición AWS globales, consulte las claves de contexto de condición AWS globales en la Guía del usuario de IAM.

Para ver una lista de claves de condición de OpenSearch ingesta, consulta Claves de condición de Amazon OpenSearch Ingestion en la Referencia de autorización de servicio. Para saber con qué acciones y recursos puede utilizar una clave de condición, consulte Acciones definidas por Amazon OpenSearch Ingestion.

ABAC con ingestión OpenSearch

Admite ABAC (etiquetas en las políticas): sí

El control de acceso basado en atributos (ABAC) es una estrategia de autorización que define permisos en función de atributos denominados etiquetas. Puede adjuntar etiquetas a las entidades y AWS los recursos de IAM y, a continuación, diseñar políticas de ABAC para permitir las operaciones cuando la etiqueta del principal coincida con la etiqueta del recurso.

Para controlar el acceso en función de etiquetas, debe proporcionar información de las etiquetas en el elemento de condición de una política utilizando las claves de condición `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` o `aws:TagKeys`.

Si un servicio admite las tres claves de condición para cada tipo de recurso, el valor es Sí para el servicio. Si un servicio admite las tres claves de condición solo para algunos tipos de recursos, el valor es Parcial.

Para obtener más información sobre ABAC, consulte Definición de permisos con la autorización de ABAC en la Guía del usuario de IAM. Para ver un tutorial con los pasos para configurar ABAC, consulte Uso del control de acceso basado en atributos (ABAC) en la Guía del usuario de IAM.

Para obtener más información sobre el etiquetado de los recursos de OpenSearch ingestión, consulte. the section called “Etiquetado de canalizaciones”

Uso de credenciales temporales con Ingestion OpenSearch

Compatibilidad con credenciales temporales: sí

Las credenciales temporales proporcionan acceso a AWS los recursos a corto plazo y se crean automáticamente cuando se utiliza la federación o se cambia de rol. AWS recomienda generar credenciales temporales de forma dinámica en lugar de utilizar claves de acceso a largo plazo. Para obtener más información, consulte Credenciales de seguridad temporales en IAM y Servicios de AWS que funcionan con IAM en la Guía del usuario de IAM.

Funciones vinculadas al servicio para Ingestion OpenSearch

Compatible con roles vinculados al servicio: sí

Un rol vinculado a un servicio es un tipo de rol de servicio que está vinculado a un. Servicio de AWS El servicio puede asumir el rol para realizar una acción en su nombre. Los roles vinculados al servicio aparecen en usted Cuenta de AWS y son propiedad del servicio. Un administrador de IAM puede ver, pero no editar, los permisos de los roles vinculados a servicios.

OpenSearch Ingestion usa un rol vinculado a un servicio llamado. `AWSServiceRoleForAmazonOpenSearchIngestionService` El rol vinculado al servicio denominado `AWSServiceRoleForOpensearchIngestionSelfManagedVpce` también está disponible para canalizaciones con puntos de conexión de VPC autogestionados. Para obtener más información sobre la creación y administración de funciones vinculadas al servicio OpenSearch de Ingestion, consulte. the section called “Rol de creación de canalizaciones”

Ejemplos de políticas de ingestión basadas en la identidad OpenSearch

De forma predeterminada, los usuarios y los roles no tienen permiso para crear o modificar OpenSearch los recursos de ingestión. Un administrador de IAM puede crear políticas de IAM para conceder permisos a los usuarios para realizar acciones en los recursos que necesitan.

Para obtener información acerca de cómo crear una política basada en identidades de IAM mediante el uso de estos documentos de políticas JSON de ejemplo, consulte Creación de políticas de IAM (consola) en la Guía del usuario de IAM.

Para obtener más información sobre las acciones y los tipos de recursos definidos por Amazon OpenSearch Ingestion, incluido el formato de cada uno de los tipos de recursos, consulte Acciones, recursos y claves de condición de Amazon OpenSearch Ingestion en la Referencia de autorización de servicio. ARNs

Temas

- Prácticas recomendadas sobre las políticas
- Uso de Ingestion en la consola OpenSearch
- Administrar las canalizaciones de ingestión OpenSearch
- Ingerir datos en una canalización de ingestión OpenSearch

Prácticas recomendadas sobre las políticas

Las políticas basadas en identidades son muy eficaces. Determinan si alguien puede crear, eliminar o acceder a los recursos de OpenSearch Ingestion de su cuenta. Estas acciones pueden generar costos adicionales para su Cuenta de AWS. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidades:

Las políticas basadas en la identidad determinan si alguien puede crear los recursos de OpenSearch ingestión de tu cuenta, acceder a ellos o eliminarlos. Estas acciones pueden generar costos

adicionales para su Cuenta de AWS. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidades:

- Comience con las políticas AWS administradas y opte por los permisos con privilegios mínimos: para empezar a conceder permisos a sus usuarios y cargas de trabajo, utilice las políticas AWS administradas que otorgan permisos para muchos casos de uso comunes. Están disponibles en su. Cuenta de AWS Le recomendamos que reduzca aún más los permisos definiendo políticas administradas por el AWS cliente que sean específicas para sus casos de uso. Con el fin de obtener más información, consulte las políticas administradas por AWS o las políticas administradas por AWS para funciones de tarea en la Guía de usuario de IAM.
- Aplique permisos de privilegio mínimo: cuando establezca permisos con políticas de IAM, conceda solo los permisos necesarios para realizar una tarea. Para ello, debe definir las acciones que se pueden llevar a cabo en determinados recursos en condiciones específicas, también conocidos como permisos de privilegios mínimos. Con el fin de obtener más información sobre el uso de IAM para aplicar permisos, consulte Políticas y permisos en IAM en la Guía del usuario de IAM.
- Utilice condiciones en las políticas de IAM para restringir aún más el acceso: puede agregar una condición a sus políticas para limitar el acceso a las acciones y los recursos. Por ejemplo, puede escribir una condición de políticas para especificar que todas las solicitudes deben enviarse utilizando SSL. También puedes usar condiciones para conceder el acceso a las acciones del servicio si se utilizan a través de una acción específica Servicio de AWS, por ejemplo CloudFormation. Para obtener más información, consulte Elementos de la política de JSON de IAM: Condición en la Guía del usuario de IAM.
- Utiliza el analizador de acceso de IAM para validar las políticas de IAM con el fin de garantizar la seguridad y funcionalidad de los permisos: el analizador de acceso de IAM valida políticas nuevas y existentes para que respeten el lenguaje (JSON) de las políticas de IAM y las prácticas recomendadas de IAM. El analizador de acceso de IAM proporciona más de 100 verificaciones de políticas y recomendaciones procesables para ayudar a crear políticas seguras y funcionales. Para más información, consulte Validación de políticas con el Analizador de acceso de IAM en la Guía del usuario de IAM.
- Requerir autenticación multifactor (MFA): si tiene un escenario que requiere usuarios de IAM o un usuario raíz en Cuenta de AWS su cuenta, active la MFA para mayor seguridad. Para exigir la MFA cuando se invoquen las operaciones de la API, añada condiciones de MFA a sus políticas. Para más información, consulte Acceso seguro a la API con MFA en la Guía del usuario de IAM.

Para obtener más información sobre las prácticas recomendadas de IAM, consulte Prácticas recomendadas de seguridad en IAM en la Guía del usuario de IAM.

Uso de Ingestion en la consola OpenSearch

Para acceder a OpenSearch Ingestion desde la consola de OpenSearch servicio, debe tener un conjunto mínimo de permisos. Estos permisos deben permitirle enumerar y ver detalles sobre los recursos de OpenSearch Ingestion de su AWS cuenta. Si se crea una política basada en identidades que es más restrictiva que los permisos necesarios mínimos, la consola no funcionará del modo esperado para las entidades (tales como roles de IAM) que tengan esa política.

No es necesario que concedas permisos mínimos de consola a los usuarios que solo realizan llamadas a la API AWS CLI o a la AWS API. En su lugar, permite acceso únicamente a las acciones que coincidan con la operación de API que intenta realizar.

La siguiente política permite a los usuarios acceder a OpenSearch Ingestion desde la consola de OpenSearch servicio:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Resource": "*",
            "Effect": "Allow",
            "Action": [
                "osis:ListPipelines",
                "osis:GetPipeline",
                "osis:ListPipelineBlueprints",
                "osis:GetPipelineBlueprint",
                "osis:GetPipelineChangeProgress"
            ]
        }
    ]
}
```

Como alternativa, puede utilizar la política the section called “AmazonOpenSearchIngestionReadOnlyAccess” AWS gestionada, que concede acceso de solo lectura a todos los recursos de OpenSearch Ingestion durante un período de tiempo. Cuenta de AWS

Administrar las canalizaciones de ingestión OpenSearch

Esta política es un ejemplo de política de «administración de canalizaciones» que permite a un usuario gestionar y administrar las canalizaciones de Amazon OpenSearch Ingestion. El usuario puede crear, ver y eliminar canalizaciones.

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"
        }
    ]
}
```

Ingerir datos en una canalización de ingestión OpenSearch

Este ejemplo de política permite a un usuario u otra entidad ingerir datos en una canalización de Amazon OpenSearch Ingestion de su cuenta. El usuario no puede modificar las canalizaciones.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Resource": "arn:aws:osis:us-east-1:123456789012:pipeline/*",
            "Action": [
                "osis:Ingest"
            ],
            "Effect": "Allow"
        }
    ]
}
```

Registro de llamadas a la API OpenSearch Amazon Ingestion mediante AWS CloudTrail

Amazon OpenSearch Ingestion está integrado con AWS CloudTrail un servicio que proporciona un registro de las acciones realizadas por un usuario, un rol o un AWS servicio en OpenSearch Ingestion.

CloudTrail captura todas las llamadas a la API de OpenSearch Ingestion como eventos. Las llamadas capturadas incluyen las llamadas de la sección de OpenSearch ingestión de la consola de OpenSearch servicio y las llamadas en código a las operaciones de la API de OpenSearch ingestión.

Si crea una ruta, puede habilitar la entrega continua de CloudTrail eventos a un bucket de Amazon S3, incluidos los eventos de OpenSearch Ingestion. Si no configura una ruta, podrá ver los eventos más recientes en la CloudTrail consola, en el historial de eventos.

Con la información recopilada por usted CloudTrail, puede determinar la solicitud que se realizó a OpenSearch Ingestion, la dirección IP desde la que se realizó la solicitud, quién la hizo, cuándo se realizó y detalles adicionales.

Para obtener más información CloudTrail, consulta la Guía del AWS CloudTrail usuario.

OpenSearch Información sobre ingestión en CloudTrail

CloudTrail está habilitada en tu cuenta Cuenta de AWS al crear la cuenta. Cuando se produce una actividad en OpenSearch Ingestion, esa actividad se registra en un CloudTrail evento junto con otros

eventos de AWS servicio en el historial de eventos. Puede ver, buscar y descargar eventos recientes en su Cuenta de AWS. Para obtener más información, consulte Visualización de eventos con el historial de CloudTrail eventos.

Para tener un registro continuo de tus eventos Cuenta de AWS, incluidos los relacionados con la OpenSearch ingestión, crea un registro. Un rastro permite CloudTrail entregar archivos de registro a un bucket de Amazon S3. De forma predeterminada, cuando se crea un registro de seguimiento en la consola, el registro de seguimiento se aplica a todas las Regiones de AWS.

La ruta registra los eventos de todas las regiones de la AWS partición y envía los archivos de registro al bucket de Amazon S3 que especifique. Además, puede configurar otros AWS servicios para analizar más a fondo los datos de eventos recopilados en los CloudTrail registros y actuar en función de ellos. Para más información, consulte los siguientes temas:

- Introducción a la creación de registros de seguimiento
- CloudTrail servicios e integraciones compatibles
- Configuración de las notificaciones de Amazon SNS para CloudTrail
- Recibir archivos de CloudTrail registro de varias regiones y recibir archivos de CloudTrail registro de varias cuentas

Todas las acciones OpenSearch de ingestión se registran CloudTrail y se documentan en la referencia de la API OpenSearch de ingestión. Por ejemplo, las llamadas a las acciones `CreateCollection`, `ListCollections`, y`DeleteCollection` generan entradas en los archivos de registro de CloudTrail .

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. La información de identidad lo ayuda a determinar:

- Si la solicitud se realizó con credenciales de usuario root o AWS Identity and Access Management (IAM).
- Si la solicitud se realizó con credenciales de seguridad temporales de un rol o fue un usuario federado.
- Si la solicitud la realizó otro AWS servicio.

Para obtener más información, consulte el elemento userIdentity de CloudTrail .

Descripción de las entradas del archivo de registro de OpenSearch ingestión

Un rastro es una configuración que permite la entrega de eventos como archivos de registro a un bucket de Amazon S3 que especifique. CloudTrail Los archivos de registro contienen una o más entradas de registro.

Un evento representa una única solicitud desde cualquier origen. Incluye información sobre la acción solicitada, la fecha y la hora de la acción, los parámetros de la solicitud, etc. CloudTrail Los archivos de registro no son un registro ordenado de las llamadas a las API públicas, por lo que no aparecen en ningún orden específico.

En el siguiente ejemplo, se muestra una entrada de CloudTrail registro que demuestra la `DeletePipeline` acción.


```
{
    "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"
}
```

API OpenSearch de Amazon Ingestion y puntos de enlace de interfaz ()AWS PrivateLink

Puede establecer una conexión privada entre los puntos de enlace de la VPC y de la API de OpenSearch ingestión mediante la creación de un punto de enlace de la VPC de interfaz. Puntos de conexión de tipo interfaz con tecnología de AWS PrivateLink.

AWS PrivateLink le permite acceder de forma privada a las operaciones OpenSearch de la API de ingestión sin una pasarela de Internet, un dispositivo NAT, una conexión VPN o una conexión. Direct Connect Los recursos de su VPC no necesitan direcciones IP públicas para comunicarse con los puntos finales de la API de OpenSearch ingestión a fin de crear, modificar o eliminar canalizaciones. El tráfico entre tu VPC e OpenSearch Ingestión no sale de la red de Amazon.

> **Note**
>
> En este tema, se describen los puntos de enlace de la VPC para acceder a la API de OpenSearch ingestión, que le permite gestionar las canalizaciones (crear, actualizar y eliminar) desde su VPC. Esto es diferente de configurar el acceso a la VPC para las propias canalizaciones, que controla la forma en que se ingieren los datos en las canalizaciones desde las fuentes de la VPC. Para obtener información sobre la configuración del acceso a la VPC para las canalizaciones, consulte. the section called “Configuración del acceso a la VPC para canalizaciones”

Cada punto de conexión de la interfaz está representado por una o más interfaces de red elásticas en las subredes. Para obtener más información sobre las interfaces de red elásticas, consulte Interfaces de red elásticas en la Guía del usuario de Amazon EC2.

Para obtener más información sobre puntos de conexión de VPC, consulte Puntos de conexión de VPC de interfaz (AWS PrivateLink) en la Guía del usuario de Amazon VPC. Para obtener más información sobre las operaciones de la API OpenSearch de ingestión, consulta la referencia de la API de OpenSearch ingestión.

Consideraciones para los puntos de enlace de VPC

Antes de configurar un punto de enlace de VPC de interfaz para los puntos de enlace de la API de OpenSearch ingestión, asegúrese de revisar las propiedades y limitaciones de los puntos de enlace de la interfaz en la Guía del usuario de Amazon VPC.

Todas las operaciones OpenSearch de la API de ingestión relevantes para la administración de los recursos de OpenSearch ingestión están disponibles en su VPC mediante. AWS PrivateLink

Las políticas de puntos de enlace de VPC son compatibles con los puntos de enlace de la API OpenSearch de ingestión. De forma predeterminada, se permite el acceso total a las operaciones de la API de OpenSearch ingestión a través del punto final. Para más información, consulte Control del acceso a los servicios con puntos de conexión de VPC en la Guía del usuario de Amazon VPC.

Disponibilidad.

OpenSearch La API de ingestión actualmente es compatible con los puntos finales de VPC en OpenSearch todas las regiones de ingestión.

En este momento, no se admiten los puntos finales FIPS.

Creación de un punto final de VPC de interfaz para OpenSearch la API de ingestión

Puede crear un punto de enlace de VPC para la API de OpenSearch ingestión mediante la consola de Amazon VPC o con (). AWS Command Line Interface AWS CLI Para obtener más información, consulte Creación de un punto de conexión de interfaz en la Guía del usuario de Amazon VPC.

Cree un punto final de VPC para la API de OpenSearch ingestión con el nombre del servicio. `com.amazonaws.region.osis`

Si habilitas el DNS privado para el punto final, puedes realizar solicitudes de API a OpenSearch Ingestion con el punto final de la VPC utilizando su nombre de DNS predeterminado para AWS la región, por ejemplo. `osis.us-east-1.amazonaws.com`

Para más información, consulte Acceso a un servicio a través de un punto de conexión de interfaz en la Guía del usuario de Amazon VPC.

Creación de una política de puntos finales de VPC para OpenSearch la API de ingestión

Puede adjuntar una política de punto final a su punto final de VPC que controle el acceso a la API de OpenSearch ingestión. La política especifica la siguiente información:

- La entidad principal que puede realizar acciones.
- Las acciones que se pueden realizar.
- Los recursos en los que se pueden llevar a cabo las acciones.

Para más información, consulte Control del acceso a los servicios con puntos de enlace de la VPC en la Guía del usuario de Amazon VPC.

Ejemplo: política de puntos finales de VPC para acciones de la API de OpenSearch ingestión

El siguiente es un ejemplo de una política de punto final para la API de OpenSearch ingestión. Cuando se adjunta a un punto final, esta política otorga acceso a las acciones de la API de OpenSearch ingestión enumeradas a todos los principales de todos los recursos.

```
{
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "osis:CreatePipeline",
            "osis:UpdatePipeline",
            "osis:DeletePipeline"
         ],
         "Resource":"*"
      }
   ]
}
```

Ejemplo: política de punto final de VPC que deniega todo acceso desde una cuenta específica AWS

La siguiente política de punto final de VPC deniega a la AWS cuenta `123456789012` todo acceso a los recursos que utilizan el punto final. La política permite todas las acciones de otras cuentas.

```
{
```

```
  "Statement": [
    {
      "Action": "*",
      "Effect": "Allow",
      "Resource": "*",
      "Principal": "*"
    },
    {
      "Action": "*",
      "Effect": "Deny",
      "Resource": "*",
      "Principal": { "AWS": [ "123456789012" ] }
     }
   ]
}
```

Etiquetado de las canalizaciones de Amazon OpenSearch Ingestion

Las etiquetas le permiten asignar información arbitraria a una canalización OpenSearch de Amazon Ingestion para que pueda categorizar y filtrar esa información. Una etiqueta es una etiqueta de metadatos que usted asigna o que AWS asigna a un recurso. AWS Cada etiqueta consta de una clave y un valor. En el caso de etiquetas que usted asigna, debe definir la clave y el valor. Por ejemplo, puede definir la clave como `stage` y el valor de un recurso como `test`.

Las etiquetas lo ayudan a hacer lo siguiente:

- Identifique y organice sus AWS recursos. Muchos AWS servicios admiten el etiquetado, por lo que puede asignar la misma etiqueta a los recursos de diferentes servicios para indicar que los recursos están relacionados. Por ejemplo, puedes asignar la misma etiqueta a una canalización de OpenSearch ingestión que a un dominio de Amazon OpenSearch Service.
- Realiza un seguimiento de tus AWS costes. Estas etiquetas se activan en el Administración de facturación y costos de AWS panel de control. AWS utiliza las etiquetas para clasificar los costes y entregarle un informe mensual de asignación de costes. Para obtener más información, consulte Uso de etiquetas de asignación de costos en la Guía del usuario de AWS Billing.
- Restrinja el acceso a las canalizaciones mediante el control de acceso basado en atributos. Para obtener más información, consulte Controlling access based on tag keys en la Guía del usuario de IAM.

En OpenSearch Ingestión, el recurso principal es una canalización. Puede usar la consola de OpenSearch servicio, la AWS CLI, OpenSearch Ingestion APIs o AWS SDKs para agregar, administrar y eliminar etiquetas de una canalización.

Temas

- Permisos necesarios
- Uso de etiquetas (consola)
- Uso de etiquetas (AWS CLI)

Permisos necesarios

OpenSearch Ingestion utiliza los siguientes permisos AWS Identity and Access Management Access Analyzer (IAM) para etiquetar las canalizaciones:

- `osis:TagResource`
- `osis:ListTagsForResource`
- `osis:UntagResource`

Para obtener más información sobre cada permiso, consulte las acciones, los recursos y las claves de condición de la OpenSearch ingestión en la Referencia de autorización de servicios.

Uso de etiquetas (consola)

La consola es la forma más sencilla para etiquetar una canalización.

Cómo crear una etiqueta

1. Inicia sesión en la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/osis/home. Estarás en la página de Pipelines.
2. Seleccione una canalización a la que desee agregar etiquetas y vaya a la pestaña Etiquetas.
3. Seleccione Administrar y Agregar nueva etiqueta.
4. Introduzca una clave de etiqueta y un valor opcional.
5. Seleccione guardar.

Para eliminar una etiqueta, siga los mismos pasos y elija Quitar en la página Administrar etiquetas.

Para obtener más información sobre cómo utilizar la consola para trabajar con etiquetas, consulte Tag Editor en la Guía de introducción a la consola de administración de AWS .

Uso de etiquetas (AWS CLI)

Para etiquetar una canalización mediante el AWS CLI, envía una `TagResource` solicitud:

```
aws osis tag-resource
  --arn arn:aws:osis:us-east-1:123456789012:pipeline/my-pipeline
  --tags Key=service,Value=osis Key=source,Value=otel
```

Elimine las etiquetas de una canalización mediante el comando `UntagResource`:

```
aws osis untag-resource
  --arn arn:aws:osis:us-east-1:123456789012:pipeline/my-pipeline
  --tag-keys service
```

Consulte las etiquetas existentes para una canalización con el comando `ListTagsForResource`:

```
aws osis list-tags-for-resource
  --arn arn:aws:osis:us-east-1:123456789012:pipeline/my-pipeline
```

Registro y supervisión de Amazon OpenSearch Ingestion con Amazon CloudWatch

Amazon OpenSearch Ingestion publica métricas y registros en Amazon CloudWatch.

Temas

- Monitoreo de registros de canalización
- Monitoreo de métricas de canalización

Monitoreo de registros de canalización

Puede habilitar el registro de las canalizaciones de Amazon OpenSearch Ingestion para exponer los mensajes de error y advertencia que se generen durante las operaciones de canalización y la actividad de ingestión. OpenSearch Ingestion publica todos los registros en Amazon CloudWatch

Logs. CloudWatch Los registros pueden monitorear la información de los archivos de registro y notificarle cuando se alcanzan ciertos umbrales. También se pueden archivar los datos del registro en un almacenamiento de larga duración. Para obtener más información, consulta la Guía del usuario CloudWatch de Amazon Logs.

Los registros de OpenSearch Ingestion pueden indicar un error en el procesamiento de las solicitudes, errores de autenticación desde el origen hasta el receptor y otras advertencias que pueden resultar útiles para solucionar problemas. Para sus registros, OpenSearch Ingestion utiliza los niveles de registro de`INFO`, `WARN`y `ERROR`. `FATAL` Recomendamos habilitar la publicación de registros en todas las canalizaciones.

Permisos necesarios

Para permitir que OpenSearch Ingestion envíe CloudWatch registros a Logs, debe iniciar sesión como usuario con determinados permisos de IAM.

Necesita los siguientes permisos de CloudWatch registro para crear y actualizar los recursos de entrega de registros:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Resource": "*",
            "Action": [
                "logs:CreateLogDelivery",
                "logs:PutResourcePolicy",
                "logs:UpdateLogDelivery",
                "logs:DeleteLogDelivery",
                "logs:DescribeResourcePolicies",
                "logs:GetLogDelivery",
                "logs:ListLogDeliveries"
            ]
        }
    ]
}
```

Habilitación de la publicación de registros

Puede habilitar la publicación de registros en las canalizaciones existentes o mientras crea una canalización. Para ver los pasos para habilitar la publicación de registros durante la creación de la canalización, consulte the section called “Creación de canalizaciones”.

Consola

Cómo habilitar la publicación de registros en una canalización existente

1. Inicia sesión en la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/osis/home. Estarás en la página de Pipelines.
2. Abra la canalización en la que desea habilitar los registros y, a continuación, seleccione Acciones y Editar las opciones de publicación de registros.
3. Habilita Publicar en CloudWatch registros.
4. Cree un grupo de registro nuevo o seleccione uno que ya exista. Se recomienda formatear el nombre como una ruta, por ejemplo`/aws/vendedlogs/OpenSearchIngestion/pipeline-name/audit-logs`. Este formato facilita la aplicación de una política de CloudWatch acceso que conceda permisos a todos los grupos de registros en una ruta específica, por ejemplo`/aws/vendedlogs/OpenSearchIngestion`.

   > **Important**
   >
   > Debe incluir el prefijo `vendedlogs` en el nombre del grupo de registro; de lo contrario, se producirá un error en la creación.

5. Seleccione Save.

CLI

Para habilitar la publicación de registros mediante el AWS CLI, envíe la siguiente solicitud:

```
aws osis update-pipeline \
  --pipeline-name my-pipeline \
  --log-publishing-options  IsLoggingEnabled=true,CloudWatchLogDestination={LogGroup="/aws/vendedlogs/OpenSearchIngestion/pipeline-name"}
```

Monitoreo de métricas de canalización

Puedes monitorizar las canalizaciones de Amazon OpenSearch Ingestion con Amazon CloudWatch, que recopila datos sin procesar y los procesa para convertirlos en métricas legibles prácticamente en tiempo real. Estas estadísticas se mantienen durante 15 meses, de forma que pueda obtener acceso a información histórica y disponer de una mejor perspectiva sobre el desempeño de su aplicación web o servicio. También puede establecer alarmas que vigilen determinados umbrales y enviar notificaciones o realizar acciones cuando se cumplan dichos umbrales. Para obtener más información, consulta la Guía del CloudWatch usuario de Amazon.

La consola OpenSearch de ingestión muestra una serie de gráficos basados en los datos sin procesar de la CloudWatch pestaña Rendimiento de cada canalización.

OpenSearch Ingestion informa de las métricas de la mayoría de los complementos compatibles. Si algunos complementos no tienen su propia tabla a continuación, significa que no muestran ninguna métrica específica de los complementos. Las métricas de canalizaciones se publican en el espacio de nombres de `AWS/OSIS`.

Temas

- Métricas comunes
- Métricas de búfer
- Métricas de la V4 de Signature
- Métricas del búfer de bloqueo limitado
- Métricas de origen de trazas de Otel
- Métricas de OTel: métricas de origen
- Métricas de Http
- Métricas de S3
- Métricas agrupadas
- Métricas de fecha
- métricas de Lambda
- Métricas de Grok
- Métricas sin formato de trazas de OTel
- Métricas del grupo de trazas de OTel
- Métricas de estado del mapa de servicio
- OpenSearch métricas

- Métricas del sistema y de medición

Métricas comunes

Las siguientes métricas son comunes a todos los procesadores y receptores.

Cada métrica lleva el nombre de la subcanalización y el nombre del complemento, en el formato **<** *sub_pipeline_name* **><** **><** **>***plugin*. *metric_name* Por ejemplo, el nombre completo de la métrica `recordsIn.count` de una subcanalización denominada `my-pipeline` y la fecha del procesador sería `my-pipeline.date.recordsIn.count`.

Sufijo de la métrica	Description (Descripción)
`recordsIn.count`	El ingreso de registros a un componente de canalización. Esta métrica se aplica a procesadores y receptores. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`recordsOut.count`	El egreso de registros de un componente de canalización. Esta métrica se aplica a procesadores y orígenes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`timeElapsed.count`	Un recuento de los puntos de datos registrados durante la ejecución de un componente de canalización. Esta métrica se aplica a procesadores y receptores. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`timeElapsed.sum`	El tiempo total transcurrido durante la ejecución de un componente de canalización. Esta métrica se aplica a procesadores y receptores, en milisegundos. Estadísticas pertinentes: suma

Sufijo de la métrica	Description (Descripción)
	Dimensión: `PipelineName`
`timeElapsed.max`	El tiempo máximo transcurrido durante la ejecución de un componente de canalización. Esta métrica se aplica a procesadores y receptores, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Métricas de búfer

Las siguientes métricas se aplican al búfer de bloqueo limitado predeterminado que OpenSearch Ingestion configura automáticamente para todas las canalizaciones.

Cada métrica lleva el nombre de la subcanalización y el nombre del búfer como prefijo, con el formato **< >< >< >**. *`sub_pipeline_name buffer_name metric_name`* Por ejemplo, el nombre completo de la métrica `recordsWritten.count` de una subcanalización denominada `my-pipeline` sería `my-pipeline.BlockingBuffer.recordsWritten.count`.

Sufijo de la métrica	Description (Descripción)
`recordsWritten.count`	El número de registros escritos en un búfer. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`recordsRead.count`	El número de registros leídos de un búfer. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`recordsInFlight.value`	El número de registros sin comprobar de un búfer. Estadísticas pertinentes: promedio

Sufijo de la métrica	Description (Descripción)
	Dimensión: `PipelineName`
`recordsInBuffer.value`	El número de registros que hay actualmente en un búfer. Estadísticas pertinentes: promedio Dimensión: `PipelineName`
`recordsProcessed.count`	El número de registros leídos de un búfer y procesados por una canalización. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`recordsWriteFailed.count`	El número de registros que la canalización no ha podido escribir en el receptor. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`writeTimeElapsed.count`	Recuento de puntos de datos registrados mientras se escribe en un búfer. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`writeTimeElapsed.sum`	El tiempo total transcurrido durante la escritura en un búfer, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`writeTimeElapsed.max`	El tiempo máximo transcurrido durante la escritura en un búfer, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`writeTimeouts.count`	El recuento de tiempos de espera de escritura en un búfer. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`readTimeElapsed.count`	Recuento de puntos de datos registrados mientras se lee de un búfer. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`readTimeElapsed.sum`	El tiempo total transcurrido mientras se lee de un búfer, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`readTimeElapsed.max`	El tiempo máximo transcurrido mientras se lee de un búfer, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`checkpointTimeElapsed.count`	Un recuento de los puntos de datos registrados durante el punto de control. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`checkpointTimeElapsed.sum`	El tiempo total transcurrido durante el punto de control, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`checkpointTimeElapsed.max`	El tiempo máximo transcurrido durante el punto de control, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Métricas de la V4 de Signature

Las siguientes métricas se aplican al punto de conexión de ingesta de una canalización y están asociadas a los complementos de origen (`http`, `otel_trace`, y `otel_metrics`). Todas las solicitudes al punto de conexión de ingesta deben firmarse con la versión 4 de Signature. Estas métricas pueden ayudarle a identificar los problemas de autorización al conectarse a su canalización o a confirmar que se está autenticando correctamente.

Cada métrica lleva el nombre de la subcanalización como prefijo y `osis_sigv4_auth`. Por ejemplo, *`sub_pipeline_name`*`.osis_sigv4_auth.httpAuthSuccess.count`.

Sufijo de la métrica	Description (Descripción)
`httpAuthSuccess.count`	El número de solicitudes de Signature V4 que se han enviado correctamente a la canalización. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`httpAuthFailure.count`	El número de solicitudes fallidas de Signature V4 enviadas a la canalización. Estadísticas pertinentes: suma

Sufijo de la métrica	Description (Descripción)
	Dimensión: `PipelineName`
`httpAuthServerError.count`	El número de solicitudes de Signature V4 enviadas a la canalización que arrojaron errores en el servidor. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Métricas del búfer de bloqueo limitado

Las siguientes métricas se aplican al búfer de bloqueo limitado. Cada métrica lleva el nombre de la subcanalización como prefijo y `BlockingBuffer`. Por ejemplo, *`sub_pipeline_name`*`.BlockingBuffer.bufferUsage.value`.

Sufijo de la métrica	Description (Descripción)
`bufferUsage.value`	El porcentaje de uso del `buffer_size` se basa en el número de registros del búfer. `buffer_size` represent a el número máximo de registros escritos en el búfer, así como los registros en movimiento que no se han comprobado. Estadísticas pertinentes: promedio Dimensión: `PipelineName`

Métricas de origen de trazas de Otel

Las siguientes métricas se aplican a la fuente de rastreo. OTel Cada métrica lleva el nombre de la subcanalización como prefijo y `otel_trace_source`. Por ejemplo, *`sub_pipeline_name`*`.otel_trace_source.requestTimeouts.count`.

Sufijo de la métrica	Description (Descripción)
`requestTimeouts.count`	El número de solicitudes que agotaron el tiempo de espera. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestsReceived.count`	El número de solicitudes recibidas por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`successRequests.count`	El número de solicitudes que el complemento procesó correctamente. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`badRequests.count`	El número de solicitudes con un formato inválido procesadas por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestsTooLarge.count`	El número de solicitudes cuya cantidad de intervalos en el contenido es superior a la capacidad del búfer. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`internalServerError.count`	El número de solicitudes procesadas por el complemento con un tipo de excepción personalizado. Estadísticas pertinentes: suma

Sufijo de la métrica	Description (Descripción)
	Dimensión: `PipelineName`
`requestProcessDuration.count`	Un recuento de los puntos de datos registrados durante el procesamiento de las solicitudes por parte del complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestProcessDuration.sum`	La latencia total de las solicitudes procesadas por el complemento, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestProcessDuration.max`	La latencia máxima de las solicitudes procesadas por el complemento, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`payloadSize.count`	Un recuento de la distribución de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`payloadSize.sum`	La distribución total de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`payloadSize.max`	La distribución máxima de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Métricas de OTel: métricas de origen

Las siguientes métricas se aplican a la fuente de OTel métricas. Cada métrica lleva el nombre de la subcanalización como prefijo y `otel_metrics_source`. Por ejemplo, *`sub_pipeline_name`*`.otel_metrics_source.requestTimeouts.count`.

Sufijo de la métrica	Description (Descripción)
`requestTimeouts.count`	El número total de solicitudes al complemento que agotaron el tiempo de espera. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestsReceived.count`	El número total de solicitudes recibidas por el complemen to. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`successRequests.count`	El número de solicitudes procesadas correctamente (código de estado de respuesta de 200) por el complemen to. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`requestProcessDuration.count`	Un recuento de la latencia de las solicitudes procesadas por el complemento, en segundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestProcessDuration.sum`	La latencia total de las solicitudes procesadas por el complemento, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestProcessDuration.max`	La latencia máxima de las solicitudes procesadas por el complemento, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`payloadSize.count`	Un recuento de la distribución de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`payloadSize.sum`	La distribución total de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`payloadSize.max`	La distribución máxima de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Métricas de Http

Las siguientes métricas se aplican al origen de HTTP. Cada métrica lleva el nombre de la subcanalización como prefijo y `http`. Por ejemplo, *`sub_pipeline_name`*`.http.requestsReceived.count`.

Sufijo de la métrica	Description (Descripción)
`requestsReceived.count`	El número de solicitudes recibidas por el punto de conexión de `/log/ingest` . Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestsRejected.count`	El número de solicitudes rechazadas (código de estado de respuesta de 429) por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`successRequests.count`	El número de solicitudes procesadas correctamente (código de estado de respuesta de 200) por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`badRequests.count`	El número de solicitudes con un tipo o formato de contenido no válido (código de estado de respuesta de 400) procesadas por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestTimeouts.count`	El número de solicitudes que agotan el tiempo de espera en el servidor origen de HTTP (código de estado de respuesta 415). Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestsTooLarge.count`	El número de solicitudes cuyo tamaño de eventos en el contenido es superior a la capacidad del búfer (código de estado de respuesta de 413). Estadísticas pertinentes: suma Dimensión: `PipelineName`
`internalServerError.count`	El número de solicitudes procesadas por el complemento con un tipo de excepción personalizado (código de estado de respuesta de 500). Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestProcessDuration.count`	Un recuento de la latencia de las solicitudes procesadas por el complemento, en segundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`requestProcessDuration.sum`	La latencia total de las solicitudes procesadas por el complemento, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestProcessDuration.max`	La latencia máxima de las solicitudes procesadas por el complemento, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`payloadSize.count`	Un recuento de la distribución de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`payloadSize.sum`	La distribución total de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`payloadSize.max`	La distribución máxima de los tamaños de carga útil de las solicitudes entrantes, en bytes. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Métricas de S3

Las siguientes métricas se aplican a la fuente de S3. Cada métrica lleva el nombre de la subcanalización como prefijo y `s3`. Por ejemplo, *`sub_pipeline_name`*`.s3.s3ObjectsFailed.count`.

Sufijo de la métrica	Description (Descripción)
`s3ObjectsFailed.count`	El número total de objetos S3 que el complemento no ha podido leer. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectsNotFound.count`	El número de objetos de S3 que el complemento no ha podido leer debido a un error `Not Found` de S3. Estas métricas también se tienen en cuenta para la métrica `s3ObjectsFailed`. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectsAccessDenied.count`	El número de objetos de S3 que el complemento no ha podido leer debido a un error `Access Denied` o `Forbidden` de S3. Estas métricas también se tienen en cuenta para la métrica `s3ObjectsFailed`. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectReadTimeElapsed.count`	El tiempo que tarda el complemento en realizar una solicitud GET para un objeto de S3, analizarla y escribir los eventos en el búfer. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`s3ObjectReadTimeElapsed.sum`	El tiempo total que tarda el complemento en realizar una solicitud GET para un objeto de S3, analizarla y escribir los eventos en el búfer, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectReadTimeElapsed.max`	El tiempo máximo que tarda el complemento en realizar una solicitud GET para un objeto de S3, analizarla y escribir los eventos en el búfer, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`s3ObjectSizeBytes.count`	El recuento de la distribución de los tamaños de los objetos de S3, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectSizeBytes.sum`	La distribución total de los tamaños de los objetos de S3, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectSizeBytes.max`	La distribución máxima de los tamaños de los objetos de S3, en bytes. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`s3ObjectProcessedBytes.count`	El recuento de la distribución de los objetos de S3 procesados por el complemento, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectProcessedBytes.sum`	La distribución total de los objetos de S3 procesados por el complemento, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectProcessedBytes.max`	La distribución máxima de los objetos de S3 procesados por el complemento, en bytes. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`s3ObjectsEvents.count`	El recuento de la distribución de los eventos de S3 recibidos por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3ObjectsEvents.sum`	La distribución total de los eventos de S3 recibidos por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`s3ObjectsEvents.max`	La distribución máxima de los eventos de S3 recibidos por el complemento. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`sqsMessageDelay.count`	Un recuento de puntos de datos registrados mientras S3 registra el tiempo de un evento desde la creación de un objeto hasta el momento en que se analiza por completo. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`sqsMessageDelay.sum`	La cantidad de tiempo total entre el momento en que S3 registra un evento para la creación de un objeto y el momento en que se analiza por completo, en milisegun dos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`sqsMessageDelay.max`	La cantidad máxima de tiempo entre el momento en que S3 registra un evento para la creación de un objeto y el momento en que se analiza por completo, en milisegun dos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`s3ObjectsSucceeded.count`	El número de objetos de S3 que el complemento ha leído correctamente. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`sqsMessagesReceived.count`	El número de mensajes de Amazon SQS recibidos de la cola por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`sqsMessagesDeleted.count`	El número de mensajes de Amazon SQS eliminados de la cola por el complemento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`sqsMessagesFailed.count`	El número de mensajes de Amazon SQS que el complemento no ha podido analizar. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Métricas agrupadas

Las siguientes métricas se aplican al procesador Aggregate. Cada métrica lleva el nombre de la subcanalización como prefijo y `aggregate`. Por ejemplo, *`sub_pipeline_name`*`.aggregate.actionHandleEventsOut.count`.

Sufijo de la métrica	Description (Descripción)
`actionHandleEvents Out.count`	El número de eventos que se devolvieron de la llamada `handleEvent` a la acción configurada. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`actionHandleEvents Dropped.count`	El número de eventos que se devolvieron de la llamada `handleEvent` a la acción configurada.

Sufijo de la métrica	Description (Descripción)
	Estadísticas pertinentes: suma Dimensión: `PipelineName`
`actionHandleEvents ProcessingErrors.count`	El número de llamadas realizadas a `handleEvent` para la acción configurada que provocaron un error. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`actionConcludeGrou pEventsOut.count`	El número de eventos que se devolvieron de la llamada `concludeGroup` a la acción configurada. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`actionConcludeGrou pEventsDropped.count`	El número de eventos que no se devolvieron de la llamada `condludeGroup` a la acción configurada. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`actionConcludeGrou pEventsProcessingE rrors.count`	El número de llamadas realizadas a `concludeGroup` para la acción configurada que provocaron un error. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`currentAggregateGr oups.value`	El número actual de grupos. Este indicador disminuye cuando se concluyen los grupos y aumenta cuando un evento inicia la creación de un grupo nuevo. Estadísticas pertinentes: promedio Dimensión: `PipelineName`

Métricas de fecha

Las siguientes métricas se aplican al procesador Date. Cada métrica lleva el nombre de la subcanalización como prefijo y `date`. Por ejemplo, *`sub_pipeline_name`*`.date.dateProcessingMatchSuccess.count`.

Sufijo de la métrica	Description (Descripción)
`dateProcessingMatchSuccess.count`	El número de registros que coinciden con al menos uno de los patrones especificados en la opción `match` de configuración. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`dateProcessingMatchFailure.count`	El número de registros que no coinciden con ninguno de los patrones especificados en la opción `match` de configuración. Estadísticas pertinentes: suma Dimensión: `PipelineName`

métricas de Lambda

Las siguientes métricas se aplican al procesador de AWS Lambda. Cada métrica lleva el nombre de la subcanalización como prefijo y `lambda`. Por ejemplo, *`sub_pipeline_name`*`.lambda.recordsSuccessfullySentToLambda.count`.

Sufijo de la métrica	Description (Descripción)
`recordsSuccessfullySentToLambda.count`	Número de registros procesados correctamente por la función de Lambda. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`recordsFailedToSendToLambda.count`	Número de registros que no se pudieron enviar a la función de Lambda. Estadísticas pertinentes: suma Dimensión: `PipelineName`

`lambdaFunctionLatency.avg`
 	
`lambdaFunctionLatency.max`
	La latencia de las invocaciones de funciones de Lambda.
	
	Estadísticas pertinentes: promedio y máximo
	
	Dimensión: `PipelineName`

`numberOfRequestsSucceeded.count`	El número total de solicitudes de invocaciones de Lambda correctas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`numberOfRequestsFailed.count`	El número total de errores de solicitud de invocación de Lambda. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`requestPayloadSize.avg`	El tamaño de las cargas de solicitud enviadas a Lambda. Estadísticas pertinentes: promedio Dimensión: `PipelineName`
`responsePayloadSize.avg`	El tamaño de las cargas útiles de respuesta recibidas de Lambda. Estadísticas pertinentes: promedio Dimensión: `PipelineName`

Métricas de Grok

Las siguientes métricas se aplican al procesador Grok. Cada métrica lleva el nombre de la subcanalización como prefijo y `grok`. Por ejemplo, *`sub_pipeline_name`*`.grok.grokProcessingMatch.count`.

Sufijo de la métrica	Description (Descripción)
`grokProcessingMatch.count`	El número de registros que encontraron al menos un patrón coincidente en la opción `match` de configuración. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`grokProcessingMismatch.count`	El número de registros que no coinciden con ninguno de los patrones especificados en la opción `match` de configuración. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`grokProcessingErrors.count`	El número de errores de procesamiento de registros. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`grokProcessingTimeouts.count`	El número de registros cuyo tiempo de espera se agotó durante la coincidencia. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`grokProcessingTime.count`	Un recuento de los puntos de datos registrados mientras un registro individual coincidía con los patrones de la opción `match` de configuración. Estadísticas pertinentes: suma

Sufijo de la métrica	Description (Descripción)
	Dimensión: `PipelineName`
`grokProcessingTime.sum`	El tiempo total que tarda cada registro individual en coincidir con los patrones de la opción `match` de configuración, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`grokProcessingTime.max`	El tiempo máximo que tarda cada registro individual en coincidir con los patrones de la opción `match` de configuración, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Métricas sin formato de trazas de OTel

Las siguientes métricas se aplican al procesador OTel Trace Raw. Cada métrica lleva el nombre de la subcanalización como prefijo y `otel_trace_raw`. Por ejemplo, *`sub_pipeline_name`*`.otel_trace_raw.traceGroupCacheCount.value`.

Sufijo de la métrica	Description (Descripción)
`traceGroupCacheCount.value`	El número de grupos de trazas en la memoria caché de grupos de trazas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`spanSetCount.value`	El número de conjuntos de intervalos de la colección de conjuntos de intervalos. Estadísticas pertinentes: suma

Sufijo de la métrica	Description (Descripción)
	Dimensión: `PipelineName`

Métricas del grupo de trazas de OTel

Las siguientes métricas se aplican al procesador de grupos de OTel trazas. Cada métrica lleva el nombre de la subcanalización como prefijo y `otel_trace_group`. Por ejemplo, *`sub_pipeline_name`*`.otel_trace_group.recordsInMissingTraceGroup.count`.

Sufijo de la métrica	Description (Descripción)
`recordsInMissingTraceGroup.count`	El número de registros de ingreso que faltan en los campos del grupo de trazas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`recordsOutFixedTraceGroup.count`	El número de registros de egreso con campos de grupos de trazas que se completaron correctamente. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`recordsOutMissingTraceGroup.count`	El número de registros de egreso que faltan en los campos del grupo de trazas. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Métricas de estado del mapa de servicio

Las siguientes métricas se aplican al procesador del estado del mapa de servicio. Cada métrica lleva el nombre de la subcanalización como prefijo y `service-map-stateful`. Por ejemplo, *`sub_pipeline_name`*`.service-map-stateful.spansDbSize.count`.

Sufijo de la métrica	Description (Descripción)
`spansDbSize.value`	Los tamaños de los bytes en memoria de los intervalos en MapDB durante la duración de la ventana actual y anterior. Estadísticas pertinentes: promedio Dimensión: `PipelineName`
`traceGroupDbSize.value`	Los tamaños de los bytes en memoria de los grupos de trazas en MapDB durante la duración de la ventana actual y anterior. Estadísticas pertinentes: promedio Dimensión: `PipelineName`
`spansDbCount.value`	El recuento de intervalos en MapDB durante la duración de la ventana actual y anterior. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`traceGroupDbCount.value`	El recuento de grupos de trazas en MapDB durante la duración de la ventana actual y anterior. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`relationshipCount.value`	El recuento de las relaciones almacenadas en la duración de la ventana actual y anterior. Estadísticas pertinentes: suma Dimensión: `PipelineName`

OpenSearch métricas

Las siguientes métricas se aplican al OpenSearchsumidero. Cada métrica lleva el nombre de la subcanalización como prefijo y `opensearch`. Por ejemplo, *`sub_pipeline_name`*`.opensearch.bulkRequestErrors.count`.

Sufijo de la métrica	Description (Descripción)
`bulkRequestErrors.count`	El número total de errores encontrados al enviar solicitud es masivas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`documentsSuccess.count`	El número de documentos enviados correctamente al OpenSearch Servicio mediante solicitud masiva, incluidos los reintentos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`documentsSuccessFirstAttempt.count`	El número de documentos enviados correctamente al OpenSearch Servicio mediante solicitud masiva en el primer intento. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`documentErrors.count`	El número de documentos que no se pudieron enviar mediante solicitudes masivas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestFailed.count`	El número de solicitudes masivas que produjeron un error. Estadísticas pertinentes: suma

Sufijo de la métrica	Description (Descripción)
	Dimensión: `PipelineName`
`bulkRequestNumberOfRetries.count`	El número de reintentos de solicitudes masivas fallidas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkBadRequestErrors.count`	El número de errores `Bad Request` encontrados al enviar solicitudes masivas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestNotAllowedErrors.count`	El número de errores `Request Not Allowed` encontrados al enviar solicitudes masivas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestInvalidInputErrors.count`	El número de errores `Invalid Input` encontrados al enviar solicitudes masivas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestNotFoundErrors.count`	El número de errores `Request Not Found` encontrados al enviar solicitudes masivas. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`bulkRequestTimeoutErrors.count`	El número de errores `Request Timeout` encontrados al enviar solicitudes masivas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestServerErrors.count`	El número de errores `Server Error` encontrados al enviar solicitudes masivas. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestSizeBytes.count`	Un recuento de la distribución de los tamaños de carga útil de las solicitudes masivas, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestSizeBytes.sum`	La distribución total de los tamaños de carga útil de las solicitudes masivas, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestSizeBytes.max`	La distribución máxima de los tamaños de carga útil de las solicitudes masivas, en bytes. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`bulkRequestLatency.count`	Un recuento de los puntos de datos registrados mientras se envían las solicitudes al complemento, incluidos los reintentos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestLatency.sum`	La latencia total de las solicitudes enviadas al complemento, incluidos los reintentos, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`bulkRequestLatency.max`	La latencia máxima de las solicitudes enviadas al complemento, incluidos los reintentos, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`
`s3.dlqS3RecordsSuccess.count`	El número de registros que se enviaron correctamente a la cola de mensajes fallidos de S3. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3.dlqS3RecordsFailed.count`	El número de registros que no se pudieron enviar a la cola de mensajes fallidos de S3. Estadísticas pertinentes: suma Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`s3.dlqS3RequestSuccess.count`	El número de solicitudes que se enviaron correctamente a la cola de mensajes fallidos de S3. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3.dlqS3RequestFailed.count`	El número de solicitudes fallidas que se enviaron a la cola de mensajes fallidos de S3. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3.dlqS3RequestLatency.count`	Un recuento de los puntos de datos registrados mientras se envían las solicitudes a la cola de mensajes fallidos de S3, incluidos los reintentos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3.dlqS3RequestLatency.sum`	La latencia total de las solicitudes enviadas a la cola de mensajes fallidos de S3, incluidos los reintentos, en milisegundos. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3.dlqS3RequestLatency.max`	La latencia máxima de las solicitudes enviadas a la cola de mensajes fallidos de S3, incluidos los reintentos, en milisegundos. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Sufijo de la métrica	Description (Descripción)
`s3.dlqS3RequestSizeBytes.count`	Un recuento de la distribución de los tamaños de carga útil de las solicitudes a la cola de mensajes fallidos de S3, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3.dlqS3RequestSizeBytes.sum`	La distribución total de los tamaños de carga útil de las solicitudes a la cola de mensajes fallidos de S3, en bytes. Estadísticas pertinentes: suma Dimensión: `PipelineName`
`s3.dlqS3RequestSizeBytes.max`	La distribución máxima de los tamaños de carga útil de las solicitudes a la cola de mensajes fallidos de S3, en bytes. Estadísticas pertinentes: máximo Dimensión: `PipelineName`

Métricas del sistema y de medición

Las siguientes métricas se aplican a todo el sistema OpenSearch de ingestión. Estas métricas no llevan ningún prefijo.

Métrica	Description (Descripción)
`system.cpu.usage.value`	El porcentaje de uso de CPU disponible para todos los nodos de datos. Estadísticas pertinentes: promedio Dimensión: `PipelineName` , `area`, `id`
`system.cpu.count.value`	La cantidad total de uso de la CPU para todos los nodos de datos.

Métrica	Description (Descripción)
	Estadísticas pertinentes: promedio Dimensión: `PipelineName` , `area`, `id`
`jvm.memory.max.value`	La cantidad máxima de memoria que se puede usar para la administración de la memoria, en bytes. Estadísticas pertinentes: promedio Dimensión: `PipelineName` , `area`, `id`
`jvm.memory.used.value`	La cantidad total de memoria usada, en bytes. Estadísticas pertinentes: promedio Dimensión: `PipelineName` , `area`, `id`signa
`jvm.memory.committ ed.value`	La cantidad de memoria asignada para el uso de la máquina virtual Java (JVM), en bytes. Estadísticas pertinentes: promedio Dimensión: `PipelineName` , `area`, `id`
`computeUnits`	El número de unidades de OpenSearch cómputo de ingestión (ingestión OCUs) que utiliza una canalización. Estadísticas pertinentes: máximo, suma, promedio Dimensión: `PipelineName`

Mejores prácticas para Amazon OpenSearch Ingestion

En este tema se describen las prácticas recomendadas para crear y gestionar los pipelines de Amazon OpenSearch Ingestion y se incluyen directrices generales que se aplican a muchos casos de uso. Cada carga de trabajo es única, con características únicas, por lo que ninguna recomendación genérica es exactamente adecuada para cada caso de uso.

Temas

- Prácticas recomendadas generales
- CloudWatch Alarmas recomendadas

Prácticas recomendadas generales

Las siguientes prácticas recomendadas generales se aplican a la creación y administración de canalizaciones.

- Para garantizar una alta disponibilidad, configure las canalizaciones de VPC con dos o tres subredes. Si solo implementa una canalización en una subred y la zona de disponibilidad deja de funcionar, no podrá incorporar datos.
- Dentro de cada canalización, recomendamos limitar la cantidad de subcanalizaciones a 5 o menos.
- Si utiliza el complemento de fuente de S3, utilice archivos S3 de tamaño uniforme para obtener un rendimiento óptimo.
- Si utiliza el complemento de fuente de S3, añada 30 segundos de tiempo de espera de visibilidad adicional por cada 0,25 GB de tamaño de archivo del bucket de S3 para obtener un rendimiento óptimo.
- Incluya una cola de mensajes fallidos (DLQ) en la configuración de canalización para poder descargar los eventos fallidos y hacerlos accesibles para su análisis. Si los receptores rechazan los datos debido a mapeos incorrectos u otros problemas, puede enrutar los datos al DLQ para solucionar el problema.
- Si utiliza un procesador agregado en una canalización, le recomendamos que utilice `“local_mode: true”` flag para obtener un rendimiento óptimo de la canalización.

CloudWatch Alarmas recomendadas

CloudWatch las alarmas realizan una acción cuando una CloudWatch métrica supera un valor especificado durante un período de tiempo determinado. Por ejemplo, es posible que AWS desee enviarle un correo electrónico si el estado del clúster es `red` superior a un minuto. En esta sección se incluyen algunas alarmas recomendadas para Amazon OpenSearch Ingestion y cómo responder a ellas.

Para obtener más información sobre la configuración de alarmas, consulte Creación de CloudWatch alarmas de Amazon en la Guía del CloudWatch usuario de Amazon.

Alarma	Problema
`computeUnits` máximas es = las `maxUnits` configuradas durante 15 minutos, 3 veces consecutivas	La canalización ha alcanzado su capacidad máxima y es posible que necesite una actualización de `maxUnits`. Aumente la capacidad máxima de su canalización
Suma de `opensearch.documentErrors.count` es = suma de *`{sub_pipeline_name}`*`.opensearch.recordsIn.count` durante 1 minuto, 1 periodo consecutivo	La canalización no puede escribir en el OpenSearch sumidero. Compruebe los permisos de la canalización y confirme que el dominio o la colección están en buen estado. También puede comprobar si hay eventos fallidos en la cola de mensajes fallidos (DLQ), si está configurada.
`bulkRequestLatency.max` máximo es >= x durante 1 minuto, 1 periodo consecutivo	La canalización está experimentando una latencia alta al enviar datos al OpenSearch sumidero. Es probable que esto se deba a que el receptor tiene un tamaño insuficiente o a una mala estrategia de partición, lo que provoca que el receptor quede rezagado. Una latencia alta y sostenida puede afectar el rendimiento de la canalización y es probable que genere una contrapresión para los clientes.
Suma de `httpAuthFailure.count` >= 1 durante 1 minuto, 1 periodo consecutivo	Las solicitudes de ingesta no se autentican. Confirme que todos los clientes tengan habilitada correctamente la autenticación de Signature Version 4.
`system.cpu.usage.value` promedio es >= 80 % durante 15 minutos,	Un uso elevado y sostenido de la CPU puede resultar problemático. Considere aumentar la capacidad máxima de la canalización.

Alarma	Problema
3 periodos consecutivos	
`bufferUsage.value` promedio es >= 80 % durante 15 minutos, 3 periodos consecutivos	Un uso elevado y sostenido del búfer puede resultar problemático. Considere aumentar la capacidad máxima de la canalización.

Otras alarmas para tener en cuenta

Considere la posibilidad de configurar las siguientes alarmas en función de las funciones OpenSearch de Amazon Ingestion que utilice habitualmente.

Alarma	Problema
`dynamodb.exportJobFailure.count` suma 1	Falló el intento de activar una exportación a Amazon S3.
El promedio de `opensearch.EndtoEndLatency.avg` es > X durante 15 minutos, 4 periodos consecutivos	La `EndtoEndLatency` es superior a la deseada para la lectura de las transmisiones de DynamoDB. Esto puede deberse a un OpenSearch clúster con una escala insuficiente o a una capacidad máxima de OCU de canalización demasiado baja para el rendimiento de la WCU de la tabla de DynamoDB. `EndtoEndLatency` será mayor después de una exportación, pero debería disminuir con el tiempo a medida que se vaya actualizando con las últimas transmisiones de DynamoDB.
`dynamodb.changeEventsProcessed.count` suma == 0 durante X minutos	No se está recopilando ningún registro de las transmisiones de DynamoDB. Esto puede deberse a una falta de actividad en la tabla o a un problema al acceder a las transmisiones de DynamoDB.

Alarma	Problema
Suma de `opensearch.s3.dlqS3RecordsSuccess.count` >= suma de `opensearch.documentSuccess.count` durante 1 minuto, 1 periodo consecutivo	Se envía un número mayor de registros al DLQ que al receptor. OpenSearch Revisa las métricas del plugin OpenSearch Sink para investigar y determinar la causa raíz.
Suma de `grok.grokProcessingTimeouts.count` = suma de recordsIn .count durante 1 minuto, 5 períodos consecutivos	Se agota el tiempo de espera de todos los datos mientras el procesador Grok intenta hacer coincidir el patrón. Es probable que esto afecte el rendimiento y ralentice la canalización. Considere ajustar los patrones para reducir los tiempos de espera.
Suma de `grok.grokProcessingErrors.count` es >= 1 durante 1 minuto, 1 periodo consecutivo	El procesador Grok no consigue hacer coincidir los patrones con los datos de la canalización, lo que provoca errores. Revise los datos y las configuraciones del complemento de Grok para asegurarse de que se espera que coincidan los patrones.
Suma de `grok.grokProcessingMismatch.count` = suma de recordsIn .count durante 1 minuto, 5 periodos consecutivos	El procesador Grok no puede hacer coincidir los patrones con los datos de la canalización. Revise los datos y las configuraciones del complemento de Grok para asegurarse de que se espera que coincidan los patrones.

Alarma	Problema
Suma de `date.dateProcessingMatchFailure.count` = suma de recordsIn.count durante 1 minuto, 5 periodos consecutivos	El procesador Date no puede hacer coincidir ningún patrón con los datos de la canalización. Revise los datos y las configuraciones del complemento de Date para asegurarse de que se espera el patrón.
Suma de `s3.s3ObjectsFailed.count` >= 1 durante 1 minuto, 1 periodo consecutivo	Este problema se debe a que el objeto S3 no existe o a que la canalización no tiene privilegios suficientes. Revise las métricas `s3ObjectsNotFound.count` y `s3ObjectsAccessDenied.count` para determinar la causa raíz. Confirme que el objeto S3 existe y and/or actualice los permisos.
Suma de `s3.sqsMessagesFailed.count` >= 1 durante 1 minuto, 1 periodo consecutivo	El complemento S3 no pudo procesar un mensaje de Amazon SQS. Si tiene una DLQ activada en la cola de SQS, revise el mensaje de error. Es posible que la cola esté recibiendo datos no válidos que la canalización está intentando procesar.
Suma de `http.badRequests.count` >= 1 durante 1 minuto, 1 periodo consecutivo	El cliente está enviando una solicitud errónea. Confirme que todos los clientes están enviando la carga útil adecuada.
Suma de `http.requestsTooLarge.count` >= 1 durante 1 minuto, 1 periodo consecutivo	Las solicitudes del complemento de fuente HTTP contienen demasiados datos, lo que supera la capacidad del búfer. Ajuste el tamaño del lote para sus clientes.

Alarma	Problema
Suma de `http.internalServerError.count` >= 0 durante 1 minuto, 1 periodo consecutivo	El complemento de fuente HTTP tiene problemas para recibir eventos.
Suma de `http.requestTimeouts.count` >= 0 durante 1 minuto, 1 periodo consecutivo	Es probable que los tiempos de espera de la fuente se deban a que la canalización está mal aprovisionada. Considere aumentar las `maxUnits` de la canalización para gestionar la carga de trabajo adicional.
Suma de `otel_trace.badRequests.count` >= 1 durante 1 minuto, 1 periodo consecutivo	El cliente está enviando una solicitud errónea. Confirme que todos los clientes están enviando la carga útil adecuada.
Suma de `otel_trace.requestsTooLarge.count` >= 1 durante 1 minuto, 1 periodo consecutivo	Las solicitudes del complemento de fuente de OTel contienen demasiados datos, lo que supera la capacidad del búfer. Ajuste el tamaño del lote para sus clientes.
Suma de `otel_trace.internalServerError.count` >= 0 durante 1 minuto, 1 periodo consecutivo	El complemento de fuente Otel Trace tiene problemas para recibir eventos.

Alarma	Problema
Suma de `otel_trace.requestTimeouts.count` >= 0 durante 1 minuto, 1 periodo consecutivo	Es probable que los tiempos de espera de la fuente se deban a que la canalización está mal aprovisionada. Considere aumentar las `maxUnits` de la canalización para gestionar la carga de trabajo adicional.
Suma de `otel_metrics.requestTimeouts.count` >= 0 durante 1 minuto, 1 periodo consecutivo	Es probable que los tiempos de espera de la fuente se deban a que la canalización está mal aprovisionada. Considere aumentar las `maxUnits` de la canalización para gestionar la carga de trabajo adicional.

Amazon OpenSearch Serverless

Amazon OpenSearch Serverless es una configuración de autoescalado bajo demanda para Amazon OpenSearch Service. A diferencia de OpenSearch los dominios aprovisionados, que requieren una administración manual de la capacidad, una colección OpenSearch Serverless escala automáticamente los recursos de cómputo en función de las necesidades de la aplicación.

OpenSearch Serverless ofrece una solución rentable para cargas de trabajo poco frecuentes, intermitentes o impredecibles. Optimiza los costos al escalar automáticamente la capacidad de cómputo según el uso de su aplicación. Las colecciones sin servidor utilizan el mismo volumen de almacenamiento de alta capacidad, distribuido y de alta disponibilidad que los dominios de servicio aprovisionados. OpenSearch

OpenSearch Las colecciones sin servidor siempre están cifradas. Puede elegir la clave de cifrado, pero no puede desactivar el cifrado. Para obtener más información, consulte the section called “Cifrado”

Ventajas

OpenSearch Serverless tiene las siguientes ventajas:

- Más simple que el aprovisionamiento: OpenSearch sin servidor elimina gran parte de la complejidad de administrar los OpenSearch clústeres y la capacidad. Ajusta y dimensiona de forma automática los clústeres, y se encarga de la administración del ciclo de vida de las particiones y los índices. También gestiona las actualizaciones del software de servicio y las actualizaciones de OpenSearch versiones. Todas las actualizaciones y mejoras no son disruptivas.
- Rentable: cuando utiliza OpenSearch Serverless, solo paga por los recursos que consume. Esto elimina la necesidad del aprovisionamiento inicial y del aprovisionamiento excesivo para las cargas de trabajo máximas.
- Alta disponibilidad: OpenSearch Serverless admite cargas de trabajo de producción con redundancia para proteger contra las interrupciones en las zonas de disponibilidad y los fallos de infraestructura.
- Escalable: OpenSearch Serverless escala automáticamente los recursos para mantener tasas de ingesta de datos y tiempos de respuesta a las consultas consistentemente rápidos.

¿Qué es Amazon OpenSearch Serverless?

Amazon OpenSearch Serverless es una opción sin servidor y bajo demanda para Amazon OpenSearch Service que elimina la complejidad operativa del aprovisionamiento, la configuración y el ajuste de los clústeres. OpenSearch Es ideal para organizaciones que prefieren no administrar sus propios clústeres o que no cuentan con los recursos y el conocimiento experto necesarios para operar implementaciones a gran escala. Con OpenSearch Serverless, puede buscar y analizar grandes volúmenes de datos sin administrar la infraestructura subyacente.

Una colección OpenSearch sin servidor es un grupo de OpenSearch índices que funcionan juntos para respaldar una carga de trabajo o un caso de uso específicos. Las colecciones simplifican las operaciones en comparación con los OpenSearch clústeres autogestionados, que requieren un aprovisionamiento manual.

Las colecciones utilizan el mismo almacenamiento de alta capacidad, distribuido y de alta disponibilidad que los dominios de OpenSearch servicio aprovisionados, pero reducen aún más la complejidad al eliminar la configuración y los ajustes manuales. Los datos de una colección se cifran en tránsito. OpenSearch Serverless también es compatible con los OpenSearch paneles de control, lo que proporciona una interfaz para el análisis de datos.

Actualmente, las colecciones sin servidor funcionan OpenSearch con la versión 2.17.x. A medida que se lanzan nuevas versiones, OpenSearch Serverless actualiza automáticamente las colecciones para incorporar nuevas funciones, correcciones de errores y mejoras de rendimiento.

OpenSearch Serverless admite las mismas operaciones de API de ingesta y consulta que la suite de código OpenSearch abierto, por lo que puede seguir utilizando sus clientes y aplicaciones actuales. Sus clientes deben ser compatibles con la versión OpenSearch 2.x para poder funcionar con Serverless. OpenSearch Para obtener más información, consulte the section called “Ingesta de datos en las colecciones”.

Temas

- Casos de uso de Serverless OpenSearch
- Funcionamiento
- Elección de un tipo de colección
- Precios
- Compatible Regiones de AWS
- Limitaciones

- Comparación entre OpenSearch servicio y sin servidor OpenSearch

Casos de uso de Serverless OpenSearch

OpenSearch Serverless admite dos casos de uso principales:

- Análisis de registros: el segmento de análisis de registros se centra en analizar grandes volúmenes de datos de series temporales semiestructurados y generados por máquinas para obtener información operativa y sobre el comportamiento de los usuarios.
- Búsqueda de texto completo: el segmento de búsqueda de texto completo potencia las aplicaciones de sus redes internas (sistemas de administración de contenido, documentos legales) y las aplicaciones orientadas a Internet, como la búsqueda de contenido de sitios web de comercio electrónico.

Al crear una colección, debe elegir uno de estos casos de uso. Para obtener más información, consulte the section called “Elección de un tipo de colección”.

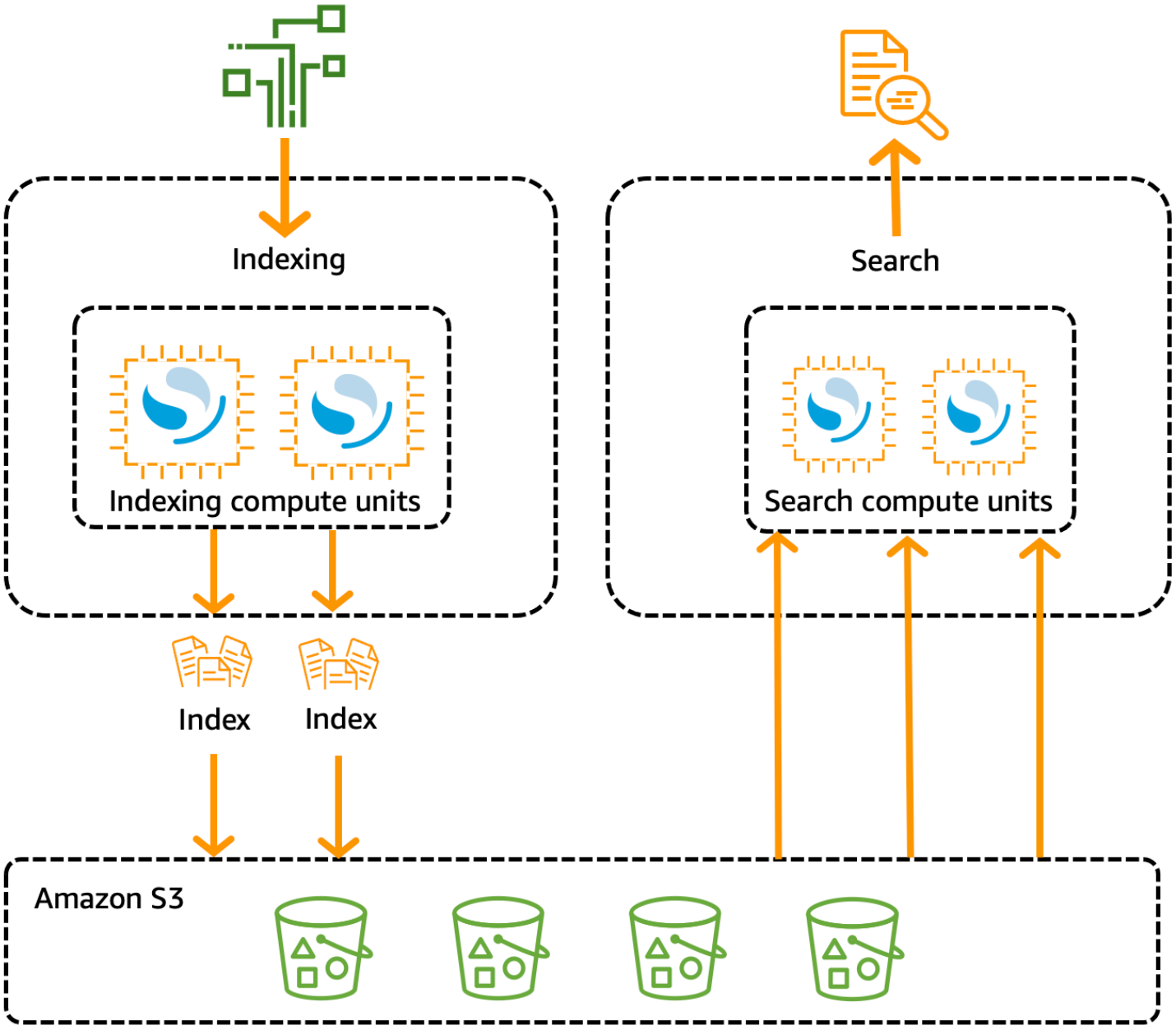
Funcionamiento

OpenSearch Los clústeres tradicionales tienen un único conjunto de instancias que realizan operaciones de indexación y búsqueda, y el almacenamiento de índices está estrechamente relacionado con la capacidad de procesamiento. Por el contrario, OpenSearch Serverless utiliza una arquitectura nativa de la nube que separa los componentes de indexación (ingesta) de los componentes de búsqueda (consulta), con Amazon S3 como almacenamiento de datos principal para los índices.

Esta arquitectura desacoplada permite escalar las funciones de búsqueda e indexación de forma independiente entre ellas y de los datos indexados en S3. La arquitectura también proporciona aislamiento para las operaciones de ingesta y consulta, de modo que puedan ejecutarse de forma simultánea sin contención de recursos.

Cuando escribe datos en una colección, OpenSearch Serverless los distribuye a las unidades informáticas de indexación. Las unidades de computación de indexación ingieren los datos entrantes y mueven los índices a S3. Al realizar una búsqueda en los datos de la recopilación, OpenSearch Serverless dirige las solicitudes a las unidades de cálculo de búsqueda que contienen los datos que se están consultando. Las unidades de computación de búsqueda descargan los datos indexados de forma directa desde S3 (si aún no están almacenados en la memoria caché local), ejecutan operaciones de búsqueda y realizan agregaciones.

La siguiente imagen ilustra esta arquitectura desacoplada:



OpenSearch La capacidad informática sin servidor para la ingesta, búsqueda y consulta de datos se mide en OpenSearch unidades de cálculo (). OCUs Cada OCU es una combinación de 6 GiB de memoria y la CPU virtual (vCPU) correspondiente, así como la transferencia de datos a Amazon S3. Cada OCU incluye suficiente almacenamiento efímero en caliente para 120 GiB de datos de índice.

Cuando creas tu primera colección, OpenSearch Serverless crea dos instancias OCUs: una para indexar y otra para buscar. Para garantizar la alta disponibilidad, también lanza un conjunto de nodos en espera en otra zona de disponibilidad. Para fines de desarrollo y pruebas, puede deshabilitar la configuración Habilitar redundancia para una colección, lo que elimina las dos réplicas en espera

y solo crea instancias de dos. OCUs De forma predeterminada, las réplicas activas redundantes están habilitadas, lo que significa que se OCUs crean instancias de un total de cuatro para la primera recopilación de una cuenta.

OCUs Existen incluso cuando no hay actividad en ningún punto final de la recopilación. Todas las colecciones posteriores las OCUs comparten. Cuando crea colecciones adicionales en la misma cuenta, OpenSearch Serverless solo agrega más OCUs para su búsqueda e ingesta según sea necesario para respaldar las colecciones, de acuerdo con los límites de capacidad que especifique. La capacidad no se reduce verticalmente a medida que disminuye el uso de computación.

Para obtener información sobre cómo se facturan estas OCUs, consulte. the section called “Precios”

Elección de un tipo de colección

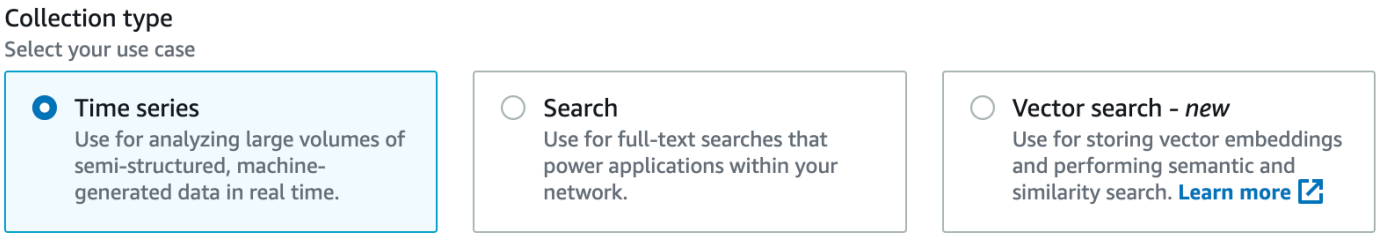
OpenSearch Serverless admite tres tipos de colecciones principales:

Series temporales: el segmento de análisis de registros que analiza grandes volúmenes de datos semiestructurados generados por máquinas en tiempo real y que proporciona información sobre operaciones, seguridad, comportamiento de usuarios y rendimiento del negocio.

Búsqueda: búsqueda de texto completo que habilita aplicaciones dentro de redes internas, como sistemas de administración de contenido y repositorios de documentos legales, así como aplicaciones orientadas a internet, como búsquedas en sitios de comercio electrónico y detección de contenido.

Búsqueda vectorial: la búsqueda semántica basada en incrustaciones vectoriales simplifica la administración de datos vectoriales y habilita experiencias de búsqueda mejoradas con machine learning (ML). Es compatible con aplicaciones de IA generativa, como chatbots, asistentes personales y detección de fraude.

El tipo de colección se elige cuando se crea una colección por primera vez:



El tipo de colección que elija dependerá del tipo de datos que piensa incorporar a la colección y de cómo piensa consultarlos. No puede cambiar el tipo de colección después de crearla.

Los tipos de colecciones presentan las siguientes diferencias notables:

- En el caso de las colecciones de búsqueda y de búsqueda vectorial, todos los datos se almacenan en un almacenamiento en caliente para garantizar tiempos de respuesta rápidos a las consultas. Las colecciones de series temporales utilizan una combinación de almacenamiento en caliente y templado, donde los datos más recientes se guardan en un almacenamiento en caliente para optimizar los tiempos de respuesta a las consultas para los datos a los que se accede con más frecuencia.
- En el caso de las colecciones de series temporales y de búsqueda vectorial, no puede indexar por identificador de documento personalizado ni actualizarlas mediante solicitudes indirectas. Esta operación se reserva para los casos de uso de búsqueda. En su lugar, puede actualizar por ID de documento. Para obtener más información, consulte the section called “Operaciones y permisos de OpenSearch API compatibles”.
- Para las recopilaciones de series temporales y de búsqueda, no puede utilizar índices de tipo k-NN.

Precios

AWS le cobra por los siguientes componentes de OpenSearch Serverless:

- Computación de la ingesta de datos
- Computación de búsquedas y consultas
- Almacenamiento retenido en Amazon S3

Una OCU consta de 6 GB de RAM, la vCPU correspondiente GP3 , el almacenamiento y la transferencia de datos a Amazon S3. La unidad más pequeña por la que se puede facturar es de 0,5 OCU. AWS factura a OCU por horas, con una granularidad de segundo nivel. En tu estado de cuenta, verás una entrada para calcular en OCU-hours con una etiqueta para la ingesta de datos y otra para la búsqueda. AWS también le factura mensualmente los datos almacenados en Amazon S3. No le cobra por usar los OpenSearch paneles de control.

Cuando crea una colección con réplicas activas redundantes, se le factura un mínimo de 1 OCU (0,5 OCU × 2) por la ingesta, incluidas las principales y en espera, y 1 OCU (0,5 OCU × 2) por la búsqueda:

- 1 OCU (0,5 OCU × 2) para ingesta, incluye tanto la primaria como la de respaldo

- 1 OCU (0,5 OCU × 2) para la búsqueda

Si usted deshabilita las réplicas activas redundantes, se le factura un mínimo de 1 OCU (0,5 OCU × 2) para la primera colección en su cuenta. OCUsTodas las colecciones posteriores pueden compartirlas.

OpenSearch Serverless agrega más OCU OCUs en incrementos de 1 OCU en función de la potencia de cómputo y el almacenamiento necesarios para respaldar sus colecciones. Puede configurar un número máximo de OCUs para su cuenta a fin de controlar los costes.

> **Note**
>
> Las colecciones con características únicas no se AWS KMS keys pueden compartir OCUs con otras colecciones.

OpenSearch Serverless intenta utilizar los recursos mínimos necesarios para tener en cuenta los cambios en las cargas de trabajo. El número de OCUs aprovisionados en cualquier momento puede variar y no es exacto. Con el tiempo, el algoritmo que utiliza OpenSearch Serverless seguirá mejorando para minimizar mejor el uso del sistema.

Para obtener información completa sobre los precios, consulta los precios OpenSearch de Amazon Service.

Compatible Regiones de AWS

OpenSearch Serverless está disponible en un subconjunto de Regiones de AWS ese OpenSearch servicio en el que está disponible. Para ver una lista de las regiones compatibles, consulta los puntos de conexión y las cuotas de Amazon OpenSearch Service en. Referencia general de AWS

Limitaciones

OpenSearch Serverless tiene las siguientes limitaciones:

- Algunas operaciones OpenSearch de la API no son compatibles. Consulte the section called “Operaciones y permisos de OpenSearch API compatibles”.
- Algunos OpenSearch complementos no son compatibles. Consulte the section called “OpenSearch Plugins compatibles”.

- Actualmente, no hay forma de migrar automáticamente los datos de un dominio de OpenSearch servicio gestionado a una colección sin servidor. Debe volver a indexar los datos desde un dominio a una colección.
- No se admiten el acceso entre cuentas a las colecciones. No puede incluir colecciones de otras cuentas en las políticas de cifrado o de acceso a los datos.
- No se admiten los OpenSearch complementos personalizados.
- Las instantáneas automatizadas son compatibles con las colecciones OpenSearch sin servidor. No se admiten las instantáneas manuales. Para obtener más información, consulte the section called "Hacer copias de seguridad de las colecciones mediante instantáneas".
- No se admiten la búsqueda y la replicación entre regiones.
- Hay límites para la cantidad de recursos sin servidor que puede tener en una sola cuenta y región. Consulte OpenSearch Cuotas sin servidor.
- El intervalo de actualización de los índices de las colecciones de búsqueda vectorial es de aproximadamente 60 segundos. El intervalo de actualización de los índices de las colecciones de búsqueda y serie temporal es de aproximadamente 10 segundos.
- El número de fragmentos, el número de intervalos y el intervalo de actualización no se pueden modificar y son gestionados por Serverless. OpenSearch La estrategia de partición se basa en el tipo de colección y el tráfico. Por ejemplo, una colección de series temporales escala las particiones principales en función de los cuellos de botella del tráfico de escritura.
- Se admiten las funciones geoespaciales disponibles en OpenSearch las versiones anteriores a la 2.1.

Comparación entre OpenSearch servicio y sin servidor OpenSearch

En OpenSearch Serverless, algunos conceptos y características son diferentes a la característica correspondiente para un dominio de servicio aprovisionado OpenSearch . Por ejemplo, una diferencia importante es que OpenSearch Serverless no tiene el concepto de clúster o nodo.

En la siguiente tabla se describe en qué se diferencian las funciones y los conceptos importantes de OpenSearch Serverless de la función equivalente de un dominio de servicio aprovisionado OpenSearch .

Característica	OpenSearch Servicio	OpenSearch Sin servidor
Dominios frente a colecciones	Los índices se guardan en dominios, que son clústeres aprovisionados previamente OpenSearch . Para obtener más información, consulte Creación y administración de dominios.	Los índices se encuentran en colecciones, que son agrupaciones lógicas de índices que representan una carga de trabajo o un caso de uso específicos. Para obtener más información, consulte the section called “Administración de colecciones”.
Tipos de nodos y administración de la capacidad	Crea un clúster con tipos de nodos que cumplen con sus especificaciones de costo y rendimiento. Debe calcular sus propios requisitos de almacenamiento y elegir un tipo de instancia para su dominio. Para obtener más información, consulte the section called “Determinación del tamaño de dominios”.	OpenSearch Serverless escala y aprovisiona automáticamente unidades de cómputo adicionales para su cuenta en función del uso de la capacidad. Para obtener más información, consulte the section called “Administración de los límites de capacidad”.
Facturación	Paga por cada hora de uso de una instancia EC2 y por el tamaño acumulado de cualquier volumen de almacenamiento de EBS adjunto a sus instancias. Para obtener más información, consulte the section called “Precios”.	Se le cobrará en horas de OCU por la computación para la ingesta de datos, la computación para las búsquedas y las consultas y el almacenamiento retenido en S3. Para obtener más información, consulte the section called “Precios”.
Cifrado	El cifrado en reposo es opcional para los dominios.	El cifrado en reposo es obligatorio para las colecciones.

Característica	OpenSearch Servicio	OpenSearch Sin servidor
	Para obtener más informaci ón, consulte the section called “Cifrado en reposo”.	Para obtener más información, consulte the section called “Cifrado”.
Control de acceso a los datos	El acceso a los datos dentro de los dominios se determina mediante políticas de IAM y un control de acceso detallado.	El acceso a los datos dentro de las colecciones se determina mediante las políticas de acceso a los datos.
Operaciones compatibles OpenSearch	OpenSearch El servicio admite un subconjunto de todas las operaciones de la OpenSearch API. Para obtener más informaci ón, consulte the section called “Operaciones admitidas”.	OpenSearch Serverless admite un subconjunto diferente de operaciones de OpenSearch API. Para obtener más información, consulte the section called “Operaciones y complementos compatibles”.
Inicio de sesión en Dashboards	Inicio de sesión con un nombre de usuario y una contraseña. Para obtener más informaci ón, consulte the section called “Acceder a los OpenSearc h paneles de control como usuario maestro”.	Si has iniciado sesión en la AWS consola y accedes a la URL de tu panel de control, iniciarás sesión automáticamente. Para obtener más información, consulte the section called “Acceso a OpenSearch los paneles”.
APIs	Interactúa programáticamente con el OpenSearch Servicio mediante las operaciones de la API del OpenSearch Servicio.	Interactúe mediante programación con OpenSearch Serverless mediante las operaciones de la API Serverless. OpenSearch

Característica	OpenSearch Servicio	OpenSearch Sin servidor
Acceso a la red	La configuración de red de un dominio se aplica tanto al punto final del dominio como al punto final de Dashboards. OpenSearch El acceso a la red para ambos está vinculado de forma estrecha.	La configuración de red para el punto final del dominio y el punto final de OpenSearch Dashboards está disociada. Puede optar por no configurar el acceso a la red para los OpenSearch paneles. Para obtener más información, consulte the section called “Acceso a la red”.
Firma de solicitudes	Utilice los clientes REST de nivel OpenSearch alto y bajo para firmar las solicitudes. Especifique el nombre del servicio como `es`.	En este momento, OpenSearch Serverless admite un subconjunto de clientes compatibles OpenSearch con Service. Al firmar las solicitudes, especifique el nombre del servicio como `aoss`. El encabezado `x-amz-content-sha256` es obligatorio. Para obtener más información, consulte the section called “Otros clientes”.
OpenSearch actualizaciones de versión	Los dominios se actualizan manualmente a medida que hay nuevas versiones OpenSearch disponibles de. Es responsable de garantizar que su dominio cumpla con los requisitos de actualización y de que solucionar cualquier cambio importante.	OpenSearch Serverless actualiza automáticamente tus colecciones a nuevas OpenSearch versiones. Las actualizaciones no siempre se producen en cuanto está disponible una nueva versión.
Actualizaciones del software del servicio	De forma manual, aplique las actualizaciones del software del servicio a su dominio a medida que estén disponibles.	OpenSearch Serverless actualiza automáticamente sus colecciones para incluir las últimas correcciones de errores, funciones y mejoras de rendimiento.

Característica	OpenSearch Servicio	OpenSearch Sin servidor
Acceso mediante VPC	Puede aprovisionar su dominio dentro de una VPC. También puede crear puntos finales de OpenSearch VPC gestionados por el servicio adicionales para acceder al dominio.	Usted crea uno o más puntos de enlace de VPC OpenSearch gestionados por Serverless para su cuenta. A continuación, incluya estos puntos de conexión dentro de las políticas de red.
Autenticación SAML	La autenticación SAML se habilita por dominio. Para obtener más informaci ón, consulte the section called “Autenticación SAML para paneles OpenSearch”.	Configura uno o más proveedores de SAML a nivel de cuenta y, a continuación, incluye el usuario y el grupo asociados en las políticas de acceso a los datos. IDs Para obtener más información, consulte the section called “Autenticación SAML”.
Transport Layer Security (TLS)	OpenSearch El servicio es compatible con TLS 1.2, pero se recomienda usar TLS 1.3.	OpenSearch Serverless es compatible con TLS 1.2, pero se recomienda usar TLS 1.3.

Tutorial: Introducción a Amazon OpenSearch Serverless

En este tutorial, se explican los pasos básicos para poner en marcha rápidamente una colección de búsquedas de Amazon OpenSearch Serverless. Una colección de búsqueda le permite potenciar las aplicaciones de sus redes internas y las aplicaciones orientadas a Internet, como la búsqueda en sitios web de comercio electrónico y la búsqueda de contenido.

Para obtener información sobre cómo utilizar una colección de búsqueda vectorial, consulte. the section called “Trabajo con colecciones de búsqueda vectorial” Para obtener información detallada sobre el uso de colecciones, consulte the section called “Administración de colecciones” y los demás temas de esta guía.

En este tutorial, deberá completar los siguientes pasos:

1. Configurar permisos
2. Crear una colección
3. Cargar y buscar datos
4. Eliminar la colección

> **Note**
>
> Recomendamos que usted utilice únicamente caracteres ASCII para su `IndexName`. Si no utiliza caracteres ASCII`IndexName`, las CloudWatch métricas de entrada se convertirán a un formato codificado `IndexName` en URL para caracteres que no sean ASCII.

Paso 1: configurar permisos

Para completar este tutorial y utilizar OpenSearch Serverless en general, debe tener los permisos de IAM correctos. En este tutorial, creará una colección, cargará y buscará datos y, a continuación, eliminará la colección.

Su usuario o rol debe tener adjunta una política basada en la identidad con los siguientes permisos mínimos:

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": "*"
    }
  ]
}
```

Para obtener más información sobre los permisos de IAM OpenSearch sin servidor, consulte. the section called “Gestión de identidad y acceso”

Paso 2: crear una colección

Una colección es un grupo de OpenSearch índices que funcionan juntos para respaldar una carga de trabajo o un caso de uso específicos.

Para crear una colección OpenSearch sin servidor

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Seleccione Colecciones en el panel de navegación de la izquierda y elija Crear colección.
3. Asigne un nombre a las películas de la colección.
4. Para el tipo de colección, seleccione Buscar. Para obtener más información, consulte Choosing a collection type (Elección de un tipo de recopilación).
5. En Seguridad, elija Creación estándar.
6. En Cifrado, seleccione Utilizar Clave propiedad de AWS. Esto es AWS KMS key lo que OpenSearch Serverless utilizará para cifrar sus datos.
7. En Red, configure los ajustes de red para la colección.
   - Para el tipo de acceso, seleccione Público.
   - Para el tipo de recurso, elija Habilitar el acceso a los OpenSearch puntos finales y Habilitar el acceso a los paneles. OpenSearch Como cargará y buscará datos mediante los OpenSearch paneles de control, debe habilitar ambas opciones.
8. Elija Siguiente.
9. En Configurar el acceso a los datos, defina los ajustes de acceso a la colección. Las políticas de acceso a los datos le permiten a los usuarios y a los roles acceder a los datos de una colección. En este tutorial, le brindaremos a un solo usuario los permisos necesarios para indexar y buscar datos en la colección de películas.

Cree una sola regla que proporcione acceso a la colección películas. Asígnele a la regla el nombre Acceso a la colección de películas.

10. Elija Agregar directores, usuarios y roles de IAM y seleccione el usuario o rol que usará para iniciar sesión en los OpenSearch paneles e indexar datos. Seleccione Save.
11. En Indexar permisos, seleccione todos los permisos.
12. Seleccione Siguiente.
13. En la configuración de la política de acceso, seleccione Crear una nueva política de acceso a datos y póngale el nombre películas a la política.
14. Seleccione Siguiente.
15. Revise la configuración de la colección y seleccione Enviar. Espere unos minutos hasta que el estado de la colección cambie a `Active`.

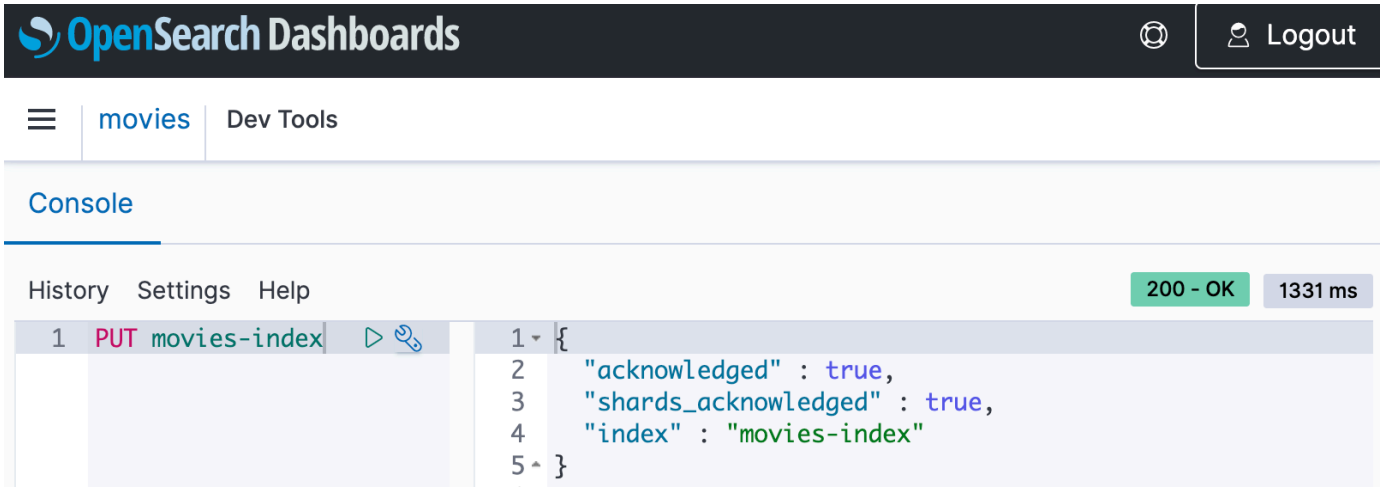
Paso 3: cargar y buscar datos

Puedes cargar datos a una colección OpenSearch sin servidor mediante Postman o cURL. Para abreviar, en estos ejemplos se utilizan las herramientas de desarrollo de la consola de Dashboards. OpenSearch

Para indexar y buscar datos en la colección de películas

1. Seleccione Colecciones en el panel de navegación de la izquierda y elija la colección de películas para abrir su página de detalles.
2. Elija la URL de los OpenSearch paneles de control para la colección. La URL toma el formato `https://dashboards.{region}.aoss.amazonaws.com/_login/?collectionId={collection-id}`.
3. En los OpenSearch paneles, abra el panel de navegación izquierdo y elija Dev Tools.
4. Para crear un índice único llamado índice de películas, envíe la siguiente solicitud:

```
PUT movies-index
```



5. Para indexar un solo documento en índice de películas, envíe la siguiente solicitud:

```
PUT movies-index/_doc/1
{
  "title": "Shawshank Redemption",
  "genre": "Drama",
  "year": 1994
}
```

6. Para buscar datos en los OpenSearch paneles, debe configurar al menos un patrón de índice. OpenSearch utiliza estos patrones para identificar los índices que desea analizar. Abra el menú principal del panel, seleccione Administración de pilas. Luego, seleccione Patrones de índices y, a continuación, seleccione Crear patrón de índice. Para este tutorial, escriba Películas.
7. seleccione Siguiente paso y, a continuación, seleccione Crear patrón de índice. Una vez creado el patrón, puede ver los diversos campos de documento, como `title` y `genre`.
8. Para comenzar a buscar los datos, abra el panel de navegación izquierdo y elija Explorar o utilice la API de búsqueda dentro de Herramientas de desarrollo.

Gestión de errores

Al ejecutar operaciones de indexación y búsqueda, usted puede recibir las siguientes respuestas de error:

- `HTTP 507`: indica que se produjo un error interno del servidor. Este error generalmente indica que sus unidades de OpenSearch cómputo (OCUs) están sobrecargadas por el volumen o la

complejidad de sus solicitudes. Si bien OpenSearch Serverless escala automáticamente para gestionar la carga, puede producirse un retraso en la implementación de recursos adicionales.

Para mitigar este error, implemente una política de reintentos con retroceso exponencial. Este enfoque reduce temporalmente la tasa de solicitudes para administrar la carga de manera eficaz. Para obtener más información, consulte el comportamiento de los reintentos en la Guía de referencia de herramientas AWS SDKs y herramientas.

- `HTTP 402`— Indica que ha alcanzado el límite máximo de capacidad de la unidad de OpenSearch cómputo (OCU). Optimice su carga de trabajo para reducir el uso de OCU o solicite un aumento de cuota.

Paso 4: eliminar la colección

Dado que la colección de películas es para hacer pruebas, debe eliminarla cuando termine de experimentar.

Para eliminar una colección OpenSearch sin servidor

1. Vuelve a la consola OpenSearch de Amazon Service.
2. Seleccione Colecciones en el panel de navegación de la izquierda y elija la colección de películas.
3. Seleccione Eliminar y confirme la eliminación.

Siguientes pasos

Ahora que sabe cómo crear una colección e indexar datos, puede que desee probar algunos de los siguientes ejercicios:

- Consulte opciones más avanzadas para crear una colección. Para obtener más información, consulte the section called “Administración de colecciones”.
- Aprenda a configurar las políticas de seguridad para gestionar la seguridad de las colecciones a escala. Para obtener más información, consulte the section called “Seguridad en Serverless OpenSearch ”.
- Descubra otras formas de indexar datos en las colecciones. Para obtener más información, consulte the section called “Ingesta de datos en las colecciones”.

Colecciones Amazon OpenSearch Serverless

Una colección en Amazon OpenSearch Serverless es una agrupación lógica de uno o más índices que representan una carga de trabajo de análisis. OpenSearch Serverless administra y ajusta automáticamente la recopilación, lo que requiere una intervención manual mínima.

Temas

- Administración de colecciones de Amazon OpenSearch Serverless
- Trabajo con colecciones de búsqueda vectorial
- Uso de políticas del ciclo de vida de los datos con Amazon OpenSearch Serverless
- Uso del AWS SDKs para interactuar con Amazon OpenSearch Serverless
- Uso CloudFormation para crear colecciones de Amazon OpenSearch Serverless
- Hacer copias de seguridad de las colecciones mediante instantáneas
- Soporte de códecs Zstandard en Amazon Serverless OpenSearch
- Ahorre almacenamiento mediante el uso de una fuente derivada

Administración de colecciones de Amazon OpenSearch Serverless

Una colección en Amazon OpenSearch Serverless es una agrupación lógica de uno o más índices que representan una carga de trabajo de análisis. OpenSearch Serverless administra y ajusta automáticamente la recopilación, lo que requiere una intervención manual mínima.

Temas

- Configuración de permisos para las colecciones
- Enriquecimiento semántico automático para Serverless
- Creación de colecciones
- Acceso a OpenSearch los paneles
- Visualización de colecciones
- Eliminación de colecciones

Configuración de permisos para las colecciones

OpenSearch Serverless utiliza los siguientes permisos AWS Identity and Access Management (IAM) para crear y administrar colecciones. Puede especificar las condiciones de IAM para restringir a los usuarios a colecciones específicas.

- `aoss:CreateCollection`: cree una colección.
- `aoss:ListCollections`: enumere las colecciones en la cuenta actual.
- `aoss:BatchGetCollection`: obtenga detalles sobre una o más colecciones.
- `aoss:UpdateCollection`: modifique una colección.
- `aoss:DeleteCollection`: elimine una colección.

El siguiente ejemplo de política de acceso basada en la identidad proporciona los permisos mínimos necesarios para que un usuario administre una única colección denominada `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` y `aoss:CreateSecurityPolicy` se incluyen porque se requieren las políticas de cifrado, de red y de acceso a los datos para que una colección funcione de forma correcta. Para más información, consulte the section called “Gestión de identidad y acceso”.

> **Note**
>
> Al crear la primera colección en su cuenta, también necesitará el permiso `iam:CreateServiceLinkedRole`. Para obtener más información, consulte the section called “Rol de creación de colecciones”.

Enriquecimiento semántico automático para Serverless

Introducción

La función de enriquecimiento semántico automático puede ayudar a mejorar la relevancia de la búsqueda hasta en un 20% con respecto a la búsqueda léxica. El enriquecimiento semántico automático elimina el pesado trabajo indiferenciado que supone gestionar tu propia infraestructura de modelos de aprendizaje automático (ML) y la integración con el motor de búsqueda. La función está disponible para los tres tipos de colecciones sin servidor: búsqueda, series temporales y vectoriales.

¿Qué es la búsqueda semántica

Los motores de búsqueda tradicionales se basan en la word-to-word coincidencia (denominada búsqueda léxica) para encontrar los resultados de las consultas. Aunque esto funciona bien para consultas específicas, como números de modelos de televisores, tiene dificultades para realizar búsquedas más abstractas. Por ejemplo, al buscar «zapatos para la playa», una búsqueda léxica solo encuentra las palabras individuales «zapatos», «playa», «para» y «el» en los artículos del catálogo, y es posible que no aparezcan productos relevantes como «sandalias resistentes al agua» o «calzado de surf» que no contienen los términos de búsqueda exactos.

La búsqueda semántica arroja resultados de consulta que incorporan no solo la coincidencia de palabras clave, sino también la intención y el significado contextual de la búsqueda del usuario. Por ejemplo, si un usuario busca «cómo tratar un dolor de cabeza», es posible que un sistema de búsqueda semántica arroje los siguientes resultados:

- Remedios para la migraña
- Técnicas de manejo del dolor
- Over-the-counter analgésicos

Detalles del modelo y punto de referencia de rendimiento

Si bien esta función resuelve las complejidades técnicas entre bastidores sin exponer el modelo subyacente, ofrecemos transparencia mediante una breve descripción del modelo y resultados comparativos para ayudarlo a tomar decisiones informadas sobre la adopción de funciones en sus cargas de trabajo críticas.

El enriquecimiento semántico automático utiliza un modelo disperso prediseñado y gestionado por el servicio que funciona de forma eficaz sin necesidad de ajustes personalizados. El modelo analiza los campos que especificas y los expande en vectores dispersos en función de las asociaciones aprendidas a partir de diversos datos de entrenamiento. Los términos ampliados y sus ponderaciones de significancia se almacenan en un formato de índice nativo de Lucene para una recuperación eficiente. Hemos optimizado este proceso mediante el modo de solo documentos, en el que la codificación solo se realiza durante la ingesta de datos. Las consultas de búsqueda simplemente se tokenizan en lugar de procesarse mediante un modelo disperso, lo que hace que la solución sea rentable y eficaz.

Nuestra validación del rendimiento durante el desarrollo de las funciones utilizó el conjunto de datos de recuperación de pasajes de MS MARCO, que incluía pasajes con un promedio de 334 caracteres. Para obtener una puntuación de relevancia, medimos la ganancia acumulada descontada normalizada (NDCG) de los 10 primeros resultados de búsqueda (ndcg @10) según el índice BEIR para contenido en inglés y la media ndcg @10 según MIRACL para contenido multilingüe. Evaluamos la latencia mediante mediciones del percentil 90 (p90) del lado del cliente, y la respuesta de búsqueda p90 tomó valores. Estos puntos de referencia proporcionan indicadores de rendimiento de referencia tanto para la relevancia de la búsqueda como para los tiempos de respuesta. Estas son las principales cifras de referencia:

- Idioma inglés: mejora la relevancia del 20% con respecto a la búsqueda léxica. También redujo la latencia de búsqueda del P90 en un 7,7% con respecto a la búsqueda léxica (BM25 es de 26 ms y el enriquecimiento semántico automático es de 24 ms).
- Multilingüe: la relevancia mejoró un 105% con respecto a la búsqueda léxica, mientras que la latencia de la búsqueda del P90 aumentó un 38,4% con respecto a la búsqueda léxica (BM25 es de 26 ms y el enriquecimiento semántico automático es de 36 ms).

Dada la naturaleza única de cada carga de trabajo, le recomendamos que evalúe esta función en su entorno de desarrollo utilizando sus propios criterios de evaluación comparativa antes de tomar decisiones de implementación.

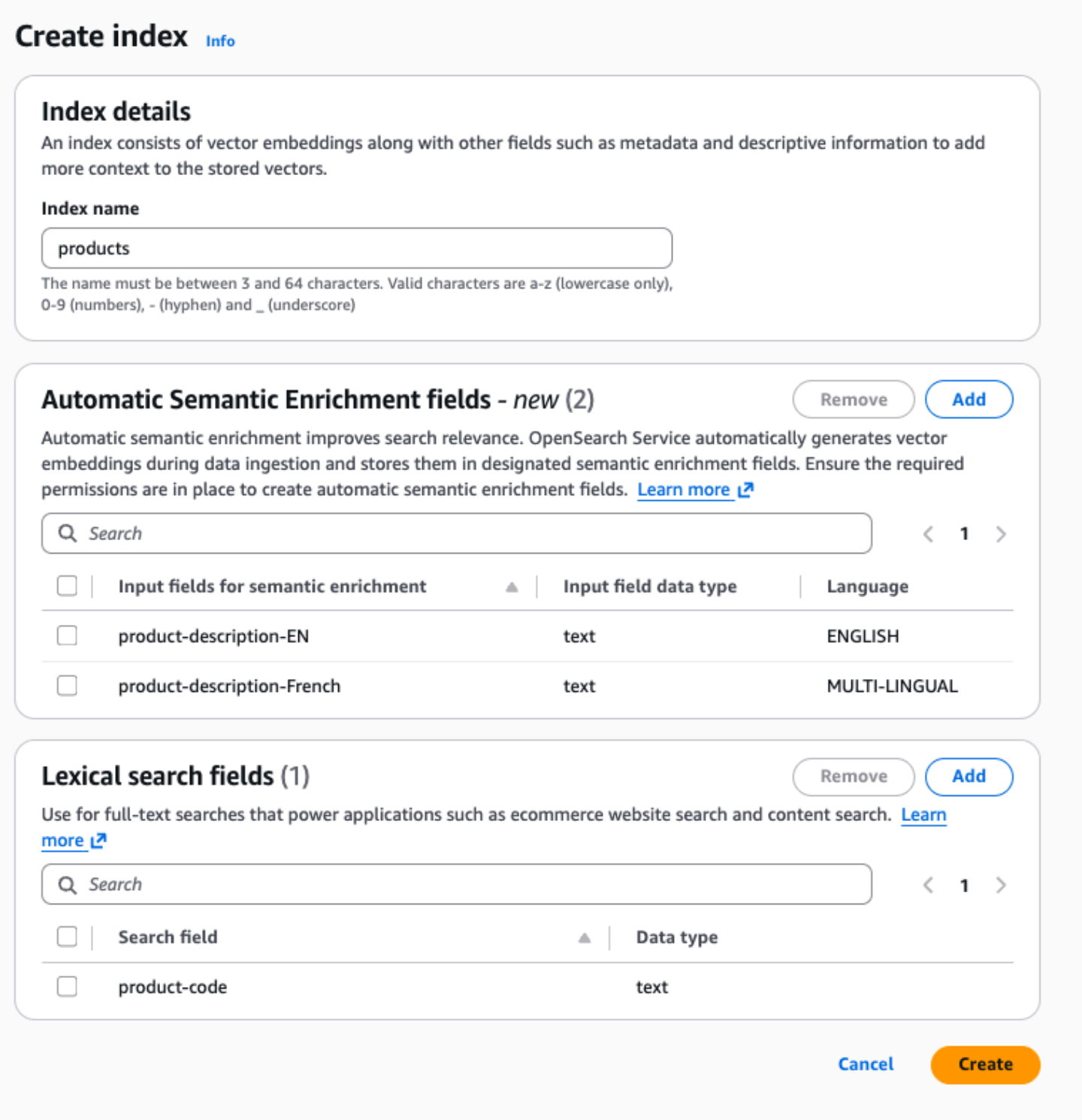
Idiomas compatibles

La función es compatible con el inglés. Además, el modelo también admite árabe, bengalí, chino, finés, francés, hindi, indonesio, japonés, coreano, persa, ruso, español, swahili y telugu.

Configure un índice de enriquecimiento semántico automático para las colecciones sin servidor

Configurar un índice con el enriquecimiento semántico automático activado para sus campos de texto es fácil y puede gestionarlo mediante la consola y las CloudFormation plantillas durante la creación de un APIs nuevo índice. Para habilitarlo en un índice existente, debe volver a crear el índice con el enriquecimiento semántico automático activado para los campos de texto.

Experiencia de consola: la AWS consola le permite crear fácilmente un índice con campos de enriquecimiento semántico automático. Cuando selecciones una colección, encontrarás el botón de creación de índice en la parte superior de la consola. Al hacer clic en el botón de creación de índice, encontrará opciones para definir campos de enriquecimiento semántico automáticos. En un índice, puede tener combinaciones de enriquecimiento semántico automático para campos en inglés y multilingüe, así como para campos léxicos.



Experiencia con las API: para crear un índice de enriquecimiento semántico automático mediante la interfaz de línea de comandos (AWS CLI), utilice el AWS comando create-index:

```
aws opensearchserverless create-index \
--id [collection_id] \
```

```
--index-name [index_name] \
--index-schema [index_body] \
```

En el siguiente ejemplo de esquema de índice, el campo title_semantic tiene un tipo de campo establecido en texto y el parámetro semantic_enrichment establecido en el estado ENABLED. Si se establece el parámetro semantic_enrichment, se habilita el enriquecimiento semántico automático en el campo title_semantic. Puede usar el campo language_options para especificar el inglés o el multilingüe.

```
    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"
            }
        }
    }
}'
```

Para describir el índice creado, utilice el siguiente comando:

```
aws opensearchserverless get-index \
--id [collection_id] \
--index-name [index_name] \
```

También puede usar CloudFormation plantillas (tipo:AWS::OpenSearchServerless::CollectionIndex) para crear una búsqueda semántica durante el aprovisionamiento de la colección y después de crearla.

Ingestión y búsqueda de datos

Una vez que haya creado un índice con el enriquecimiento semántico automático activado, la función funcionará automáticamente durante el proceso de ingesta de datos, sin necesidad de realizar ninguna configuración adicional.

Ingesta de datos: al añadir documentos al índice, el sistema automáticamente:

- Analiza los campos de texto que designó para el enriquecimiento semántico
- Genera codificaciones semánticas mediante OpenSearch un modelo disperso gestionado por el servicio
- Almacena estas representaciones enriquecidas junto con los datos originales

Este proceso utiliza los conectores OpenSearch de aprendizaje automático integrados y las canalizaciones de ingesta, que se crean y administran automáticamente entre bastidores.

Búsqueda: los datos de enriquecimiento semántico ya están indexados, por lo que las consultas se ejecutan de manera eficiente sin tener que volver a invocar el modelo de aprendizaje automático. Esto significa que obtiene una mayor relevancia de búsqueda sin sobrecargar la latencia de búsqueda adicional.

Configuración de permisos para el enriquecimiento semántico automático

Antes de crear un índice de enriquecimiento semántico automatizado, debe configurar los permisos necesarios. En esta sección, se explican los permisos necesarios y cómo configurarlos.

Permisos de política de IAM

Utilice la siguiente política AWS Identity and Access Management (IAM) para conceder los permisos necesarios para trabajar con el enriquecimiento semántico automático:

JSON

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

```
        {
            "Sid": "AutomaticSemanticEnrichmentPermissions",
            "Effect": "Allow",
            "Action": [
                "aoss:CreateIndex",
                "aoss:GetIndex",
                "aoss:UpdateIndex",
                "aoss:DeleteIndex",
                "aoss:APIAccessAll"
            ],
            "Resource": "*"
        }
    ]
}
```

Permisos de claves

- Los permisos `aoss:*Index` permiten la administración de índices.
- El `aoss:APIAccessAll` permiso permite las operaciones de la API OpenSearch
- Para restringir los permisos a una colección específica, reemplace `"Resource": "*"` con el ARN de la colección.

Configurar los permisos de acceso a datos

A fin de configurar un índice para el enriquecimiento semántico automático, debe contar con las políticas de acceso a datos adecuadas que concedan permiso de acceso a los recursos de índices, canalizaciones y colecciones de modelos. Para obtener más información acerca de las políticas de acceso a datos, consulte Control de acceso a datos para Amazon OpenSearch Serverless. Para obtener información acerca del procedimiento de configuración de una política de acceso a datos, consulte Creación de políticas de acceso a los datos (consola).

Permisos de acceso a datos

```
[
    {
        "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"
        ]
    },
]
```

Permisos de acceso a la red

Para permitir que el servicio acceda APIs a las colecciones privadas, debe configurar políticas de red que permitan el acceso necesario entre la API del servicio y la colección. Para obtener más información sobre las políticas de red, consulte Acceso a la red para 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"
      ],
   }
]
```

Configuración de los permisos de acceso a la red para una colección privada

1. Inicie sesión en la consola de OpenSearch servicio de https://console.aws.amazon.com/aos/sucasa.
2. En el panel de navegación de la izquierda, seleccione Políticas de red. A continuación, realice alguna de las operaciones siguientes:
   - Elija un nombre de política existente y seleccione Editar.
   - Elija Crear política de red y configure los detalles de la política.
3. En el área Tipo de acceso, seleccione Privado (recomendado) y, a continuación, seleccione Acceso privado al servicio de AWS .
4. En el campo de búsqueda, seleccione Servicio y, a continuación, aoss.amazonaws.com.
5. En el área Tipo de recurso, selecciona la casilla Habilitar el acceso al OpenSearch punto final.

6. Para Buscar en colecciones o introducir términos prefijos específicos, en el campo de búsqueda, seleccione Nombre de la colección. A continuación, escriba o seleccione el nombre de las colecciones que desee asociar a la política de red.
7. Seleccione Crear para una nueva política de red o Actualizar para una política de red existente.

La consulta se reescribe

El enriquecimiento semántico automático convierte automáticamente las consultas «coincidentes» existentes en consultas de búsqueda semántica sin necesidad de modificar las consultas. Si una consulta de coincidencia forma parte de una consulta compuesta, el sistema recorre la estructura de la consulta, busca las consultas coincidentes y las reemplaza por consultas de dispersión neuronal. Actualmente, la función solo permite reemplazar las consultas «coincidentes», tanto si se trata de una consulta independiente como si forma parte de una consulta compuesta. No se admite la opción «multi_match». Además, la función admite todas las consultas compuestas para reemplazar sus consultas de coincidencia anidadas. Las consultas compuestas incluyen: bool, boost, constant_score, dis_max, function_score e hybrid.

Limitaciones del enriquecimiento semántico automático

La búsqueda semántica automática es más eficaz cuando se aplica a campos small-to-medium dimensionados que contienen contenido en lenguaje natural, como títulos de películas, descripciones de productos, reseñas y resúmenes. Si bien la búsqueda semántica aumenta la relevancia en la mayoría de los casos de uso, puede que no sea óptima en algunos escenarios. Ten en cuenta las siguientes limitaciones a la hora de decidir si vas a implementar el enriquecimiento semántico automático para tu caso de uso específico.

- Documentos muy largos: el modelo disperso actual procesa solo los primeros 8.192 símbolos de cada documento en inglés. En el caso de los documentos multilingües, se trata de 512 fichas. En el caso de artículos extensos, considere la posibilidad de implementar la fragmentación de documentos para garantizar un procesamiento completo del contenido.
- Cargas de trabajo de análisis de registros: el enriquecimiento semántico aumenta significativamente el tamaño del índice, lo que puede resultar innecesario en el análisis de registros, donde la coincidencia exacta suele ser suficiente. El contexto semántico adicional rara vez mejora la eficacia de la búsqueda de registros lo suficiente como para justificar el aumento de los requisitos de almacenamiento.
- El enriquecimiento semántico automático no es compatible con la función Fuente derivada.

- Limitación: las solicitudes de inferencia de indexación actualmente tienen un límite de 100 TPS para sistemas sin servidor. OpenSearch Este es un límite flexible; ponte en contacto con AWS Support para obtener límites más altos.

Precios

OpenSearch Serverless factura el enriquecimiento semántico automático en función de las unidades de OpenSearch cálculo (OCUs) consumidas durante la generación de vectores dispersos en el momento de la indexación. Solo se le cobrará por el uso real durante la indexación. Puedes monitorizar este consumo mediante la SemanticSearch OCU CloudWatch métrica de Amazon. Para obtener detalles específicos sobre los límites de los tokens del modelo, el volumen de procesamiento por OCU y un ejemplo de cálculo, visita OpenSearch Precios de servicios.

Creación de colecciones

Puede usar la consola o la AWS CLI para crear una colección sin servidor. En estos pasos se explica cómo crear una búsqueda o una colección de series temporales. Para crear una colección de búsquedas vectoriales, consulte the section called “Trabajo con colecciones de búsqueda vectorial”.

Temas

- Crear una colección (consola)
- Para crear una colección (CLI)

Crear una colección (consola)

Utilice los procedimientos de esta sección para crear una colección mediante la Consola de administración de AWS. En estos pasos se explica cómo crear una búsqueda o una colección de series temporales. Para crear una colección de búsquedas vectoriales, consulte the section called “Trabajo con colecciones de búsqueda vectorial”.

Temas

- Configurar los ajustes de la colección
- Configurar campos de búsqueda adicionales

Configurar los ajustes de la colección

Utilice el siguiente procedimiento para configurar la información de la colección.

Para configurar los ajustes de la colección mediante la consola

1. Ve a la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/casa/.
2. Expanda Sin servidor en el panel de navegación de la izquierda y seleccione Colecciones.
3. Seleccione Crear colección.
4. Escriba un nombre y una descripción para la colección. El nombre debe cumplir los siguientes criterios:

   - Es exclusivo de tu cuenta y Región de AWS
   - Contener solo letras minúsculas de la a a la z, números del 0 al 9 y un guion (-)
   - Contener entre 3 y 32 caracteres
5. Seleccione un tipo de colección:

   - Series temporales: segmento de análisis de registros que se centra en analizar grandes volúmenes de datos semiestructurados generados por máquinas. Se almacenan al menos 24 horas de datos en los índices calientes y el resto permanece en almacenamiento templado.
   - Búsqueda: búsqueda de texto completo que potencia las aplicaciones de sus redes internas y las aplicaciones orientadas a Internet. Todos los datos de búsqueda se almacenan en almacenamiento en caliente para garantizar un tiempo de respuesta rápido a las consultas.

     > **Note**
     >
     > Elija esta opción si va a habilitar la búsqueda semántica automática, tal y como se describe en Configurar los ajustes de la colección.
   - Búsqueda vectorial: búsqueda semántica en incrustaciones vectoriales que simplifica la administración de datos vectoriales. Impulsa las experiencias de búsqueda aumentada con machine learning (ML) y las aplicaciones de IA generativa, como los chatbots, los asistentes personales y la detección de fraudes.

   Para obtener más información, consulte the section called “Elección de un tipo de colección”.
6. En Tipo de implementación, elija la configuración de redundancia de la colección. De forma predeterminada, cada colección tiene redundancia, lo que significa que cada una de las unidades de OpenSearch cálculo de indexación y búsqueda (OCUs) tiene sus propias réplicas en espera en una zona de disponibilidad diferente. Para fines de desarrollo y pruebas, puedes

optar por deshabilitar la redundancia, lo que reduce el número de unidades de OCUs tu colección a dos. Para obtener más información, consulte the section called “Funcionamiento”.

7. En Seguridad, elija Creación estándar.
8. Para el cifrado, elija una AWS KMS clave con la que cifrar los datos. OpenSearch Serverless le notifica si el nombre de la colección que ha introducido coincide con un patrón definido en una política de cifrado. Puede optar por mantener esta coincidencia o anularla con una configuración de cifrado única. Para obtener más información, consulte the section called “Cifrado”.
9. En Configuración de acceso a la red, configure el acceso a la red para la colección.

   - En Tipo de acceso, seleccione acceso público o privado.

     Si elige el privado, especifique qué puntos de conexión de VPC y Servicios de AWS pueden acceder a la colección.

     - Puntos de conexión de VPC para el acceso: especifique uno o más puntos de conexión de VPC para permitir el acceso. Para crear un punto de conexión de VPC, consulte the section called “Punto final de VPC del plano de datos”.
     - Servicio de AWS acceso privado: seleccione uno o más servicios compatibles a los que permitir el acceso.
   - En Tipo de recurso, seleccione si los usuarios pueden acceder a la colección a través de su OpenSearchpunto final (para realizar llamadas a la API a través de cURL, Postman, etc.), a través del punto final de OpenSearch Dashboards (para trabajar con visualizaciones y realizar llamadas a la API a través de la consola) o ambos.

     > **Note**
     >
     > Servicio de AWS El acceso privado solo se aplica al OpenSearch punto final, no al punto final de Dashboards. OpenSearch

   OpenSearch Serverless le notifica si el nombre de la colección que ha introducido coincide con un patrón definido en una política de red. Puede optar por mantener esta coincidencia o anularla con una configuración de red personalizada. Para más información, consulte the section called “Acceso a la red”.
10. (Opcional) Agregue una o varias etiquetas a la colección. Para más información, consulte the section called “Etiquetado de colecciones”.
11. Elija Siguiente.

Configurar campos de búsqueda adicionales

Las opciones que aparecen en la página dos del flujo de trabajo de creación de colecciones dependen del tipo de colección que se esté creando. En esta sección, se describe cómo configurar campos de búsqueda adicionales para cada tipo de colección. En esta sección, también se describe cómo configurar el enriquecimiento semántico automático. Omita cualquier sección que no se aplique a su tipo de colección.

Temas

- Configurar el enriquecimiento semántico automático
- Configurar los campos de búsqueda de series temporales
- Configurar los campos de búsqueda léxica
- Configurar los campos de búsqueda vectorial

Configurar el enriquecimiento semántico automático

Al crear o editar una colección, puede configurar el enriquecimiento semántico automático, lo que simplifica la implementación y las capacidades de la búsqueda semántica en Amazon Service. OpenSearch La búsqueda semántica arroja resultados de consulta que incorporan no solo la coincidencia de palabras clave, sino también la intención y el significado contextual de la búsqueda del usuario. Para obtener más información, consulte Enriquecimiento semántico automático para Serverless.

Para configurar el enriquecimiento semántico automático

1. En la sección Detalles del índice, en Nombre del índice, especifique un nombre.
2. En la sección Campos de enriquecimiento semántico automático, elija Agregar campo de búsqueda semántica.
3. En el campo Nombre del campo de entrada para el enriquecimiento semántico, escriba el nombre del campo que desea enriquecer.
4. El Tipo de datos es Texto. No puede cambiar este valor.
5. En Idioma, seleccione Inglés o Multilingüe.
6. Seleccione Agregar campo.
7. Cuando termine de configurar los campos opcionales para la colección, seleccione Siguiente. Revise los cambios y seleccione Enviar para crear la colección.

Configurar los campos de búsqueda de series temporales

Las opciones de la sección Campos de búsqueda de series temporales se refieren a los datos de series temporales y a los flujos de datos. Para obtener más información acerca de estos temas, consulte Gestión de datos de series temporales en Amazon OpenSearch Service con flujos de datos.

Para configurar los campos de búsqueda de series temporales

1. En la sección Campos de búsqueda de series temporales, seleccione Agregar campo de series temporales.
2. En Nombre del campo, escriba un nombre.
3. En Tipo de datos, seleccione un tipo de la lista.
4. Seleccione Agregar campo.
5. Cuando termine de configurar los campos opcionales para la colección, seleccione Siguiente. Revise los cambios y seleccione Enviar para crear la colección.

Configurar los campos de búsqueda léxica

La búsqueda léxica busca una coincidencia exacta entre una consulta de búsqueda y términos o palabras clave indexados.

Para configurar los campos de búsqueda léxica

1. En la sección Campos de búsqueda léxica, seleccione Agregar campo de búsqueda.
2. En Nombre del campo, escriba un nombre.
3. En Tipo de datos, seleccione un tipo de la lista.
4. Seleccione Agregar campo.
5. Cuando termine de configurar los campos opcionales para la colección, seleccione Siguiente. Revise los cambios y seleccione Enviar para crear la colección.

Configurar los campos de búsqueda vectorial

Para configurar los campos de búsqueda vectorial

1. En la sección Campos vectoriales, seleccione Agregar campo vectorial.
2. En Nombre del campo, escriba un nombre.
3. En Motor, seleccione un tipo de la lista.

4. Escriba el número de dimensiones.
5. En Métrica de distancia, seleccione un tipo de la lista.
6. Cuando termine de configurar los campos opcionales para la colección, seleccione Siguiente.
7. Revise los cambios y seleccione Enviar para crear la colección.

Para crear una colección (CLI)

Utilice los procedimientos de esta sección para crear una colección OpenSearch sin servidor mediante. AWS CLI

Temas

- Antes de empezar
- Creación de una colección
- Crear una colección con un índice de enriquecimiento semántico automático

Antes de empezar

Antes de crear una colección mediante el AWS CLI, utilice el siguiente procedimiento para crear las políticas necesarias para la colección.

> **Note**
>
> En cada uno de los siguientes procedimientos, al especificar un nombre para una colección, el nombre debe cumplir los siguientes criterios:
>
> - Es exclusivo de su cuenta y Región de AWS
> - Contener solo letras minúsculas de la a a la z, números del 0 al 9 y un guion (-)
> - Contener entre 3 y 32 caracteres

Para crear las políticas necesarias de una colección

1. Abra AWS CLI y ejecute el siguiente comando para crear una política de cifrado con un patrón de recursos que coincida con el nombre deseado de la colección.

```
&aws opensearchserverless create-security-policy \
```

```
  --name policy name \
  --type encryption --policy "{\"Rules\":[{\"ResourceType\":\"collection\",
\"Resource\":[\"collection\/collection name\"]}],\"AWSOwnedKey\":true}"
```

Por ejemplo, si planea nombrar la colección como registros de la aplicación, puede crear una política de cifrado como esta:

```
&aws opensearchserverless create-security-policy \
  --name logs-policy \
  --type encryption --policy "{\"Rules\":[{\"ResourceType\":\"collection\",
\"Resource\":[\"collection\/logs-application\"]}],\"AWSOwnedKey\":true}"
```

Si piensa utilizar la política para colecciones adicionales, puede hacer que la regla sea más amplia, como `collection/logs*` o `collection/*`.

2. Ejecute el siguiente comando para configurar los ajustes de la red para la colección mediante una política de red. Es posible crear políticas de red después de crear una colección, pero le recomendamos que lo haga de antemano.

```
&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}]"
```

Para seguir con el ejemplo anterior, logs-application, puede crear la siguiente política de red:

```
&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}]"
```

Creación de una colección

En el siguiente procedimiento, se utiliza la acción de la CreateCollectionAPI para crear una colección del tipo `SEARCH` o`TIMESERIES`. Si no especifica ningún tipo de colección en la solicitud, se usa `TIMESERIES` de forma predeterminada. Para obtener más información acerca de estos tipos,

consulte Elección de un tipo de colección. Para crear una colección de búsquedas vectoriales, consulte the section called “Trabajo con colecciones de búsqueda vectorial”.

Si su colección está cifrada con un Clave propiedad de AWS, `kmsKeyArn` es `auto` en lugar de un ARN.

> **Important**
>
> Después de crear una colección, no podrá acceder a ella a menos que coincida con una política de acceso a los datos. Para obtener más información, consulte Control de acceso a datos para Amazon OpenSearch Serverless.

Creación de una colección

1. Compruebe que ha creado las políticas necesarias que se describen en Antes de empezar.
2. Ejecute el comando siguiente. En `type`, especifique `SEARCH` o `TIMESERIES`.

```
&aws opensearchserverless create-collection --name "collection name" --
type collection type --description "description"
```

Crear una colección con un índice de enriquecimiento semántico automático

Utilice el siguiente procedimiento para crear una nueva colección OpenSearch sin servidor con un índice que esté configurado para el enriquecimiento semántico automático. El procedimiento utiliza la acción de la API OpenSearch sin servidor CreateIndex.

Para crear una nueva colección con un índice configurado con el enriquecimiento semántico automático

Ejecute el siguiente comando para crear la colección y un índice.

```
&aws opensearchserverless create-index \
--region Region ID \
--id collection name --index-name index name \
--index-schema \
'mapping in json'
```

A continuación se muestra un ejemplo.

```
&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"
```

```
                }
            }
        }
    }
}'
```

Acceso a OpenSearch los paneles

Después de crear una colección con Consola de administración de AWS, puede navegar hasta la URL de los OpenSearch paneles de control de la colección. Para encontrar la URL de Dashboards, seleccione Colecciones en el panel de navegación izquierdo y elija la colección para abrir su página de detalles. La URL toma el formato `https://dashboards.us-east-1.aoss.amazonaws.com/_login/?collectionId=07tjusf2h91cunochc`. Cuando navegue hasta la URL, iniciará sesión automáticamente en Dashboards.

Si ya tiene disponible la URL de los OpenSearch paneles, pero no está en ella Consola de administración de AWS, si llama a la URL de los paneles desde el navegador, se redirigirá a la consola. Una vez que introduzcas tus AWS credenciales, iniciarás sesión automáticamente en Dashboards. Para obtener información sobre cómo acceder a las colecciones para SAML, consulte Acceder a los OpenSearch paneles de control con SAML.

El tiempo de espera de la consola de OpenSearch Dashboards es de una hora y no se puede configurar.

> **Note**
>
> El 10 de mayo de 2023, OpenSearch introdujimos un punto final global común para los OpenSearch paneles. Ahora puede navegar a los OpenSearch paneles de control en el navegador con una URL que adopte el formato. `https://dashboards.us-east-1.aoss.amazonaws.com/_login/?collectionId=07tjusf2h91cunochc` Para garantizar la compatibilidad con versiones anteriores, seguiremos admitiendo los puntos finales de OpenSearch Dashboards específicos de la colección existente con este formato. `https://07tjusf2h91cunochc.us-east-1.aoss.amazonaws.com/_dashboards`

Visualización de colecciones

Puedes ver las colecciones existentes Cuenta de AWS en la pestaña Colecciones de la consola de Amazon OpenSearch Service.

Para enumerar las colecciones junto con sus colecciones IDs, envía una ListCollectionssolicitud.

```
&aws opensearchserverless list-collections
```

Respuesta de ejemplo

```
{
   "collectionSummaries":[
      {
         "arn":"arn:aws:aoss:us-east-1:123456789012:collection/07tjusf2h91cunochc",
         "id":"07tjusf2h91cunochc",
         "name":"my-collection",
         "status":"CREATING"
      }
   ]
}
```

Para limitar los resultados de la búsqueda, utilice filtros de colección. Esta solicitud filtra la respuesta a las colecciones en estado ACTIVE:

```
&aws opensearchserverless list-collections --collection-filters '{ "status":
 "ACTIVE" }'
```

Para obtener información más detallada sobre una o más colecciones, incluidos el OpenSearch punto final y el punto final de OpenSearch Dashboards, envía una BatchGetCollectionsolicitud:

```
&aws opensearchserverless batch-get-collection --ids "07tjusf2h91cunochc"
 "1iu5usc4rame"
```

> **Note**
>
> Puede incluir `--names` o `--ids` en la solicitud, pero no ambos.

Respuesta de ejemplo

```
{
   "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":[]
}
```

Eliminación de colecciones

Al eliminar una colección, se eliminan todos los datos e índices en la colección. No se pueden recuperar las colecciones después de eliminarlas.

Para eliminar una colección mediante la consola

1. En el panel Colecciones de la consola de Amazon OpenSearch Service, selecciona la colección que deseas eliminar.

2. Seleccione Eliminar y confirme la eliminación.

Para eliminar una colección mediante el AWS CLI, envía una DeleteCollectionsolicitud:

```
&aws opensearchserverless delete-collection --id 07tjusf2h91cunochc
```

Respuesta de ejemplo

```
{
   "deleteCollectionDetail":{
      "id":"07tjusf2h91cunochc",
      "name":"my-collection",
      "status":"DELETING"
   }
}
```

Trabajo con colecciones de búsqueda vectorial

El tipo de colección de búsqueda vectorial en OpenSearch Serverless proporciona una capacidad de búsqueda por similitud que es escalable y de alto rendimiento. Le facilita la creación de experiencias de búsqueda aumentada de machine learning (ML) modernas y aplicaciones de inteligencia artificial (IA) generativa sin tener que gestionar la infraestructura de base de datos vectorial subyacente.

Los casos de uso de las colecciones de búsquedas vectoriales incluyen búsquedas de imágenes, búsquedas de documentos, recuperación de música, recomendaciones de productos, búsquedas de vídeos, búsquedas basadas en la ubicación, detección de fraudes y detección de anomalías.

Como el motor vectorial de OpenSearch Serverless funciona con la función de búsqueda del vecino más cercano (k-NN) de la empresa OpenSearch, se obtiene la misma funcionalidad con la sencillez de un entorno sin servidor. El motor es compatible con la API del complemento k-NN. Con estas operaciones, puede sacar provecho de la búsqueda de texto completo, el filtrado avanzado, las agregaciones, las consultas geoespaciales, las consultas anidadas para una recuperación más rápida de los datos y la mejora de los resultados de búsqueda.

El motor vectorial proporciona métricas de distancia, como la distancia euclidiana, la similitud de coseno y la similitud de productos de puntos, y puede acomodar 16 000 dimensiones. Puede almacenar campos con varios tipos de datos para los metadatos, como números, valores booleanos, fechas, palabras clave y geopuntos. También puede almacenar campos con texto para obtener

información descriptiva y añadir más contexto a los vectores almacenados. La ubicación de los tipos de datos reduce la complejidad, aumenta la capacidad de mantenimiento y evita la duplicación de datos, los problemas de compatibilidad de versiones y los problemas de licencia.

> **Note**
>
> Tenga en cuenta la siguiente información:
>
> - Amazon OpenSearch Serverless admite la cuantificación escalar de 16 bits de Faiss, que se puede utilizar para realizar conversiones entre vectores flotantes de 32 bits y vectores de 16 bits. Para obtener más información, consulte Cuantificación escalar de 16 bits de Faiss. También puede utilizar vectores binarios para reducir los costos de memoria. Para obtener más información, consulte Vectores binarios.
> - Amazon OpenSearch Serverless admite la búsqueda vectorial basada en disco. La búsqueda vectorial basada en disco reduce considerablemente los costos operativos de las cargas de trabajo de vector en entornos con poca memoria. Para obtener más información, consulte Búsqueda vectorial basada en disco.

Introducción a las colecciones de búsqueda vectorial

En este tutorial, debe completar los siguientes pasos para almacenar, buscar y recuperar incrustaciones vectoriales en tiempo real:

1. Configurar permisos
2. Crear una colección
3. Cargar y buscar datos
4. Eliminar la colección

Paso 1: configurar permisos

Para completar este tutorial (y para usar OpenSearch Serverless en general), debe tener los permisos AWS Identity and Access Management (de IAM) correctos. En este tutorial, crea una colección, carga y busca datos y, a continuación, elimina la colección.

Su usuario o rol debe tener adjunta una política basada en la identidad con los siguientes permisos mínimos:

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": "*"
    }
  ]
}
```

Para obtener más información sobre los permisos de IAM OpenSearch sin servidor, consulte. the section called “Gestión de identidad y acceso”

Paso 2: crear una colección

Una colección es un grupo de OpenSearch índices que funcionan juntos para respaldar una carga de trabajo o un caso de uso específicos.

Para crear una colección OpenSearch sin servidor

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Seleccione Colecciones en el panel de navegación de la izquierda y elija Crear colección.
3. Asigne un nombre a la carcasa de la colección.
4. Para el tipo de colección, seleccione Búsqueda vectorial. Para obtener más información, consulte the section called “Elección de un tipo de colección”.

5. En Tipo de implementación, elimine Habilitar redundancia (réplicas activas). Esto crea una colección en modo de desarrollo o prueba y reduce el número de unidades de OpenSearch cómputo (OCUs) de la colección a dos. Si desea crear un entorno de producción en este tutorial, deje la casilla seleccionada.
6. En Seguridad, seleccione Crear fácilmente para optimizar la configuración de seguridad. De forma predeterminada, todos los datos del motor vectorial se cifran en tránsito y en reposo. El motor vectorial admite permisos de IAM detallados para que pueda definir quién puede crear, actualizar y eliminar cifrados, redes, colecciones e índices.
7. Seleccione Siguiente.
8. Revise la configuración de la colección y seleccione Enviar. Espere unos minutos hasta que el estado de la colección cambie a `Active`.

Paso 3: cargar y buscar datos

Un índice es una colección de documentos con un esquema de datos común que permite almacenar, buscar y recuperar las incrustaciones vectoriales y otros campos. Puedes crear y cargar datos en los índices de una colección OpenSearch sin servidor mediante la consola Dev Tools de OpenSearch Dashboards o una herramienta HTTP como Postman o awscurl. En este tutorial se utilizan Herramientas para desarrolladores.

Para indexar y buscar datos en la colección de viviendas

1. Para crear un índice único para su nueva colección, envíe la siguiente solicitud en la consola de Herramientas para desarrolladores. De forma predeterminada, se crea un índice con un motor `nmslib` y una distancia euclidiana.

```
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. Para indexar un solo documento en housing-index, envíe la siguiente solicitud:

```
POST housing-index/_doc
{
  "housing-vector": [
    10,
    20,
    30
  ],
  "title": "2 bedroom in downtown Seattle",
  "price": "2800",
  "location": "47.71, 122.00"
}
```

3. Para buscar propiedades similares a las del índice, envíe la siguiente consulta:

```
GET housing-index/_search
{
    "size": 5,
    "query": {
        "knn": {
            "housing-vector": {
                "vector": [
                    10,
                    20,
                    30
                ],
                "k": 5
            }
        }
    }
}
```

Paso 4: eliminar la colección

Dado que la colección de housing es para hacer pruebas, debe eliminarla cuando termine de experimentar.

Para eliminar una colección OpenSearch sin servidor

1. Vuelve a la consola OpenSearch de Amazon Service.
2. Seleccione Colecciones en el panel de navegación izquierdo y elija la colección de propiedades.
3. Para confirmar la eliminación, seleccione Eliminar.

Búsqueda filtrada

Puede usar filtros para mejorar los resultados de búsqueda semántica. Para crear un índice y realizar una búsqueda filtrada en sus documentos, sustituya los datos de carga y búsqueda en el tutorial anterior por las siguientes instrucciones. Los demás pasos siguen siendo los mismos. Para obtener más información acerca de los filtros, consulte búsqueda k-NN con filtros.

Para indexar y buscar datos en la colección de viviendas

1. Para crear un índice único para su colección, envíe la siguiente solicitud en la consola de Herramientas para desarrolladores:

```
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. Para indexar un solo documento housing-index-filtered, envíe la siguiente solicitud:

```
POST housing-index-filtered/_doc
{
  "housing-vector": [
    10,
    20,
    30
  ],
  "title": "2 bedroom in downtown Seattle",
  "price": "2800",
  "location": "47.71, 122.00"
}
```

3. Para buscar los datos de un apartamento en Seattle por un precio determinado y dentro de una distancia determinada de un punto geográfico, envíe la siguiente solicitud:

```
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
                  }
                }
              }
            ]
          }
        }
      }
    }
  }
}
```

Cargas de trabajo a escala de mil millones

Las colecciones de búsquedas vectoriales admiten cargas de trabajo con miles de millones de vectores. No necesita reindexar para escalar porque el escalado automático lo hace por usted. Si tienes millones de vectores (o más) con un gran número de dimensiones y necesitas más de 200 OCUs, ponte en contacto con AWS Support para aumentar el número máximo de unidades de OpenSearch cómputo (OCUs) de tu cuenta.

Limitaciones

Las colecciones de búsqueda vectorial presentan las siguientes limitaciones:

- Las colecciones de búsqueda vectorial no son compatibles con el motor ANN Apache Lucene.
- Las colecciones de búsqueda vectorial solo admiten el algoritmo HNSW con Faiss y no admiten FIV ni IVFQ.
- Las colecciones de búsqueda vectorial no admiten las operaciones de la API de calentamiento, estadística y entrenamiento de modelos.
- Las colecciones de búsqueda vectorial no admiten scripts en línea ni almacenados.
- La información sobre el recuento de índices no está disponible en Consola de administración de AWS las colecciones de búsquedas vectoriales.
- El intervalo de actualización de los índices de las colecciones de búsqueda vectorial es de 60 segundos.

Siguientes pasos

Ahora que sabe cómo crear una colección de búsquedas vectoriales e indexar datos, puede que desee probar algunos de los siguientes ejercicios:

- Utilice el cliente OpenSearch Python para trabajar con colecciones de búsqueda vectorial. Consulte este tutorial en GitHub.
- Utilice el cliente OpenSearch Java para trabajar con colecciones de búsquedas vectoriales. Consulte este tutorial en GitHub.
- Configurado LangChain para usarse OpenSearch como una tienda de vectores. LangChain es un marco de código abierto para desarrollar aplicaciones basadas en modelos de lenguaje. Para obtener más información, consulte la Documentación de LangChain .

Uso de políticas del ciclo de vida de los datos con Amazon OpenSearch Serverless

Una política del ciclo de vida de los datos en Amazon OpenSearch Serverless define durante cuánto tiempo OpenSearch Serverless conserva los datos de una recopilación de series temporales. Por ejemplo, puede establecer una política para conservar los datos de registro durante 30 días antes de que OpenSearch Serverless los elimine.

Puede configurar una política independiente para cada índice dentro de cada colección de series temporales de su. Cuenta de AWS OpenSearch Serverless conserva los documentos durante al menos el tiempo que especifique en la política. A continuación, borra automáticamente los

documentos según lo mejor que pueda, normalmente en un plazo de 48 horas o en un 10 % del periodo de retención, lo que sea más largo.

Solo las recopilaciones de series temporales admiten las políticas del ciclo de vida de los datos. Las colecciones de búsqueda y búsqueda vectorial no las admiten.

Temas

- Políticas de ciclo de vida de los datos
- Permisos necesarios
- Prioridad política
- Sintaxis de la política
- Creación de políticas del ciclo de vida de los datos
- Actualización de políticas de ciclo de vida de los datos
- Eliminación de políticas de ciclo de vida de los datos

Políticas de ciclo de vida de los datos

En una política de ciclo de vida de los datos, especifique una serie de reglas. La política de ciclo de vida de los datos le permite administrar el período de retención de los datos asociados a los índices o colecciones que cumplen estas reglas. Estas reglas definen el período de retención de los datos de un índice o grupo de índices. Cada regla consta de un tipo de recurso (`index`), un período de retención y una lista de recursos (índices) a los que se aplica el período de retención.

El período de retención se define con uno de los siguientes formatos:

- `"MinIndexRetention": "24h"`— OpenSearch Serverless conserva los datos de índice durante el período especificado en horas o días. Puede configurar este período para que sea de `24h` a `3650d`.
- `"NoMinIndexRetention": true`— OpenSearch Serverless conserva los datos de índice de forma indefinida.

En el siguiente ejemplo de política, la primera regla especifica un período de retención de 15 días para todos los índices de la colección `marketing`. La segunda regla especifica que todos los nombres de índice que comiencen por `log` en la colección `finance` no tienen un período de retención establecido y se conservarán indefinidamente.

```
{
   "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
   }
}
```

En el siguiente ejemplo de regla de política, OpenSearch Serverless conserva indefinidamente los datos de todos los índices de todas las colecciones de la cuenta.

```
{
   "Rules": [
      {
         "ResourceType": "index",
         "Resource": [
            "index/*/*"
         ]
      }
   ],
   "NoMinIndexRetention": true
}
```

Permisos necesarios

Las políticas de ciclo de vida de OpenSearch Serverless utilizan los siguientes permisos AWS Identity and Access Management (IAM). Puede especificar condiciones de IAM para restringir a los usuarios a las políticas de ciclo de vida de los datos asociadas con colecciones e índices específicos.

- `aoss:CreateLifecyclePolicy`: cree una política de ciclo de vida de los datos.
- `aoss:ListLifecyclePolicies`: enumere todas las políticas de ciclo de vida de los datos de la cuenta actual.
- `aoss:BatchGetLifecyclePolicy`: consulte una política del ciclo de vida de los datos asociada a un nombre de cuenta o política.
- `aoss:BatchGetEffectiveLifecyclePolicy`: consulte una política del ciclo de vida de los datos para un recurso determinado (`index` es el único recurso compatible).
- `aoss:UpdateLifecyclePolicy`: modifique una política de ciclo de vida de los datos determinada y cambie su configuración o recurso de retención.
- `aoss:DeleteLifecyclePolicy`: elimine una política de ciclo de vida de los datos.

La siguiente política de acceso basada en identidades permite al usuario ver todas las políticas de ciclo de vida de los datos y actualizarlas según el patrón de recursos `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": "*"
        }
    ]
}
```

Prioridad política

Puede haber situaciones en las que las reglas de la política de ciclo de vida de los datos se superpongan, dentro de las políticas o entre ellas. Cuando esto sucede, una regla con un nombre o patrón de recurso más específico para un índice anula una regla con un nombre o patrón de recurso más general para cualquier índice que sea común a ambas reglas.

Por ejemplo, en la siguiente política, se aplican dos reglas a un índice `index/sales/logstash`. En esta situación, la segunda regla tiene prioridad porque `index/sales/log*` es la que coincide más tiempo con `index/sales/logstash`. Por lo tanto, OpenSearch Serverless no establece ningún período de retención para el índice.

```
{
      "Rules":[
         {
            "ResourceType":"index",
            "Resource":[
               "index/sales/*",
            ],
            "MinIndexRetention": "15d"
         },
         {
            "ResourceType":"index",
            "Resource":[
               "index/sales/log*",
            ],
            "NoMinIndexRetention": true
         }
      ]
   }
```

Sintaxis de la política

Proporcione una o más reglas. Estas reglas definen la configuración del ciclo de vida de los datos para los índices OpenSearch Serverless.

Cada regla contiene los siguientes elementos. Puede proporcionar `MinIndexRetention` o `NoMinIndexRetention` en cada regla, pero no ambas.

Element	Description (Descripción)
Tipo de recurso	El tipo de recurso al que se aplica la regla. La única opción admitida para las políticas del ciclo de vida de los datos es `index`.
Resource	Una lista de patrones de nombres and/or de recursos. Los patrones son un prefijo seguido de un comodín (*), que permiten que los permisos asociados se apliquen a varios recursos. Por ejemplo, `index/`*`<collection-name|pattern>`*`/`*`<index-name|pattern>`* .
MinIndexRetention	El período mínimo, en días (d) u horas (h), para retener el documento en el índice. El límite inferior es `24h` y el límite superior es `3650d`.
NoMinIndexRetention	Si `true`, OpenSearch Serverless conserva los documentos indefinidamente.

En el siguiente ejemplo, la primera regla se aplica a todos los índices incluidos en el patrón `autoparts-inventory` (`index/autoparts-inventory/*`) y exige que los datos se conserven durante al menos 20 días antes de que se pueda realizar cualquier acción, como eliminarlos o archivarlos.

La segunda regla se dirige a los índices que coinciden con el patrón `auto*/gear` (`index/auto*/gear`) y establece un periodo de retención mínimo de 24 horas.

La tercera regla se aplica específicamente al índice `tires` y no tiene un periodo mínimo de retención, lo que significa que los datos de este índice se pueden eliminar o archivar de forma inmediata o en función de otros criterios. Estas reglas ayudan a administrar la retención de los datos del índice con tiempos de retención variables o sin restricciones de retención.

```
{
  "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
    }
  ]
}
```

Creación de políticas del ciclo de vida de los datos

Para crear una política del ciclo de vida de los datos, debe definir reglas que administren la retención y la eliminación de los datos en función de criterios específicos.

Consola

Para crear una política del ciclo de vida de los datos

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.

2. En el panel de navegación de la izquierda, seleccione Políticas de ciclo de vida de los datos.
3. Seleccione Creación de una política de ciclo de vida de los datos.
4. Escriba un nombre descriptivo para la política.
5. Para Ciclo de vida de los datos, elija Agregar y seleccione las colecciones e índices de la política.

   Comience por elegir las colecciones a las que pertenecen los índices. A continuación, seleccione el índice de la lista o ingrese un patrón de índice. Para seleccionar todas las colecciones como orígenes, escriba un asterisco (*).
6. Para la retención de datos, puede optar por conservar los datos indefinidamente o deseleccionar Ilimitado (nunca eliminar) y especificar un período de tiempo después OpenSearch del cual Serverless eliminará automáticamente los datos de Amazon S3.
7. Seleccione Guardar y, a continuación, Crear.

AWS CLI

Para crear una política de ciclo de vida de los datos mediante el AWS CLI, utilice el create-lifecycle-policycomando con las siguientes opciones:

- `--name`: el nombre de la política.
- `--type`: el tipo de política. El único valor disponible actualmente es `retention`.
- `--policy`: la política del ciclo de vida de los datos. Este parámetro acepta tanto políticas insertadas como archivos .json. Debe codificar las políticas insertadas como una cadena de escape de JSON. Para proporcionar la política en un archivo, utilice el 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}]}"
```

Actualización de políticas de ciclo de vida de los datos

Para actualizar una política del ciclo de vida de los datos, puede modificar las reglas existentes para reflejar los cambios en sus requisitos de retención o eliminación de datos. Esto le permite adaptar sus políticas a medida que evolucionen sus necesidades de administración de datos.

Es posible que transcurran algunos minutos entre el momento en que se actualiza la política y el momento en que OpenSearch Serverless comienza a aplicar los nuevos períodos de retención.

Consola

Para actualizar una política del ciclo de vida de los datos

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, seleccione Políticas de ciclo de vida de los datos.
3. Seleccione la política del ciclo de vida de los datos que desee actualizar y, a continuación, elija Editar.
4. Modifique la política mediante el editor visual o el editor JSON.
5. Seleccione Save.

AWS CLI

Para actualizar una política del ciclo de vida de los datos mediante el AWS CLI, usa el update-lifecycle-policycomando.

Debe incluir el parámetro `--policy-version` en la solicitud. Puede recuperar el control de versiones de la política mediante los comandos list-lifecycle-policies o batch-get-lifecycle-policy. Recomendamos incluir la versión más reciente de la política para evitar que se sobrescriban accidentalmente los cambios de otros usuarios.

La siguiente solicitud actualiza una política del ciclo de vida de los datos con un nuevo documento JSON de política.

Example

```
aws opensearchserverless update-lifecycle-policy \
  --name my-policy \
  --type retention \
  --policy-version MTY2MzY5MTY1MDA3Ml8x \
```

```
  --policy file://my-new-policy.json
```

Eliminación de políticas de ciclo de vida de los datos

Al eliminar una política del ciclo de vida de los datos, OpenSearch Serverless ya no la aplica en ningún índice coincidente.

Consola

Para eliminar una política del ciclo de vida de los datos

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, seleccione Políticas de ciclo de vida de los datos.
3. Seleccione la política de configuración que desea eliminar y, a continuación, elija Eliminar y confirme la eliminación.

AWS CLI

Para eliminar una política del ciclo de vida de los datos mediante el AWS CLI, usa el delete-lifecycle-policycomando.

Example

```
aws opensearchserverless delete-lifecycle-policy \
  --name my-policy \
  --type retention
```

Uso del AWS SDKs para interactuar con Amazon OpenSearch Serverless

En esta sección se incluyen ejemplos de cómo utilizarlos AWS SDKs para interactuar con Amazon OpenSearch Serverless. Estos ejemplos de código muestran cómo crear políticas de seguridad y colecciones, y cómo consultar colecciones.

> **Note**
>
> Actualmente estamos creando estos ejemplos de código. Si quieres contribuir con un ejemplo de código (Java, Go, etc.), abre una solicitud de extracción directamente en el GitHub repositorio.

Temas

- Python
- JavaScript

Python

El siguiente script de ejemplo utiliza AWS SDK para Python (Boto3), así como también el cliente para Python opensearch-py para crear políticas de acceso a datos, de cifrado y de redes, crear colecciones de coincidencias e indexar algunos datos de ejemplo.

Ejecute los siguientes comandos para asegurarse de que dispone de todas las dependencias necesarias:

```
pip install opensearch-py
pip install boto3
pip install botocore
pip install requests-aws4auth
```

En el script, reemplace `Principal` por el nombre de recurso de Amazon (ARN) del usuario o rol que firma la solicitud. También puede modificar opcionalmente el archivo `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

El siguiente script de ejemplo utiliza el SDK de Node.js, así como el cliente opensearch-js JavaScript, para crear políticas de cifrado, redes y acceso a los datos, crear una colección coincidente, crear un índice e indexar algunos datos de ejemplo. JavaScript

Ejecute los siguientes comandos para asegurarse de que dispone de todas las dependencias necesarias:

```
npm i aws-sdk
npm i aws4
npm i @opensearch-project/opensearch
```

En el script, reemplace `Principal` por el nombre de recurso de Amazon (ARN) del usuario o rol que firma la solicitud. También puede modificar opcionalmente el archivo `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()
```

Uso CloudFormation para crear colecciones de Amazon OpenSearch Serverless

Puede utilizarlos CloudFormation para crear recursos de Amazon OpenSearch Serverless, como colecciones, políticas de seguridad y puntos de enlace de VPC. Para obtener una CloudFormation referencia completa sobre OpenSearch Serverless, consulte Amazon OpenSearch Serverless en la Guía del CloudFormation usuario.

La siguiente CloudFormation plantilla de ejemplo crea una política sencilla de acceso a los datos, una política de red y una política de seguridad, así como una colección coincidente. Es una buena forma de empezar a utilizar Amazon OpenSearch Serverless rápidamente y de aprovisionar los elementos necesarios para crear y usar una colección.

> **Important**
>
> En este ejemplo, se utiliza el acceso a la red pública, que no se recomienda para las cargas de trabajo de producción. Recomendamos usar el acceso de VPC para proteger tus colecciones. Para obtener más información, consulte AWS::OpenSearchServerless::VpcEndpoint y the section called “Punto final de VPC del plano de datos”.

```
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
```

Hacer copias de seguridad de las colecciones mediante instantáneas

Las instantáneas son point-in-time copias de seguridad de sus colecciones de Amazon OpenSearch Serverless que proporcionan funciones de recuperación ante desastres. OpenSearch Serverless crea y administra automáticamente las instantáneas de sus colecciones, lo que garantiza la continuidad empresarial y la protección de los datos. Cada instantánea contiene metadatos de índice (ajustes y mapeos de sus índices), metadatos de clúster (plantillas de índice y alias) y datos de índice (todos los documentos y datos almacenados en sus índices).

OpenSearch Serverless ofrece copias de seguridad automáticas cada hora, sin configuración manual, sin gastos de mantenimiento ni costes de almacenamiento adicionales, con una recuperación rápida en caso de pérdida accidental de datos y la posibilidad de restaurar índices específicos a partir de una instantánea.

Antes de trabajar con instantáneas, comprenda estas importantes consideraciones. La creación de una instantánea tarda en completarse y no es instantánea. Los documentos nuevos o las actualizaciones que se realicen durante la creación de la instantánea no se incluirán en la instantánea. Solo puede restaurar instantáneas en su colección original y no en una nueva. Cuando se restauran, los índices reciben versiones nuevas UUIDs que difieren de sus versiones originales. La restauración a un índice abierto existente en OpenSearch Serverless sobrescribirá los datos de ese índice siempre que no se proporcione un nuevo nombre de índice o un patrón de prefijo. Esto difiere del comportamiento principal. OpenSearch Solo puede ejecutar una operación de restauración a la vez y no puede iniciar varias operaciones de restauración en la misma colección simultáneamente. Si se intenta restaurar los índices durante una operación de restauración activa, se produce un error en la operación. Durante una operación de restauración, las solicitudes a los índices fallan.

Permisos necesarios

Para trabajar con instantáneas, configure los siguientes permisos en su política de acceso a datos. Para obtener más información acerca de las políticas de acceso a datos, consulte Políticas de acceso a datos frente a políticas de IAM.

Política de acceso a los datos	APIs
Aoss: DescribeSnapshot	AUTOMATIZA /_ cat/snapshots/aoss

Política de acceso a los datos	APIs
	OBTENGA _/snapshot/aoss-automated/snapshot
aoss: RestoreSnapshot	POST /_ /_restore snapshot/aoss-automated/snapshot
aoss: DescribeCollectionItems	GET /_cat/recovery

Puede configurar las políticas mediante los siguientes AWS CLI comandos:

1. create-access-policy
2. delete-access-policy
3. get-access-policy
4. update-access-policy

Este es un ejemplo de comando CLI para crear una política de acceso. En el comando, sustituya el *example* contenido por su información específica.

```
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."
  }
]'
```

Uso de instantáneas

De forma predeterminada, al crear una nueva colección, OpenSearch Serverless crea instantáneas automáticamente cada hora. No es necesario realizar ninguna acción. Cada instantánea incluye todos los índices de la colección. Una vez que OpenSearch Serverless haya creado las instantáneas, puede enumerarlas y revisar los detalles de la instantánea mediante los siguientes procedimientos.

Enumere las instantáneas

Utilice los siguientes procedimientos para enumerar todas las instantáneas de una colección y revisar sus detalles.

Console

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, selecciona Serverless y, a continuación, Colecciones.
3. Elige el nombre de tu colección para abrir su página de detalles.
4. Seleccione la pestaña Instantáneas para ver todas las instantáneas generadas.
5. Revise la información de las instantáneas, que incluye:
   - ID de la instantánea: identificador único de la instantánea
   - Estado: estado actual (disponible, en curso)
   - Hora de creación: momento en que se tomó la instantánea

AWS CLI

- Utilice el siguiente comando para enumerar todas las instantáneas de una colección.

  ```
  GET /_cat/snapshots/aoss-automated
  ```

  OpenSearch Serverless devuelve una respuesta como la siguiente:

```
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
```

Obtenga detalles de la instantánea

Utilice los siguientes procedimientos para recuperar información detallada sobre una instantánea específica.

Console

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, selecciona Serverless y, a continuación, Colecciones.
3. Elige el nombre de tu colección para abrir su página de detalles.
4. Elija la pestaña Snapshots (Instantáneas).
5. Elija el identificador del trabajo de la instantánea para mostrar información detallada sobre la instantánea, incluidos los metadatos, los índices incluidos y la información de temporización.

AWS CLI

- Utilice el siguiente comando para recuperar información sobre una instantánea. En el comando, sustituya el *example* contenido por su información específica.

  ```
  GET _snapshot/aoss-automated/snapshot/
  ```

  Solicitud de ejemplo:

  ```
  GET _snapshot/aoss-automated/snapshot-ExampleSnapshotID1/
  ```

  Respuesta de ejemplo:

```
{
    "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 respuesta a la instantánea incluye varios campos clave: `id` proporciona un identificador único para la operación de instantánea, `status` devuelve el estado actual `SUCCESS` o `IN_PROGRESS` `duration` indica el tiempo necesario para completar la operación de instantánea y `indexes` devuelve el número de índices incluidos en la instantánea.

Restauración a partir de una instantánea

Al restaurar a partir de una instantánea, se recuperan los datos de una copia de seguridad realizada anteriormente. Este proceso es crucial para la recuperación ante desastres y la administración de datos en OpenSearch Serverless. Antes de realizar la restauración, tenga en cuenta que los

índices restaurados tendrán versiones diferentes UUIDs a las originales; la restauración a un índice abierto existente en OpenSearch Serverless sobrescribirá los datos de ese índice siempre que no se proporcione un nuevo nombre de índice o un patrón de prefijo; las instantáneas solo se pueden restaurar a su colección original (no se admite la restauración cruzada de colecciones); además, las operaciones de restauración afectarán al rendimiento del clúster, así que planifique en consecuencia.

Utilice los siguientes procedimientos para restaurar los índices respaldados a partir de una instantánea.

Console

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, selecciona Serverless y, a continuación, Colecciones.
3. Elige el nombre de tu colección para abrir su página de detalles.
4. Selecciona la pestaña Instantáneas para ver las instantáneas disponibles.
5. Elija la instantánea desde la que desee realizar la restauración y, a continuación, seleccione Restaurar a partir de una instantánea.
6. En el cuadro de diálogo Restaurar a partir de una instantánea:
   - En el nombre de la instantánea, compruebe el ID de la instantánea seleccionada.
   - Para el ámbito de la instantánea, elija una de las siguientes opciones:
     - Todos los índices de la colección: restaure todos los índices de la instantánea
     - Índices específicos: seleccione los índices individuales que desee restaurar
   - En Destination, selecciona la colección en la que deseas realizar la restauración.
   - (Opcional) Configure los ajustes de cambio de nombre para cambiar el nombre de los índices restaurados:
     - No cambie el nombre: conserve los nombres de los índices originales
     - Añadir un prefijo a los nombres de índice restaurados: añada un prefijo para evitar conflictos
     - Cambie el nombre mediante una expresión regular: utilice patrones de cambio de nombre avanzados
   - (Opcional) Configure los ajustes de notificación para que se le notifique cuando la restauración se complete o detecte errores.
7. Seleccione Guardar para iniciar la operación de restauración.

OpenSearch API

1. Ejecute el siguiente comando para identificar la instantánea adecuada.

   ```
   GET /_snapshot/aoss-automated/_all
   ```

   Para obtener una lista de instantáneas más pequeña, ejecute el siguiente comando.

   ```
   GET /_cat/snapshots/aoss-automated
   ```

2. Ejecute el siguiente comando para verificar los detalles de la instantánea antes de restaurar. En el comando, sustituya el *example* contenido por su información específica.

   ```
   GET _snapshot/aoss-automated/snapshot-ExampleSnapshotID1/
   ```

3. Ejecute el siguiente comando para restaurar a partir de una instantánea específica.

   ```
   POST /_snapshot/aoss-automated/snapshot-ID/_restore
   ```

   Puede personalizar la operación de restauración incluyendo el cuerpo de la solicitud. A continuación se muestra un ejemplo.

   ```
   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. Ejecute el siguiente comando para ver el progreso de la restauración.

   ```
   GET /_cat/recovery
   ```

> **Note**
>
> Al restaurar una instantánea con un comando que incluye el cuerpo de la solicitud, puede usar varios parámetros para controlar el comportamiento de la restauración. El `indices` parámetro especifica qué índices se van a restaurar y admite patrones comodín. Configúrelo `ignore_unavailable` para continuar con la operación de restauración incluso si falta un índice en la instantánea. Se utiliza `include_global_state` para determinar si se debe restaurar el estado del clúster y `include_aliases` para controlar si se restauran los alias asociados. Los `rename_replacement` parámetros `rename_pattern` y cambian el nombre de los índices durante la operación de restauración.

Soporte de códecs Zstandard en Amazon Serverless OpenSearch

Los códecs de índice determinan cómo se comprimen y almacenan los campos almacenados de un índice en el disco y en S3. El códec de índice se controla mediante la `index.codec` configuración estática que especifica el algoritmo de compresión. Esta configuración afecta tanto al tamaño del fragmento del índice como al rendimiento de la operación del índice.

De forma predeterminada, los índices de OpenSearch Serverless utilizan el códec predeterminado con el algoritmo de compresión. LZ4 OpenSearch Serverless también admite `zstd` `zstd_no_dict` códecs con niveles de compresión configurables del 1 al 6.

> **Important**
>
> Como `index.codec` es una configuración estática, no se puede cambiar después de la creación del índice.

Para obtener más información, consulte la documentación de los códecs OpenSearch indexados.

Creación de un índice con el códec ZSTD

Puede especificar el códec ZSTD durante la creación del índice mediante la configuración: `index.codec`

```
PUT /your_index
{
  "settings": {
```

```
    "index.codec": "zstd"
  }
}
```

Niveles de compresión

Los códecs ZSTD admiten niveles de compresión opcionales mediante la `index.codec.compression_level` configuración y aceptan números enteros en el rango [1, 6]. Los niveles de compresión más altos dan como resultado mejores relaciones de compresión (almacenamiento más pequeño) pero velocidades de compresión y descompresión más lentas. El nivel de compresión predeterminado es 3.

```
PUT /your_index
{
  "settings": {
    "index.codec": "zstd",
    "index.codec.compression_level": 2
  }
}
```

Análisis comparativo de rendimiento

Según las pruebas comparativas realizadas con el conjunto de datos nyc_taxi, la compresión ZSTD logró una compresión entre un 26 y un 32% mejor en comparación con la línea base en diferentes combinaciones de niveles de compresión `zstd` y, `zstd_no_dict`

Métrica	ZSTD L1	ZSTD L6	ZSTD_NO_DICT L1	ZSTD_NO_DICT L6
Reducción del tamaño del índice	28,10%	32%	26,90%	28,70%
Cambio en el rendimiento de indexación	-0,50%	-23,80%	-0,50%	-5,30%
Mejora de la latencia de Match-all Query p90	-16,40%	29,50%	-16,40%	23,40%
Mejora de la latencia de Range Query p90	90,90%	92,40%	-282,90%	92,50%

Métrica	ZSTD L1	ZSTD L6	ZSTD_NO_DICT L1	ZSTD_NO_DICT L6
Mejora de la latencia Agg de la cantidad de distancia p90	2%	24,70%	2%	13,80%

Para obtener más información, consulte el AWS OpenSearch blog.

Ahorre almacenamiento mediante el uso de una fuente derivada

De forma predeterminada, OpenSearch Serverless almacena cada documento ingerido en el `_source` campo, que contiene el cuerpo del documento JSON original, e indexa los campos individuales para su búsqueda. Si bien el `_source` campo no se puede buscar, se conserva para poder devolver el documento completo al ejecutar solicitudes de recuperación, como get y search. Cuando la fuente derivada está habilitada, OpenSearch Serverless omite el almacenamiento del `_source` campo y, en su lugar, lo reconstruye dinámicamente bajo demanda, por ejemplo, durante las operaciones de búsqueda, obtención, mget, reindexación o actualización. El uso de la configuración de fuente derivada puede reducir el uso del almacenamiento hasta en un 50%.

Configuración

Para configurar la fuente derivada para su índice, cree el índice con la siguiente `index.derived_source.enabled` configuración:

```
PUT my-index1
{
  "settings": {
    "index": {
      "derived_source": {
        "enabled": true
      }
    }
  }
}
```

Consideraciones importantes

- Solo se admiten ciertos tipos de campos. Para obtener una lista de los campos y limitaciones admitidos, consulta la OpenSearch documentación. Si crea un índice con una fuente derivada y un

campo no compatible, no se podrá crear el índice. Si intenta ingerir un documento con un campo no compatible en un índice con código fuente derivado activado, la ingestión fallará. Utilice esta función solo cuando conozca los tipos de campos que se agregarán al índice.

- La configuración `index.derived_source.enabled` es estática. Esto no se puede cambiar una vez creado el índice.

Limitaciones en las respuestas a las consultas

Cuando la fuente derivada está habilitada, se imponen ciertas limitaciones a la forma en que se generan y devuelven las respuestas a las consultas.

- Los campos de fecha con varios formatos especificados utilizan siempre el primer formato de la lista para todos los documentos solicitados, independientemente del formato original introducido.
- Los valores de los puntos geográficos se devuelven en un `{"lat": lat_val, "lon": lon_val}` formato fijo y pueden perder algo de precisión.
- Las matrices con varios valores se pueden ordenar y los campos de palabras clave se pueden deduplicar.

Para obtener más información, consulte el blog. OpenSearch

Evaluación comparativa del rendimiento

Según las pruebas comparativas realizadas con el conjunto de datos nyc_taxi, la fuente derivada logró una reducción del 58% en el tamaño del índice en comparación con el valor inicial.

Métrica	Fuente derivada
Reducción del tamaño del índice	58,3%
Cambio en el rendimiento de indexación	3,7%
Indexando el cambio de latencia del p90	6,9%
Mejora de la latencia de Match-all Query p90	19%
Mejora de la latencia de Range Query p90	-18,8%
Mejora de la latencia Agg de la cantidad de distancia p90	-7,3%

Para obtener más información, consulte el OpenSearch blog.

Grupos de recopilación de Amazon OpenSearch Serverless

Los grupos de colecciones de Amazon OpenSearch Serverless organizan varias colecciones y permiten compartir recursos de cómputo entre colecciones con diferentes claves de KMS. Este modelo de cómputo compartido reduce los costos al eliminar la necesidad de unidades de OpenSearch cómputo (OCUs) independientes para cada clave de KMS.

Cada colección OpenSearch sin servidor que cree está protegida mediante el cifrado de los datos en reposo, que se utiliza AWS KMS para almacenar y administrar las claves de cifrado. Las colecciones del mismo grupo de colecciones comparten los recursos informáticos y el espacio de memoria de la OCU, incluso cuando utilizan claves KMS diferentes para el cifrado.

Los grupos de colecciones proporcionan aislamiento para cumplir con los requisitos de seguridad y rendimiento. Puede agrupar las colecciones con la misma clave de KMS en un único grupo de recopilación para aislarlas por motivos de seguridad, o bien combinar las colecciones con diferentes claves de KMS en el mismo grupo para optimizar los costes. Esta flexibilidad le permite equilibrar los requisitos de seguridad con la eficiencia de los recursos.

Al añadir una colección a un grupo de colecciones, OpenSearch Serverless la asigna a los recursos informáticos compartidos del grupo. El sistema administra automáticamente la distribución de las cargas de trabajo entre estos recursos y, al mismo tiempo, mantiene la seguridad al cifrar los datos de cada colección con la clave KMS designada. Los controles de acceso siguen aplicándose a nivel de recopilación, y los recursos informáticos compartidos acceden a varias claves de KMS según sea necesario para atender las colecciones del grupo.

La asignación de recursos se controla mediante el establecimiento de límites mínimos y máximos de OCU a nivel de grupo de recopilación. Estos límites se aplican a todas las colecciones del grupo y le ayudan a gestionar los costes y, al mismo tiempo, a garantizar un rendimiento uniforme.

De forma predeterminada, hay una cuota de servicio (límite) para el número de colecciones de un grupo de colecciones, el número de índices de una colección y el número de colecciones de OCUs un grupo de colecciones. Para obtener más información, consulte Cuotas OpenSearch sin servidor.

Conceptos clave

Grupo de recopilación

AWS Recurso que actúa como contenedor de una o más colecciones. Cada grupo de colecciones tiene un identificador único y puede tener sus propios límites de capacidad y ajustes de configuración.

Computación compartida

Las colecciones del mismo grupo de colecciones comparten el mismo conjunto de OCUs, independientemente de las claves de KMS que utilicen. Este uso compartido reduce los costos al eliminar la necesidad de recursos de cómputo independientes para cada clave de KMS.

Límites de capacidad

Puede establecer límites de OCU mínimos y máximos para las operaciones de indexación y búsqueda a nivel de grupo de recopilación. Estos límites ayudan a controlar los costes y a garantizar un rendimiento uniforme.

Límites de capacidad de los grupos de recolección

Los grupos de recolección proporcionan un control detallado sobre la asignación de recursos mediante límites mínimos y máximos de OCU. Estos límites se aplican a todas las colecciones del grupo y funcionan independientemente de la configuración de capacidad a nivel de cuenta.

De forma predeterminada, hay una cuota de servicio (límite) para el número de colecciones de un grupo de colecciones, el número de índices de una colección y el número de colecciones de un grupo OCUs de colecciones. Para obtener más información, consulte Cuotas OpenSearch sin servidor.

Descripción de los límites de capacidad de los grupos de recopilación

Puede configurar los límites mínimos y máximos de OCU para las operaciones de indexación y búsqueda a nivel de grupo de colecciones. Estos límites controlan la forma en que OpenSearch Serverless escala los recursos de las colecciones del grupo:

- OCU mínima: la cantidad mínima OCUs que OpenSearch Serverless mantiene para el grupo de recopilación, lo que garantiza un rendimiento de referencia uniforme.

  - Si la carga de trabajo requiere menos OCU que el valor mínimo especificado, OpenSearch Serverless seguirá manteniendo el valor mínimo especificado de OCU y la facturación reflejará lo mismo.
  - Si la carga de trabajo requiere un número de OCU superior al valor mínimo especificado, OpenSearch Serverless mantendrá el nivel de OCU necesario para la carga de trabajo y la facturación reflejará la mayor utilización de las OCU.
- OCU máxima: la cantidad máxima de unidades OpenSearch sin servidor OCUs que pueden ampliarse para el grupo de recolección, lo que le ayuda a controlar los costos.

Los límites de capacidad de los grupos de cobranza están disociados de los límites a nivel de cuenta. La configuración máxima de OCU a nivel de cuenta se aplica solo a las colecciones que no están asociadas a ningún grupo de colecciones, mientras que la configuración máxima de OCU por grupo de colecciones se aplica a las colecciones de ese grupo específico.

Valores límite de capacidad válidos

Al establecer los límites mínimos y máximos de OCU para un grupo de recopilación, solo puede usar valores del siguiente conjunto: 1, 2, 4, 8, 16 y múltiplos de 16 (por ejemplo, 32, 48, 64, 80, 96) hasta un máximo de 1696. OCUs

Los límites mínimos y máximos de OCU son opcionales al crear un grupo de recopilación. Si no especificas un límite máximo de OCU, OpenSearch Serverless utiliza un valor predeterminado de 96. OCUs

El límite mínimo de OCU debe ser inferior o igual al límite máximo de OCU.

Comprender la relación entre los límites de OCU a nivel de cuenta y de grupo de cobranza

Al planificar su capacidad OpenSearch sin servidor, es importante entender cómo interactúan los límites de las OCU a nivel de cuenta y los límites de las OCU de los grupos de cobranza. La suma de la configuración máxima de OCU en todos los grupos de recopilación más la configuración máxima de OCU a nivel de cuenta debe ser inferior o igual al límite de la cuota de servicio por cuenta. Para ver los valores límite actuales, consulte Cuotas OpenSearch sin servidor.

> **Note**
>
> La configuración máxima de OCU a nivel de cuenta solo se aplica a las colecciones que no están asociadas a ningún grupo de colecciones. Los cobros dentro de los grupos de cobranza se rigen por sus respectivos límites de grupo de cobranza, no por los límites a nivel de cuenta.

Esta restricción se aplica tanto a la indexación como a la búsqueda de forma independiente. OCUs Por ejemplo, si configura los ajustes a nivel de cuenta y los grupos de recopilación, debe asegurarse de que el total no supere el límite de la cuota de servicio para la indexación y, por separado, no OCUs supere el límite de la cuota de servicio para la búsqueda. OCUs Además, puede crear un máximo de 300 grupos de recopilación por cuenta.

Ejemplo: planificar la capacidad con límites a nivel de cuenta y de grupo de cobranza

Si estableces la OCU máxima de búsqueda a nivel de cuenta en 500 y el límite de cuota de servicio es de 1700:

- Y cree 2 grupos de recopilación, la suma de la OCU máxima para los 2 grupos de recopilación no debe ser superior a 1200 (1700 - 500)
- Puede dejar cada grupo de colecciones en la OCU máxima predeterminada de 96 (96 + 96 + 500 = 692), dejando espacio para un futuro crecimiento.
- O puede aumentar el máximo de cada grupo de recolección a 600 (600 + 600 + 500 = 1700), utilizando toda la capacidad permitida por la cuota de servicio

Esta relación es fundamental para la planificación de la capacidad. Antes de crear nuevos grupos de recopilación o aumentar los límites máximos de la OCU, compruebe que la asignación total no supere el límite de la cuota de servicio. Si alcanza este límite, debe reducir la configuración máxima de OCU en los grupos de cobranza existentes o disminuir la configuración máxima de OCU a nivel de cuenta para dejar espacio para nuevas asignaciones.

Configuración de los límites de capacidad

Puede establecer límites de capacidad al crear un grupo de recopilación o actualizarlos más adelante. Para configurar los límites de AWS CLI capacidad mediante los UpdateCollectionGroupcomandos CreateCollectionGroupo:

```
aws opensearchserverless create-collection-group \
    --name my-collection-group \
    --capacity-limits
 maxIndexingCapacityInOCU=32,maxSearchCapacityInOCU=32,minIndexingCapacityInOCU=4,minSearchCapa
```

Para actualizar los límites de capacidad de un grupo de recopilación existente:

```
aws opensearchserverless update-collection-group \
    --id abcdef123456 \
    --capacity-limits
 maxIndexingCapacityInOCU=48,maxSearchCapacityInOCU=48,minIndexingCapacityInOCU=8,minSearchCapa
```

Supervisión de la capacidad de un grupo de recopilación

OpenSearch Serverless emite las siguientes métricas de Amazon CloudWatch Logs a intervalos de un minuto para ayudarle a supervisar los límites de uso y capacidad de la OCU a nivel de grupo de recopilación:

- `IndexingOCU`— El número de indexación que utiliza OCUs actualmente el grupo de recopilación.
- `SearchOCU`— El número de búsquedas que utiliza OCUs actualmente el grupo de colecciones.

OpenSearch Serverless también emite métricas de OCU a nivel de cuenta para las colecciones que no están asociadas a ningún grupo de colecciones. Puede agregar estas métricas CloudWatch para visualizar la suma de OCUs todos los grupos de cobros y cobros a nivel de cuenta.

Configure alarmas para que le notifiquen cuando su grupo de recolección se acerque a sus límites de capacidad para que pueda ajustar la configuración según sea necesario. Para obtener más información sobre las métricas de OpenSearch Serverless, consultethe section called “Supervisión sin servidor OpenSearch ”.

Cómo se aplican los límites de capacidad

OpenSearch Serverless impone los límites de capacidad de los grupos de recopilación durante las operaciones de escalado. Cuando sus colecciones necesitan recursos adicionales, OpenSearch Serverless se amplía hasta el límite máximo de OCU. Cuando la demanda disminuye, OpenSearch Serverless se reduce pero mantiene al menos el límite mínimo de OCU para garantizar un rendimiento uniforme.

Los límites de capacidad se aplican solo cuando el grupo de recopilación contiene al menos una colección. Los grupos de recolección vacíos no consumen OCUs ni imponen límites de capacidad.

Si una operación de escalado supera el límite máximo de OCU o infringe el requisito mínimo de OCU, OpenSearch Serverless rechaza la operación para mantener el cumplimiento de los límites configurados.

Claves de cifrado y KMS en grupos de recopilación

Cada colección OpenSearch sin servidor que cree está protegida mediante el cifrado de los datos en reposo, que se utiliza AWS KMS para almacenar y administrar sus claves de cifrado. Al trabajar con grupos de colecciones, tiene flexibilidad a la hora de especificar la clave KMS para sus colecciones.

Puede proporcionar la clave KMS asociada a una colección de dos maneras:

- En la CreateCollection solicitud: especifique la clave de KMS directamente al crear la colección mediante el `encryption-config` parámetro.
- En las políticas de seguridad: defina la asociación de claves KMS en una política de seguridad de cifrado.

Al especificar una clave de KMS en ambas ubicaciones, la clave de KMS proporcionada en la CreateCollection solicitud tiene prioridad sobre la configuración de la política de seguridad.

Esta flexibilidad simplifica la administración de las colecciones a escala, especialmente cuando se necesitan crear varias colecciones con claves KMS únicas. En lugar de crear y administrar miles de políticas de cifrado, puede especificar la clave KMS directamente durante la creación de la colección.

Compartir OCUs entre diferentes claves de KMS

Los grupos de colecciones permiten compartir recursos de cómputo entre colecciones con diferentes claves de KMS. Las colecciones del mismo grupo de colecciones comparten el espacio de memoria de la OCU, independientemente de sus claves de cifrado. Este modelo de procesamiento compartido reduce los costos al eliminar la necesidad de separar OCUs cada clave KMS.

Los grupos de recopilación proporcionan aislamiento para cumplir con los requisitos de seguridad y rendimiento. Puede agrupar las colecciones con la misma clave de KMS en un único grupo de recopilación para aislarlas por motivos de seguridad, o bien combinar las colecciones con diferentes claves de KMS en el mismo grupo para optimizar los costes. Esta flexibilidad le permite equilibrar los requisitos de seguridad con la eficiencia de los recursos.

El sistema mantiene la seguridad al cifrar los datos de cada colección con la clave KMS designada. Los controles de acceso siguen aplicándose a nivel de recopilación y los recursos informáticos compartidos acceden a varias claves de KMS según sea necesario para atender las recopilaciones del grupo.

Permisos de KMS necesarios

Cuando especificas una clave KMS en la CreateCollection solicitud, necesitas los siguientes permisos adicionales:

- `kms:DescribeKey`— Permite a OpenSearch Serverless recuperar información sobre la clave KMS.
- `kms:CreateGrant`— Permite a OpenSearch Serverless crear una concesión para la clave KMS a fin de habilitar las operaciones de cifrado.

Estos permisos no son necesarios cuando se utilizan claves AWS propias.

Crea grupos de recopilación

En este tema se describe cómo crear, configurar y administrar grupos de colecciones en Amazon OpenSearch Serverless. Utilice los grupos de colecciones para organizar las colecciones y compartir los recursos de cómputo a fin de optimizar los costos. Establezca límites mínimos y máximos de OCU a nivel de grupo de recopilación para controlar el rendimiento y los gastos.

Cree un grupo de recolección

Utilice los siguientes procedimientos para crear un nuevo grupo de recopilación y configurar sus ajustes. Cree un grupo de colecciones mediante la consola OpenSearch Serverless o la AWS SDKs . AWS CLI Al crear un grupo de recopilación, se especifican los límites de capacidad y otras opciones de configuración.

Console

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, selecciona Serverless y, a continuación, selecciona Grupos de colección
3. Selecciona Crear grupo de colecciones.

4. En Nombre del grupo de colecciones, introduce un nombre para tu grupo de colecciones. El nombre debe tener entre 3 y 32 caracteres, empezar con una letra minúscula y contener únicamente letras minúsculas, números y guiones.
5. (Opcional) En Descripción, introduce una descripción para tu grupo de colección.
6. En la sección Administración de la capacidad, configure los límites de la OCU:
   - Capacidad máxima de indexación: la cantidad máxima de indexación a la OCUs que pueden ampliarse las colecciones de este grupo.
   - Capacidad máxima de búsqueda: el número máximo de búsquedas al OCUs que pueden ampliarse las colecciones de este grupo.
   - Capacidad mínima de indexación: la cantidad mínima de indexación que se debe mantener OCUs para lograr un rendimiento uniforme.
   - Capacidad mínima de búsqueda: la cantidad mínima de búsquedas que se deben mantener OCUs para lograr un rendimiento uniforme.
7. (Opcional) En la sección Etiquetas, añade etiquetas para ayudar a organizar e identificar tu grupo de recopilación.
8. Selecciona Crear grupo de colecciones.

AWS CLI

- Utilice el create-collection-groupcomando para crear un nuevo grupo de colecciones. En el comando, sustituya el *example* contenido por su propia información específica.

```
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
```

El comando devuelve detalles sobre el grupo de colecciones creado, incluidos su identificador único y su ARN.

Agrega una colección nueva a un grupo de colecciones

Al crear una nueva colección, especifique el nombre de un grupo de colecciones existente al que asociar la colección. Utilice los siguientes procedimientos para añadir una nueva colección a un grupo de colecciones.

Console

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, selecciona Serverless y, a continuación, Colecciones
3. Seleccione Crear colección.
4. En Nombre de colección, ingresa un nombre para tu colección. El nombre debe tener entre 3 y 28 caracteres, empezar con una letra minúscula y contener únicamente letras minúsculas, números y guiones.
5. (Opcional) En Descripción, introduce una descripción para tu colección.
6. En la sección Grupo de colecciones, seleccione el grupo de colecciones al que desea que se asigne la colección. Una colección solo puede pertenecer a un grupo de colecciones a la vez.

   (Opcional) También puedes elegir crear un grupo nuevo. Esto te llevará al flujo de trabajo Crear un grupo de colecciones. Cuando termine de crear el grupo de colecciones, vuelva al paso 1 de este procedimiento para empezar a crear la nueva colección.
7. Continúe con el flujo de trabajo para crear la colección.

   > **Important**
   >
   > No te salgas del flujo de trabajo de creación de colecciones. Si lo hace, se detendrá la configuración de la colección. La página de detalles de la colección aparece una vez finalizada la configuración.

AWS CLI

- Utilice el comando create-collection para crear una nueva colección y añadirla a un grupo de colecciones existente. En el comando, sustituya el *example* contenido por su propia información específica.

```
aws opensearchserverless create-collection \
```

```
    --name my-collection \
    --type SEARCH \
    --collection-group-name my-collection-group \
    --description "Collection for search workloads"
```

Administrar grupos de recopilación de Amazon OpenSearch Serverless

Tras crear los grupos de colecciones de Amazon OpenSearch Serverless, puede modificar su configuración a medida que cambien sus necesidades. Utilice estas operaciones de administración para actualizar los límites de capacidad y ver los detalles de los grupos de recopilación. Estos cambios le ayudan a optimizar la asignación de recursos y a mantener una organización eficiente de sus colecciones.

Vea los grupos de colecciones

Muestre sus grupos de colecciones OpenSearch sin servidor para revisar sus configuraciones, las colecciones asociadas y el estado actual.

Console

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, selecciona Serverless y, a continuación, Colecciones
3. Seleccione la pestaña Grupos de colecciones. Se muestran los grupos de colecciones de tu cuenta.
4. Elija el nombre de un grupo de recopilación para mostrar sus detalles.

AWS CLI

- Utilice el list-collection-groupscomando para enumerar todos los grupos de colecciones de su cuenta. Utilice el batch-get-collection-groupcomando para ver los detalles sobre grupos de recopilación específicos. En los siguientes comandos, sustituya el *example* contenido por su propia información específica.

  Para enumerar todos los grupos de colecciones:

```
aws opensearchserverless list-collection-groups
```

Para obtener detalles sobre grupos de colecciones específicos:

```
aws opensearchserverless batch-get-collection-group \
    --names my-collection-group another-group
```

Actualiza la configuración de los grupos de colecciones

Actualice la configuración del grupo de recopilación OpenSearch Serverless para modificar configuraciones como los límites de capacidad y la descripción.

Console

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, selecciona Serverless y, a continuación, Colecciones
3. Seleccione la pestaña Grupos de colecciones. Se muestran los grupos de colecciones de tu cuenta.
4. Elija el nombre de un grupo de recopilación para mostrar sus detalles.
5. En Detalles del grupo de colecciones, seleccione Editar.
6. Realice los cambios que desee y, a continuación, seleccione Guardar.

AWS CLI

- Utilice el update-collection-groupcomando para actualizar la descripción y los límites de capacidad de un grupo de recopilación existente. En el siguiente comando, sustituya el *example* contenido por su propia información.

```
aws opensearchserverless update-collection-group \
    --id abcdef123456 \
    --description "Updated description for production workloads" \
    --capacity-limits
 maxIndexingCapacityInOCU=30,maxSearchCapacityInOCU=30,minIndexingCapacityInOCU=4,minSea
```

Los cambios en los límites de capacidad se aplican inmediatamente y pueden afectar al comportamiento de escalado de las colecciones del grupo.

Eliminar grupos de colecciones

Para poder eliminar un grupo de colecciones, primero debe eliminar todas las colecciones del grupo. No puede eliminar un grupo de colecciones que contenga colecciones.

Console

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, selecciona Serverless y, a continuación, Colecciones
3. Seleccione la pestaña Grupos de colecciones. Se muestran los grupos de colecciones de tu cuenta.
4. Elige el nombre del grupo de recopilación que deseas eliminar.

   > **Important**
   >
   > Elimine todas las colecciones del grupo de colecciones actualizando cada colección para eliminar la asociación de grupos de colecciones o moviéndolas a otros grupos de colecciones.

5. En la parte superior de la página, elija Eliminar.
6. Confirme la eliminación y, a continuación, seleccione Eliminar.

AWS CLI

- Utilice el delete-collection-groupcomando para eliminar un grupo de colecciones.

  > **Important**
  >
  > Elimine todas las colecciones del grupo de colecciones actualizando cada colección para eliminar la asociación del grupo de colecciones o moviéndolas a otros grupos de colecciones.

  En el siguiente comando, sustituya el *example* contenido por su propia información.

Elimine el grupo de colecciones vacío:

```
aws opensearchserverless delete-collection-group \
    --id abcdef123456
```

Administración de los límites de capacidad de Amazon OpenSearch Serverless

Con Amazon OpenSearch Serverless, no tiene que administrar la capacidad usted mismo. OpenSearch Serverless escala automáticamente la capacidad de cómputo de su cuenta en función de la carga de trabajo actual. La capacidad de cómputo sin servidor se mide en unidades de OpenSearch cómputo ()OCUs. Cada OCU es una combinación de 6 GiB de memoria y la CPU virtual (vCPU) correspondiente, así como la transferencia de datos a Amazon S3. Para obtener más información sobre la arquitectura desacoplada de OpenSearch Serverless, consulte. the section called “Funcionamiento”

Al crear la primera colección, OpenSearch Serverless crea instancias en OCUs función de la configuración de redundancia. De forma predeterminada, las réplicas activas redundantes están habilitadas, lo que crea cuatro instancias (dos para la indexación y dos para la OCUs búsqueda). Esto garantiza una alta disponibilidad con los nodos en espera en otra zona de disponibilidad.

Para el desarrollo y las pruebas, puede deshabilitar la configuración Habilitar redundancia para una colección. Esto elimina las réplicas en espera y utiliza solo dos OCUs (una para la indexación y otra para la búsqueda).

OCUs Siempre existen, incluso cuando no hay actividad de indexación o búsqueda. Todas las colecciones posteriores pueden compartirlas OCUs, excepto las colecciones con AWS KMS claves únicas, que crean una instancia de su propio conjunto de claves. OCUs Todas las colecciones asociadas a un grupo de colecciones pueden compartir el mismo conjunto de. OCUs Solo se puede incluir un tipo de colección (búsqueda, serie temporal o búsqueda vectorial) en un único grupo de colecciones. Para obtener más información, consulte the section called “Grupos de colecciones”.

OpenSearch Serverless se amplía y agrega automáticamente a OCUs medida que aumenta el uso de la indexación y las búsquedas. Cuando el tráfico disminuye, la capacidad se reduce a la cantidad mínima OCUs requerida para el tamaño de los datos.

En el caso de las recopilaciones de series temporales y de búsqueda, la cantidad OCUs necesaria cuando están inactivas es proporcional al tamaño de los datos y al recuento de índices. En el caso de las colecciones vectoriales, los requisitos de la OCU dependen de la memoria (RAM) para almacenar los gráficos vectoriales y del espacio en disco para almacenar los índices. Cuando no está inactiva, los requisitos de la OCU tienen en cuenta ambos factores.

Las colecciones vectoriales almacenan los datos de índice en el almacenamiento local de la OCU. Los límites de RAM de la OCU se alcanzan más rápido que los límites del disco, lo que restringe las colecciones de vectores por espacio de RAM.

Con la redundancia habilitada, la capacidad de la OCU se reduce a un mínimo de 1 OCU (0,5 OCU x 2) para la indexación y 1 OCU (0,5 OCU x 2) para la búsqueda. Al deshabilitar la redundancia, la colección se puede reducir a 0,5 OCU para la indexación y a 0,5 OCU para la búsqueda.

La escalabilidad también tiene en cuenta la cantidad de particiones necesarias para su colección o índice. Cada OCU admite un número específico de fragmentos, y el número de índices debe ser proporcional al número de fragmentos. La cantidad total de bases OCUs requerida es el máximo de sus requisitos de datos, memoria y partición. Para obtener más información, consulte las funcionalidades de búsqueda rentables de Amazon OpenSearch Serverless, a cualquier escala, en el blog sobre AWS macrodatos.

En el caso de las colecciones de búsqueda y de búsquedas vectoriales, todos los datos se almacenan en índices calientes para garantizar tiempos de respuesta rápidos a las consultas. Las colecciones de series temporales utilizan una combinación de almacenamiento en caliente y templado, lo que almacena los datos más recientes en un almacenamiento en caliente para optimizar los tiempos de respuesta a las consultas para los datos a los que se accede con mayor frecuencia. Para obtener más información, consulte the section called “Elección de un tipo de colección”.

> **Note**
>
> Una colección de búsqueda vectorial no se puede compartir OCUs con colecciones de búsqueda y series temporales, incluso si la colección de búsqueda vectorial utiliza la misma clave KMS que las colecciones de búsqueda o series temporales. Se creará un nuevo conjunto de OCUs vectores para tu primera colección vectorial. Las colecciones OCUs de vectores se comparten entre las mismas colecciones de claves de KMS.

Para administrar la capacidad de sus colecciones y controlar los costos, puede especificar la capacidad máxima general de indexación y búsqueda para la cuenta corriente y la región, y

OpenSearch Serverless amplía los recursos de su colección automáticamente en función de estas especificaciones.

Como la capacidad de indexación y búsqueda se escalan por separado, debe especificar los límites de cuenta para cada una de ellas:

- Capacidad máxima de indexación: OpenSearch Serverless puede aumentar la capacidad de indexación hasta este número de. OCUs
- Capacidad máxima de búsqueda: OpenSearch Serverless puede aumentar la capacidad de búsqueda hasta este número de. OCUs

> **Note**
>
> En este momento, la configuración de capacidad solo se aplica según el nivel de cuenta. No puede configurar los límites de capacidad por colección.

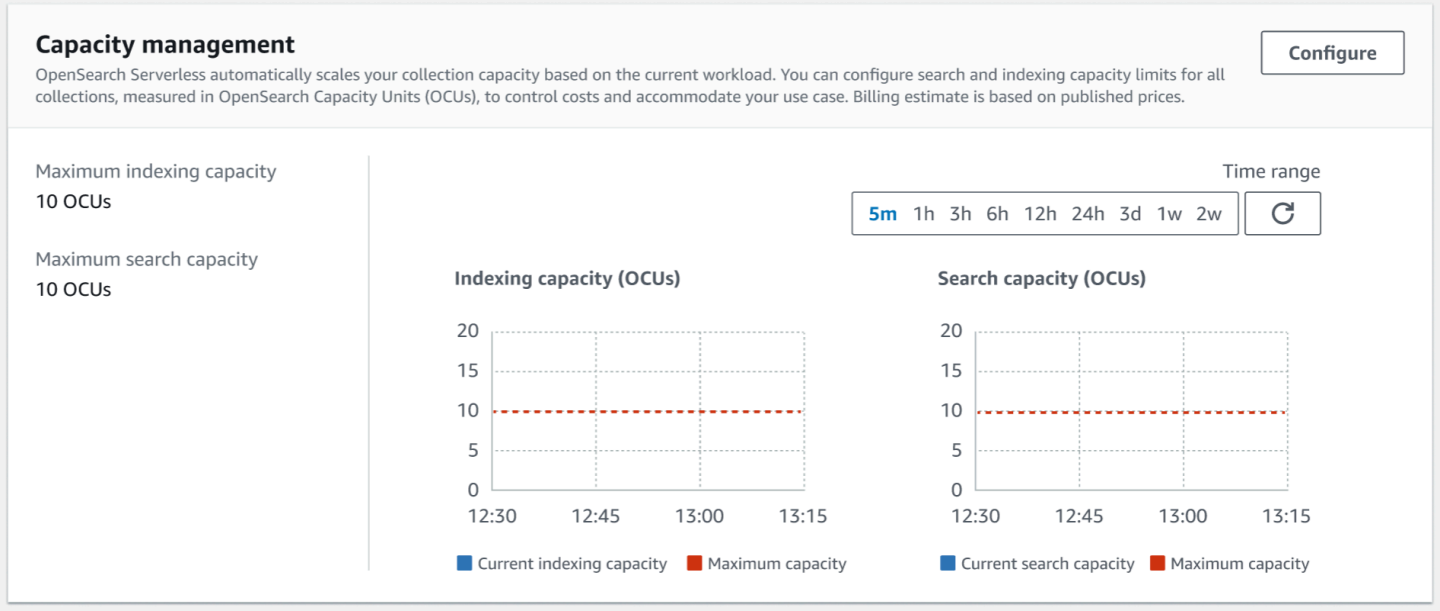
El objetivo debe ser garantizar que la capacidad máxima sea lo bastante alta como para asumir los picos de carga de trabajo. Según su configuración, OpenSearch Serverless reduce automáticamente el número de colecciones OCUs para procesar la carga de trabajo de indexación y búsqueda.

Temas

- Como establecer los parámetros de capacidad
- Límites de la capacidad máxima
- Monitoreo del uso de la capacidad

Como establecer los parámetros de capacidad

Para configurar los ajustes de capacidad en la consola OpenSearch Serverless, expanda Serverless en el panel de navegación izquierdo y seleccione Panel de control. Especifique la capacidad máxima de indexación y búsqueda en Administración de capacidades:



Para configurar la capacidad mediante el AWS CLI, envíe una UpdateAccountSettingssolicitud:

```
aws opensearchserverless update-account-settings \
    --capacity-limits '{ "maxIndexingCapacityInOCU": 8,"maxSearchCapacityInOCU": 9 }'
```

Límites de la capacidad máxima

El total máximo de índices que puede contener una colección es 1000. Para los tres tipos de recopilaciones, la capacidad máxima predeterminada de la OCU es de 10 OCUs para la indexación y 10 OCUs para la búsqueda. La capacidad mínima de OCU permitida para una cuenta es de 1 OCU [0,5 OCU x 2] para indexación y 1 OCU [0,5 OCU x 2] para búsqueda. Para todas las colecciones, la capacidad máxima permitida es de 1700 para la indexación y de 1700 OCUs para la búsqueda. OCUs Puede configurar el recuento de OCU para que sea cualquier número desde 2 hasta la capacidad máxima permitida, en múltiplos de 2.

Cada OCU incluye suficiente almacenamiento efímero en caliente para 120 GiB de datos de índice. OpenSearch Serverless admite hasta 1 TiB de datos por índice en las colecciones de búsqueda y búsqueda vectorial, y 100 TiB de datos importantes por índice en una colección de series temporales. En el caso de las recopilaciones de series temporales, puede incorporar aún más datos, que se pueden almacenar como datos templados en S3.

Para ver una lista de todas las cuotas, consulta OpenSearch Cuotas sin servidor.

Monitoreo del uso de la capacidad

Puedes monitorizar las métricas `Search0CU` y las CloudWatch métricas a `Indexing0CU` nivel de cuenta para entender cómo están escalando tus colecciones. Le recomendamos que configure alarmas para que le notifiquen si su cuenta se está acercando a un límite para las métricas relacionadas con la capacidad, de modo que pueda ajustar su configuración de capacidad en consecuencia.

Además, puede utilizar estas métricas para determinar si su configuración de capacidad máxima es adecuada o si necesita ajustarla. Analice estas métricas para centrar sus esfuerzos para optimizar la eficiencia de sus colecciones. Para obtener más información sobre las métricas a las que envía OpenSearch Serverless, consulte. CloudWatch the section called “Supervisión sin servidor OpenSearch ”

Ingerir datos en colecciones de Amazon OpenSearch Serverless

En estas secciones se proporcionan detalles sobre las canalizaciones de ingesta compatibles para la ingesta de datos en las colecciones de Amazon OpenSearch Serverless. También incluyen algunos de los clientes que puede utilizar para interactuar con las operaciones de la API. OpenSearch Sus clientes deben ser compatibles con la versión OpenSearch 2.x para poder integrarse con OpenSearch Serverless.

Temas

- Permisos mínimos necesarios
- OpenSearch Ingestión
- Fluent Bit
- Amazon Data Firehose
- Go
- Java
- JavaScript
- Logstash
- Python
- Ruby
- Firma de solicitudes HTTP con otros clientes

Permisos mínimos necesarios

Para incorporar datos a una colección OpenSearch sin servidor, el director que escribe los datos debe tener los siguientes permisos mínimos asignados en una política de acceso a los datos:

```
[
   {
      "Rules":[
         {
            "ResourceType":"index",
            "Resource":[
               "index/target-collection/logs"
            ],
            "Permission":[
               "aoss:CreateIndex",
               "aoss:WriteDocument",
               "aoss:UpdateIndex"
            ]
         }
      ],
      "Principal":[
         "arn:aws:iam::123456789012:user/my-user"
      ]
   }
]
```

Los permisos pueden ser más amplios si planea escribir en índices adicionales. Por ejemplo, en lugar de especificar un índice objetivo único, puede conceder permisos a todos los índices (index/*target-collection* /*) o a un subconjunto de índices (index//). *target-collection logs**

Para obtener una referencia de todas las operaciones de OpenSearch API disponibles y sus permisos asociados, consulte. the section called “Operaciones y complementos compatibles”

OpenSearch Ingestión

En lugar de utilizar un cliente de terceros para enviar datos directamente a una colección OpenSearch sin servidor, puede utilizar Amazon OpenSearch Ingestion. Usted configura sus generadores de datos para que envíen datos a OpenSearch Ingestion y esta entrega automáticamente los datos a la colección que usted especifique. También puede configurar OpenSearch Ingestion para transformar los datos antes de entregarlos. Para obtener más información, consulte OpenSearch Ingestión de Amazon.

Una canalización de OpenSearch ingestión necesita permiso para escribir en una colección OpenSearch sin servidor que esté configurada como sumidero. Estos permisos incluyen la capacidad de describir la colección y enviarle solicitudes HTTP. Para obtener instrucciones sobre cómo usar OpenSearch Ingestión para añadir datos a una colección, consulte. the section called “Concesión de acceso a colecciones para canalizaciones”

Para empezar a usar OpenSearch Ingestion, consulte. the section called “Tutorial: incorporar datos en una colección”

Fluent Bit

Puede utilizar la imagen AWS de Fluent Bit y el complemento OpenSearch de salida para introducir datos en colecciones OpenSearch sin servidor.

> **Note**
>
> Debe tener la versión 2.30.0 o posterior de la imagen de Fluent Bit AWS para poder integrarla con Serverless. OpenSearch

Ejemplo de configuración:

Este ejemplo de sección de resultados del archivo de configuración muestra cómo utilizar una colección OpenSearch Serverless como destino. La adición importante es el parámetro `AWS_Service_Name`, que es `aoss`. `Host` es el punto de conexión de la colección.

```
[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 admite OpenSearch Serverless como destino de entrega. Para obtener instrucciones sobre cómo enviar datos a OpenSearch Serverless, consulte Crear una transmisión de entrega de Kinesis Data Firehose y OpenSearch elegir Serverless para su destino en la Guía para desarrolladores de Amazon Data Firehose.

El rol de IAM que proporcione a Firehose para la entrega se debe especificar dentro de una política de acceso a datos con el permiso mínimo `aoss:WriteDocument` para la colección de destino y debe tener un índice preexistente al que enviar datos. Para obtener más información, consulte the section called “Permisos mínimos necesarios”.

Antes de enviar datos a OpenSearch Serverless, es posible que deba realizar transformaciones en los datos. Para obtener más información acerca de cómo utilizar las funciones de Lambda para realizar esta tarea, consulte Amazon Kinesis Data Firehose Data Transformation en la misma guía.

Go

El siguiente código de ejemplo utiliza el cliente opensearch-go para Go a fin de establecer una conexión segura con la colección OpenSearch Serverless especificada y crear un índice único. Debe proporcionar valores para `region` y `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

En el siguiente código de ejemplo, se utiliza el cliente opensearch-java para Java a fin de establecer una conexión segura con la colección OpenSearch Serverless especificada y crear un índice único. Debe proporcionar valores para `region` y `host`.

La diferencia importante en comparación con los dominios de OpenSearch servicio es el nombre del servicio (`aoss`en lugar de). `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();
```

El siguiente código de ejemplo vuelve a establecer una conexión segura y, a continuación, busca en un índice.

```
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

El siguiente código de ejemplo utiliza el cliente opensearch-js JavaScript para establecer una conexión segura con la colección OpenSearch Serverless especificada, crear un índice único, agregar un documento y eliminar el índice. Debe proporcionar valores para `node` y `region`.

La diferencia importante en comparación con los dominios de OpenSearch servicio es el nombre del servicio (`aoss`en lugar de). `es`

Version 3

En este ejemplo, se utiliza la versión 3 del SDK para JavaScript 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

En este ejemplo, se usa la versión 2 del SDK para JavaScript 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

Puede usar el OpenSearch complemento Logstash para publicar registros en colecciones OpenSearch sin servidor.

Para usar Logstash para enviar datos a Serverless OpenSearch

1. Instale la versión 2.0.0 o posterior del logstash-output-opensearchcomplemento mediante Docker o Linux.

   Docker

   Docker aloja el software Logstash OSS con el complemento de OpenSearch salida preinstalado: opensearchproject/ -output-plugin. logstash-oss-with-opensearch Puede extraer la imagen como cualquier otra:

   ```
   docker pull opensearchproject/logstash-oss-with-opensearch-output-plugin:latest
   ```

   Linux

   En primer lugar, instale la última versión de Logstash si aún no lo hizo. A continuación, instale la versión 2.0.0 del complemento de salida:

   ```
   cd logstash-8.5.0/
   bin/logstash-plugin install --version 2.0.0 logstash-output-opensearch
   ```

   Si el complemento ya está instalado, actualícelo a la versión más reciente:

   ```
   bin/logstash-plugin update logstash-output-opensearch
   ```

   A partir de la versión 2.0.0 del complemento, el SDK usa la versión 3. AWS Si utilizas una versión de Logstash anterior a la 8.4.0, debes eliminar todos los AWS complementos preinstalados e instalar el complemento: `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. Para que el complemento de OpenSearch salida funcione con OpenSearch Serverless, debes realizar las siguientes modificaciones en la sección de salida de logstash.conf: `opensearch`

   - Especifique `aoss` como el `service_name` en `auth_type`.
   - Especifique el punto de conexión de la colección para `hosts`.
   - Agregue los parámetros `default_server_major_version` y `legacy_template`. Estos parámetros son necesarios para que el complemento funcione con Serverless. OpenSearch

```
output {
  opensearch {
    hosts => "collection-endpoint:443"
    auth_type => {
      ...
      service_name => 'aoss'
    }
    default_server_major_version => 2
    legacy_template => false
  }
}
```

Este ejemplo de archivo de configuración toma las entradas de los archivos de un bucket de S3 y las envía a una colección OpenSearch 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. A continuación, ejecute Logstash con la nueva configuración para probar el complemento:

```
bin/logstash -f config/test-plugin.conf
```

Python

El siguiente código de ejemplo usa el cliente opensearch-py para Python para establecer una conexión segura con la colección OpenSearch Serverless especificada, crear un índice único y buscar en ese índice. Debe proporcionar valores para `region` y `host`.

La diferencia importante en comparación con los dominios de OpenSearch servicio es el nombre del servicio (`aoss`en lugar de). `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` gema proporciona acceso a OpenSearch Serverless, junto con OpenSearch Service, de forma inmediata. Dispone de todas las funciones del cliente opensearch-ruby, ya que es una dependencia de esta gema.

Cuando cree una instancia del firmante de Sigv4, especifique `aoss` como nombre del servicio:

```
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 de solicitudes HTTP con otros clientes

Los siguientes requisitos se aplican a la hora de firmar solicitudes en colecciones OpenSearch sin servidor cuando se crean solicitudes HTTP con otros clientes.

- Debe especificar el nombre del servicio como `aoss`.
- El encabezado `x-amz-content-sha256` es obligatorio para todas las solicitudes de Signature Version 4 de AWS . Proporciona un hash de la carga de solicitud. Si hay una carga útil de solicitud, defina el valor en su hash criptográfico () del algoritmo de hash seguro (SHA). SHA256 Si no hay ninguna carga de solicitud, establezca el valor en `e3b0c44298fc1c149afbf4c8996fb92427ae41e4649b934ca495991b7852b855`, que es el hash de una cadena vacía.

Temas

- Indexación con cURL
- Indexación con Postman

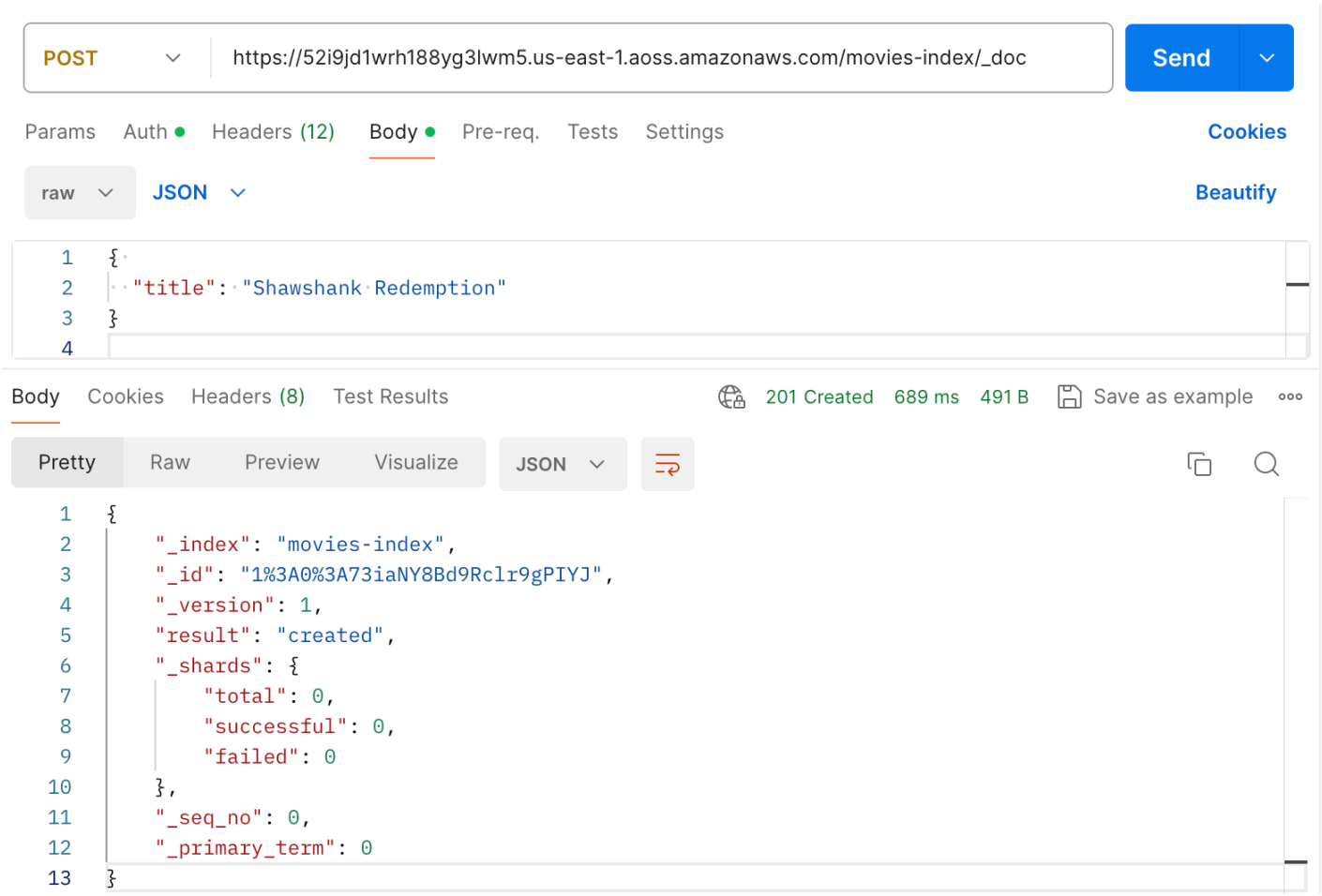
Indexación con cURL

La siguiente solicitud de ejemplo utiliza la biblioteca de solicitudes de URL de cliente (cURL) para enviar un único documento a un índice denominado `movies-index` dentro de una colección:

```
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"}'
```

Indexación con Postman

La siguiente imagen muestra cómo enviar solicitudes a una colección con Postman. Para obtener instrucciones sobre cómo autenticarse, consulte el flujo de trabajo de autenticación con AWS firma en Postman.



Configuración de Machine Learning en Amazon OpenSearch Serverless

Machine Learning

Machine Learning (ML) proporciona capacidades de ML en forma de algoritmos de ML y modelos remotos. Con el acceso a estos modelos, puede ejecutar varios flujos de trabajo de IA, como RAG o la búsqueda semántica. ML permite la experimentación y la implementación en producción de casos de uso de IA generativa mediante el uso de los últimos modelos alojados externamente que se pueden configurar con conectores. Después de configurar un conector, debe configurarlo en un modelo y, a continuación, implementarlo para realizar una predicción.

Connectors

Los conectores facilitan el acceso a los modelos de alojados en plataformas de ML de terceros. Sirven como puerta de enlace entre su OpenSearch clúster y un modelo remoto. Para obtener más información, consulte la siguiente documentación sobre :

- Creación de conectores para plataformas de aprendizaje automático de terceros en el sitio web de OpenSearch documentación
- Conectores para plataformas externas
- Conectores para Servicios de AWS

> **Important**
>
> - Al crear una política de confianza, **`ml.opensearchservice.amazonaws.com`** agréguela como principio OpenSearch de servicio.
> - Omita los pasos de la página Conectores que muestran cómo configurar un dominio en la política.
> - Agregue la sentencia `iam:PassRole` en el paso de Configuración de permisos.
> - Omita el paso Asignar la función de aprendizaje automático en los OpenSearch paneles de control. No es necesaria la configuración del rol de backend. Esto se aplica a los conectores para Servicios de AWS plataformas externas y a los conectores para plataformas externas.
> - En su solicitud de SigV4 al punto de conexión de la colección, defina el nombre del servicio como **`aoss`** en lugar de **`es`**.

Modelos

Un modelo es la funcionalidad principal que se utiliza en varios flujos de trabajo de IA. Por lo general, se asocia el conector a un modelo para realizar una predicción mediante el conector. Una vez que el modelo está en estado desplegando, puede ejecutar la predicción. Para obtener más información, consulte Registrar un modelo alojado en una plataforma de terceros en el sitio web de OpenSearch documentación.

> **Note**
>
> No todas las funciones del modelo son compatibles con OpenSearch Serverless, como los modelos locales. Para obtener más información, consulte Machine Learning APIs y funciones no compatibles.

Configurar permisos para Machine Learning

En la siguiente sección, se describen las políticas de acceso a los datos de recopilación necesarias para Machine Learning (ML). *`placeholder values`*Sustitúyala por su información específica. Para obtener más información, consulte Permisos de política compatibles.

```
{
    "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: otorga permiso para buscar y consultar conectores, modelos y grupos de modelos.
- AOSS:CreateMLResource: otorga permiso para crear conectores, modelos y grupos de modelos.

- AOSS:Update MLResource — Otorga permiso para actualizar conectores, modelos y grupos de modelos.
- AOSS:Delete MLResource — Otorga permiso para eliminar conectores, modelos y grupos de modelos.
- AOSS:Execute MLResource — Otorga permiso para realizar predicciones en los modelos.

Machine Learning APIs y funciones no compatibles

No es compatible APIs

Amazon OpenSearch Serverless no admite los siguientes sistemas de Machine Learning (ML) APIs :

- Funcionalidad de modelo local
- API de entrenamiento de modelos
- API de algoritmos de predicción de modelos
- API de predicción por lotes
- API de agentes y sus herramientas correspondientes
- Servidor MCP APIs
- Memoria APIs
- Controlador APIs
- API de ejecución de algoritmos
- API de perfil de ML
- API de estadísticas de ML

Para obtener más información sobre ML APIs, consulte ML APIs en el sitio web de OpenSearch documentación.

Características no admitidas

Amazon OpenSearch Serverless no admite las siguientes funciones de aprendizaje automático:

- Agentes y herramientas
- Modelos locales

- El procesador de inferencia de machine learning dentro de las canalizaciones de búsqueda y de ingesta
  - Procesador de ingesta de inferencia de ML
  - Procesador de respuesta de búsqueda de inferencia de ML
  - Procesador de petición de búsqueda de inferencia de ML

Para obtener más información sobre estas funciones, consulte la siguiente documentación en el sitio web de OpenSearch documentación:

- Machine learning
- Procesador de inferencias de ML
- Busque canalizaciones

Configurar la búsqueda neuronal y la búsqueda híbrida en entornos OpenSearch sin servidor

Búsqueda neuronal

Amazon OpenSearch Serverless admite la funcionalidad de búsqueda neuronal para realizar operaciones de búsqueda semántica en sus datos. La búsqueda neuronal utiliza modelos de machine learning para comprender el significado semántico y el contexto de sus peticiones, lo que proporciona resultados de búsqueda más relevantes que las búsquedas tradicionales basadas en palabras clave. En esta sección, se explica cómo configurar la búsqueda neuronal en OpenSearch Serverless, incluidos los permisos necesarios, los procesadores compatibles y las principales diferencias con respecto a la implementación estándar. OpenSearch

Con la búsqueda neuronal, usted puede realizar búsquedas semánticas en sus datos, lo que considera el significado semántico para comprender la intención de sus peticiones de búsqueda. Esta capacidad se basa en los siguientes componentes:

- Procesador de canalización de ingesta de incrustaciones de texto
- Consulta neuronal
- Consulta neuronal dispersa

Búsqueda híbrida

Con la búsqueda híbrida, usted puede mejorar la relevancia de la búsqueda combinando las capacidades de búsqueda por palabras clave y búsqueda semántica. Para usar la búsqueda híbrida, cree una canalización de búsqueda que procese sus resultados y combine las puntuaciones de los documentos. Para obtener más información, consulte las canalizaciones de búsqueda en el sitio web de OpenSearch documentación. Use los siguientes componentes para implementar la búsqueda híbrida:

- Procesador de canalización de búsqueda de normalización

  Técnicas de normalización compatibles

  - `min_max`
  - `l2`

  Técnicas de combinación compatibles

  - `arithmetic_mean`
  - `geometric_mean`
  - `harmonic_mean`

  Para obtener más información sobre las técnicas de normalización y combinación, consulte los campos del cuerpo de la solicitud en el sitio web de OpenSearchdocumentación.
- Consulta híbrida

Consultas neuronales e híbridas

De forma predeterminada, OpenSearch calcula las puntuaciones de los documentos mediante el BM25 algoritmo Okapi basado en palabras clave, que funciona bien para las consultas de búsqueda que contienen palabras clave. La búsqueda neuronal proporciona nuevos tipos de consultas para consultas en lenguaje natural y la capacidad de combinar búsqueda semántica y por palabras clave.

Example: neural

```
"neural": {
  "vector_field": {
    "query_text": "query_text",
    "query_image": "image_binary",
    "model_id": "model_id",
    "k": 100
  }
}
```

Para obtener más información, consulte la consulta neuronal en el sitio web de documentación. OpenSearch

Example: hybrid

```
"hybrid": {
      "queries": [
        array of lexical, neural, or combined queries
      ]
    }
```

Para obtener más información, consulte la consulta híbrida en el sitio web de OpenSearch documentación.

Para configurar los componentes de búsqueda semántica en Amazon OpenSearch Serverless, siga los pasos del tutorial de Neural Search del sitio web de OpenSearch documentación. Tenga en cuenta estas diferencias importantes:

- OpenSearch Serverless solo admite modelos remotos. Debe configurar los conectores para los modelos alojados de forma remota. No es necesario implementar o eliminar modelos remotos. Para obtener más información, consulte Introducción a la búsqueda semántica e híbrida en el sitio web de OpenSearch documentación.
- Espere hasta 15 segundos de latencia cuando realice búsquedas en su índice vectorial o busque canalizaciones de búsqueda e ingesta creadas recientemente.

Configuración de permisos

La búsqueda neuronal en OpenSearch Serverless requiere los siguientes permisos. Para obtener más información, consulte Permisos de política compatibles.

Example: Política de búsqueda neuronal

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 índice vectorial donde se almacenan las incrustaciones de texto.
- aoss: * CollectionItems — Crea canalizaciones de ingesta y búsqueda.
- aoss: * MLResource — Crea y registra modelos de incrustación de texto.
- aoss: APIAccess All: proporciona acceso a las operaciones de OpenSearch APIs búsqueda e ingesta.

A continuación, se describen las políticas de acceso a datos de la colección necesarias para la búsqueda neuronal. Sustitúyalo *`placeholder values`* por su información específica.

```
            {
                "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"
        ]
    }
]
```

Configuración de flujos de trabajo en Amazon OpenSearch Serverless

Flujos de trabajo

Los flujos de trabajo ayudan a los desarrolladores a innovar en las aplicaciones de IA. OpenSearch El proceso actual de utilizar las ofertas de aprendizaje automático (ML) OpenSearch, como la búsqueda semántica, requiere tareas complejas de configuración y preprocesamiento, además de consultas detalladas de los usuarios, que pueden llevar mucho tiempo y ser propensas a errores. Los flujos de trabajo son un marco de simplificación para encadenar múltiples llamadas a la API. OpenSearch

Para obtener información sobre la configuración y el uso, consulte Automatizar las configuraciones en el OpenSearchsitio web. Cuando utilice flujos de trabajo en OpenSearch Serverless, tenga en cuenta estas importantes diferencias:

- OpenSearch Serverless usa solo modelos remotos en los pasos del flujo de trabajo. No es necesario implementar estos modelos.
- OpenSearch Serverless no admite el paso de reindexación del flujo de trabajo.
- Cuando busque flujos de trabajo y estados de flujos de trabajo después de otras llamadas a la API, espere hasta 15 segundos de latencia para que aparezcan las actualizaciones.

OpenSearch Las colecciones sin servidor admiten los flujos de trabajo cuando se utilizan como fuente de datos en la aplicación de interfaz de usuario. OpenSearch Para obtener más información, consulte Administración de asociaciones de orígenes de datos.

Configuración de permisos

Antes de crear y aprovisionar una plantilla, verifique que tenga los permisos necesarios. Si necesita ayuda, póngase en contacto con el administrador de su cuenta. OpenSearch Los flujos de trabajo sin servidor requieren los siguientes permisos. Usted puede delimitar los permisos a una colección específica definiendo el ARN del recurso de la colección en su política de IAM.

Example: Política de flujos de trabajo

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"
    }
  ]
}
```

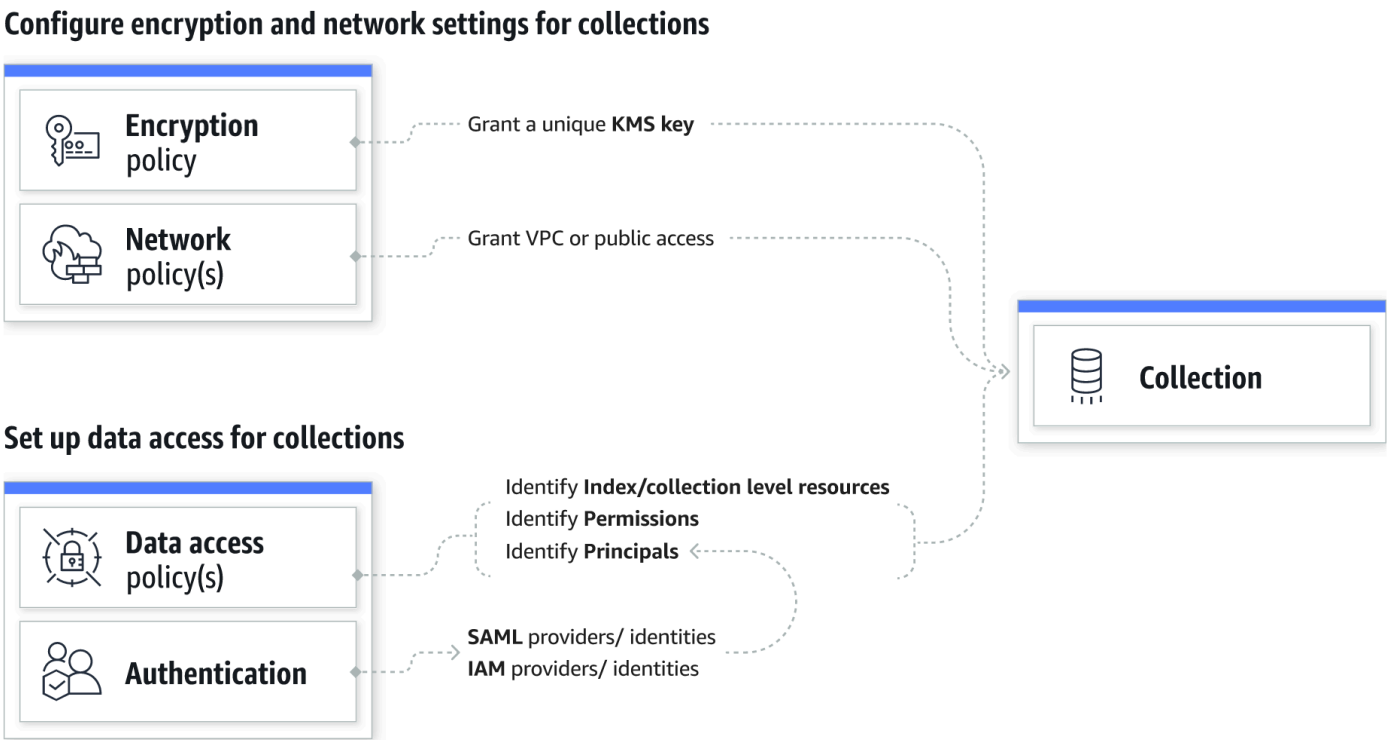
- aoss: * CollectionItems — Otorga permiso para crear y administrar plantillas y aprovisionar canalizaciones de búsqueda e ingesta.
- aoss: *Index: otorga permiso para crear y eliminar índices mediante operaciones de API. OpenSearch
- aoss: * MLResource — Otorga permiso para aprovisionar los pasos del flujo de trabajo que utilizan Configure Machine Learning.

Descripción general de la seguridad en Amazon OpenSearch Serverless

La seguridad de Amazon OpenSearch Serverless se diferencia fundamentalmente de la seguridad de Amazon OpenSearch Service en los siguientes aspectos:

Característica	OpenSearch Servicio	OpenSearch Sin servidor
Control de acceso a los datos	El acceso a los datos está determinado por las políticas de IAM y el control de acceso detallado.	El acceso a los datos está determinado por las políticas de acceso a los datos.
Cifrado en reposo	El cifrado en reposo es opcional para los dominios.	El cifrado en reposo es obligator io para las colecciones.
Configuración y administración de seguridad	Debe configurar la red, el cifrado y el acceso a los datos de forma individual para cada dominio.	Puede usar políticas de seguridad para administrar la configuración de seguridad de varias colecciones a escala.

El siguiente diagrama ejemplifica los componentes de seguridad que componen una colección funcional. A una colección se le debe asignar una clave de cifrado, una configuración de acceso a la red y una política de acceso a los datos coincidente que otorgue permisos a sus recursos.



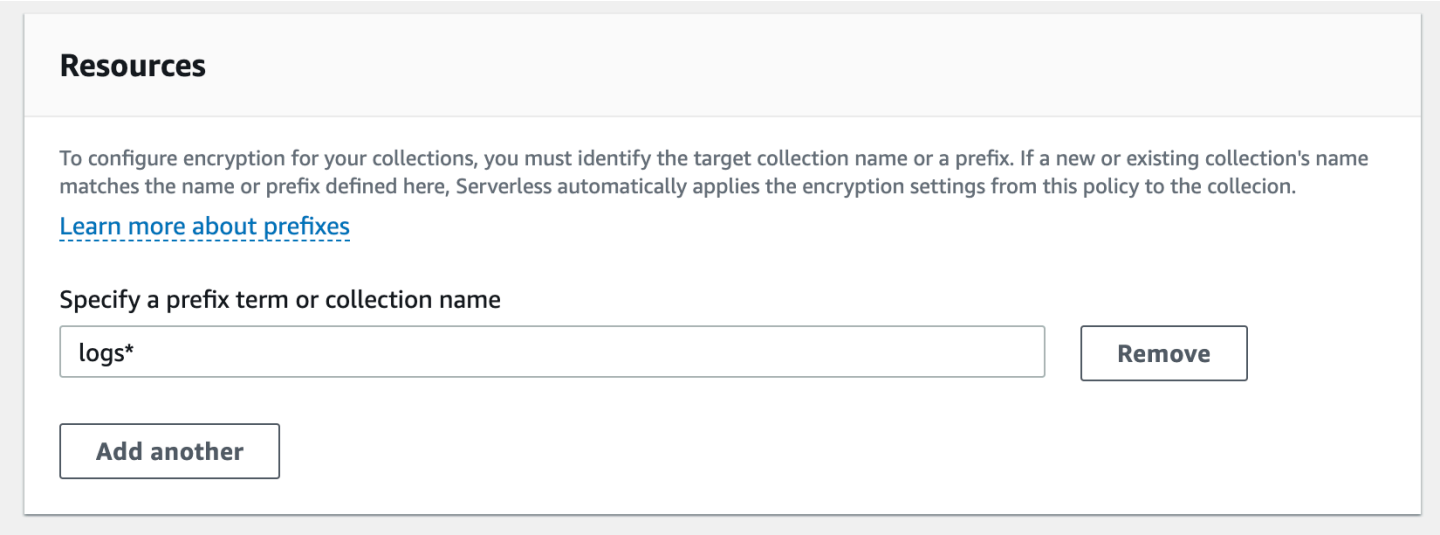
Temas

- Políticas de cifrado
- Políticas de red
- Políticas de acceso a datos
- Autenticación SAML e IAM
- Seguridad de la infraestructura
- Introducción a la seguridad en Amazon OpenSearch Serverless
- Identity and Access Management para Amazon OpenSearch Serverless
- Cifrado en Amazon OpenSearch Serverless
- Acceso a la red para Amazon OpenSearch Serverless
- Conformidad con FIPS en Amazon Serverless OpenSearch
- Control de acceso a datos para Amazon OpenSearch Serverless
- Acceso al plano de datos a través de AWS PrivateLink
- Acceso al plano de control a través de AWS PrivateLink
- Autenticación SAML para Amazon Serverless OpenSearch
- Validación de conformidad para Amazon OpenSearch Serverless

Políticas de cifrado

Las políticas de cifrado definen si sus colecciones se cifran con una clave gestionada por el cliente Clave propiedad de AWS o con una clave gestionada por el cliente. Las políticas de cifrado constan de dos componentes: un patrón de recursos y una clave de cifrado. El patrón de recursos define a qué colección o colecciones se aplica la política. La clave de cifrado determina cómo se protegerán las colecciones asociadas.

Para aplicar una política a varias colecciones debe incluir un comodín (*) en la regla de política. Por ejemplo, la siguiente política se aplica a todas las colecciones cuyos nombres comiencen por “logs” (registros).

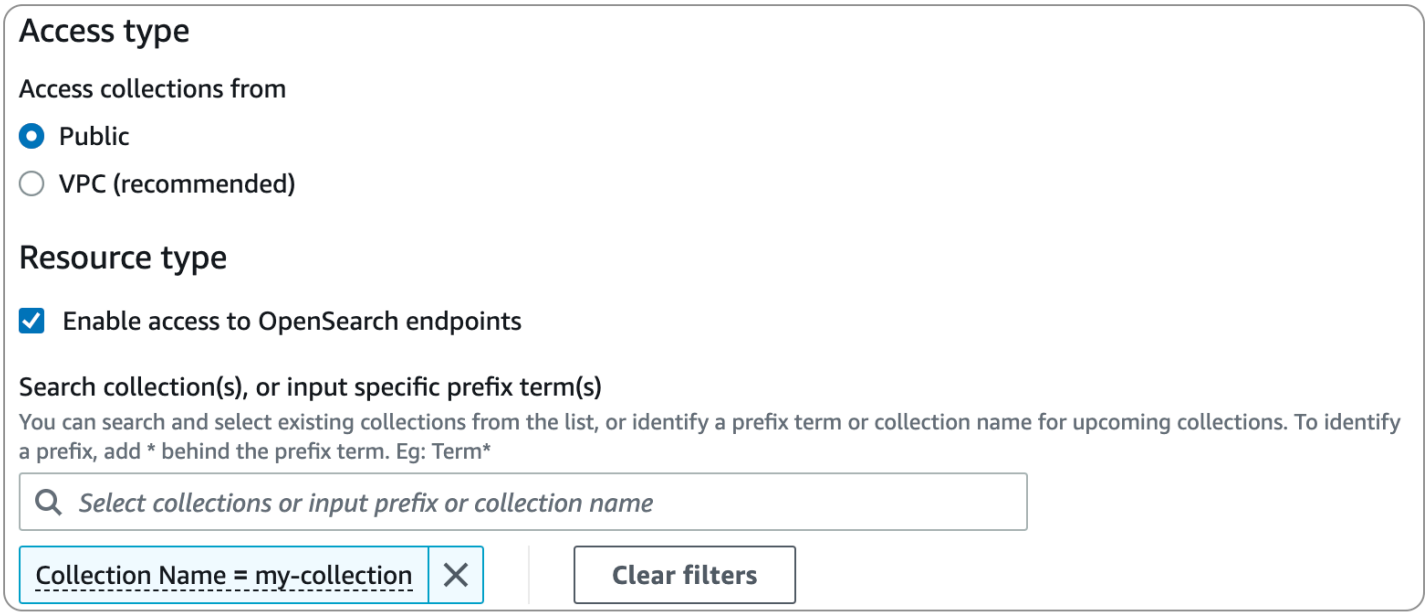


Las políticas de cifrado agilizan el proceso de creación y administración de colecciones, especialmente cuando se hace mediante programación. Puede crear una colección especificando un nombre y, al crearla, se le asigna de forma automática una clave de cifrado.

Políticas de red

Las políticas de red definen si se puede acceder a las colecciones de forma privada o a través de Internet desde redes públicas. Se puede acceder a las colecciones privadas a través de puntos de enlace de VPC OpenSearch gestionados sin servidor o mediante dispositivos específicos, Servicios de AWS como Amazon Bedrock, mediante acceso privado.Servicio de AWS Al igual que las políticas de cifrado, las políticas de red se pueden aplicar a varias colecciones, lo que le permite administrar el acceso a la red para muchas colecciones a gran escala.

Las políticas de red constan de dos componentes: un tipo de acceso y un tipo de recurso. El tipo de acceso puede ser público o privado. El tipo de recurso determina si el acceso que elija se aplica al punto final de la colección, al punto final de Dashboards o a ambos. OpenSearch



Si planea configurar el acceso a la VPC dentro de una política de red, primero debe crear uno o más puntos de enlace de VPC OpenSearch administrados sin servidor. Estos puntos de enlace le permiten acceder a OpenSearch Serverless como si estuviera en su VPC, sin el uso de una puerta de enlace a Internet, un dispositivo NAT, una conexión VPN o una conexión. Direct Connect

El acceso privado a solo Servicios de AWS se puede aplicar al punto final de la colección, no al OpenSearch punto final de Dashboards. OpenSearch Servicios de AWS no se le puede conceder acceso a los OpenSearch paneles.

Políticas de acceso a datos

Las políticas de acceso a datos definen la forma en que los usuarios acceden a los datos de sus colecciones. Las políticas de acceso a datos le ayudan a administrar las colecciones a escala mediante la asignación automática de permisos de acceso a las colecciones e índices que coinciden con un patrón específico. Se pueden aplicar varias políticas a un solo recurso.

Las políticas de acceso a datos constan de un conjunto de reglas, cada una con tres componentes: un tipo de recurso, los recursos concedidos y un conjunto de permisos. El tipo de recurso puede ser una colección o un índice. Los recursos otorgados pueden ser collection/index nombres o patrones con un comodín (*). La lista de permisos especifica a qué operaciones de OpenSearch API concede

acceso la política. Además, la política contiene una lista de entidades principales, que especifican los roles, los usuarios y las identidades SAML de IAM a los que se debe conceder acceso.


Selected principals

Principals

arn:aws:iam::478253424788:user/Administrator

saml/478253424788/myprovider/user/Annie

Granted resources and permissions (2)

Granted resources	Resource type	Permissions
collection/autopartsinventory	collection	aoss:CreateCollectionItems aoss:UpdateCollectionItems
index/test-collection/*	index	aoss:ReadDocument aoss:DescribeIndex

Para obtener más información sobre el formato de una política de acceso a datos, consulte la sintaxis de la política.

Antes de crear una política de acceso a datos, debe tener uno o más roles o usuarios de IAM, o identidades SAML, a los que proporcionar acceso en la política. Para más información, consulte la siguiente sección.

> **Note**
>
> Al cambiar del acceso público al privado para su colección, se eliminará la pestaña Índices de la consola de colecciones OpenSearch sin servidor.

Autenticación SAML e IAM

La entidad principal de IAM y las identidades de SAML son uno de los componentes básicos de una política de acceso a los datos. En la instrucción `principal` de una política de acceso puede incluir roles, usuarios e identidades SAML de IAM. A estas entidades principales se le conceden los permisos que usted especifique en las reglas de política asociadas.

```
[
   {
      "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"
      ]
   }
]
```

La autenticación SAML se configura directamente en Serverless. OpenSearch Para obtener más información, consulte the section called “Autenticación SAML”.

Seguridad de la infraestructura

Amazon OpenSearch Serverless está protegido por la seguridad de AWS la red global. Para obtener información sobre los servicios AWS de seguridad y cómo se AWS protege la infraestructura, consulte Seguridad AWS en la nube. Para diseñar su AWS entorno utilizando las mejores prácticas de seguridad de la infraestructura, consulte Protección de infraestructuras en un marco de buena AWS arquitectura basado en el pilar de la seguridad.

Utiliza las llamadas a la API AWS publicadas para acceder a Amazon OpenSearch Serverless a través de la red. Los clientes deben admitir Transport Layer Security (TLS). Exigimos TLS 1.2 y recomendamos TLS 1.3. Para obtener una lista de los cifrados compatibles con TLS 1.3, consulte los protocolos y cifrados TLS en la documentación de Elastic Load Balancing.

Además, debe firmar las solicitudes mediante un ID de clave de acceso y una clave de acceso secreta que esté asociada a una entidad principal de IAM. También puede utilizar AWS Security Token Service (AWS STS) para generar credenciales de seguridad temporales para firmar solicitudes.

Introducción a la seguridad en Amazon OpenSearch Serverless

Los siguientes tutoriales le ayudan a empezar a utilizar Amazon OpenSearch Serverless. Ambos tutoriales siguen los mismos pasos básicos, pero uno usa la consola mientras que el otro usa la AWS CLI.

Tenga en cuenta que los casos de uso de estos tutoriales están simplificados. Las políticas de red y seguridad son bastante abiertas. En las cargas de trabajo de producción, se recomienda configurar funciones de seguridad más sólidas, como la autenticación SAML, el acceso a la VPC y las políticas de acceso restrictivo a los datos.

Temas

- Tutorial: Introducción a la seguridad en Amazon OpenSearch Serverless (consola)
- Tutorial: Introducción a la seguridad en Amazon OpenSearch Serverless (CLI)

Tutorial: Introducción a la seguridad en Amazon OpenSearch Serverless (consola)

En este tutorial, se explican los pasos básicos para crear y gestionar políticas de seguridad mediante la consola Amazon OpenSearch Serverless.

En este tutorial, debe completar los siguientes pasos:

1. Configurar permisos
2. Crear una política de cifrado
3. Crear una política de red
4. Configurar la política de acceso a datos
5. Crear una colección
6. Cargar y buscar datos

En este tutorial se explica cómo configurar una colección mediante. Consola de administración de AWS Para ver los mismos pasos con el AWS CLI, consultethe section called “Tutorial: Introducción a la seguridad (CLI)”.

Paso 1: Configurar los permisos

> **Note**
>
> Puedes omitir este paso si ya utilizas una política más amplia basada en la identidad, como `Action":"aoss:*"` o `Action":"*"`. Sin embargo, en entornos de producción, le recomendamos que siga la entidad principal del privilegio mínimo y solo asigne los permisos mínimos necesarios para completar una tarea.

Para completar este tutorial, debe tener los permisos de IAM correctos. Su usuario o rol debe tener adjunta una política basada en la identidad con los siguientes permisos mínimos:

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": "*"
    }
  ]
}
```

Para obtener una lista completa de los permisos OpenSearch sin servidor, consultethe section called "Gestión de identidad y acceso".

Paso 2: Crear una política de cifrado

Las políticas de cifrado especifican la AWS KMS clave que OpenSearch Serverless utilizará para cifrar la colección. Puede cifrar las colecciones con una clave Clave administrada de AWS o una clave diferente. Para simplificar este tutorial, cifraremos nuestra colección con un Clave administrada de AWS.

Para crear una política de cifrado

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Expanda Sin servidor en el panel de navegación de la izquierda y seleccione Políticas de red.

3. Seleccione Crear política de cifrado.
4. Asigne el nombre de esta política a los libros de políticas. Para la descripción, introduzca Política de cifrado para la colección de libros.
5. En Recursos, introduzca libros, que es el nombre con el que llamará su colección. Si desea que este sea más extenso, puede incluir un asterisco (`books*`) para que la política se aplique a todas las colecciones que comiencen por la palabra “libros”.
6. Para el cifrado, mantén seleccionada la AWS opción Usar clave propia.
7. Seleccione Crear.

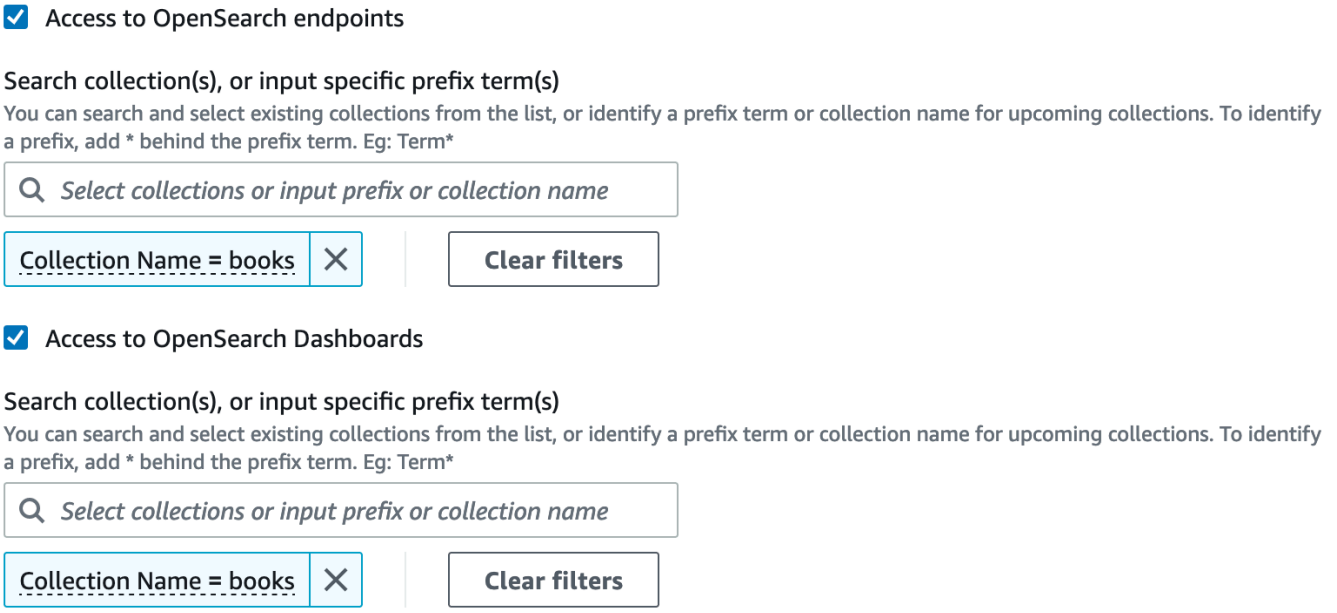
Paso 3: crear una política de red

Las políticas de red determinan si se puede acceder a la colección a través de Internet desde redes públicas o si se debe acceder a ella a través de puntos de conexión de VPC OpenSearch gestionados sin servidor. En este tutorial, configuraremos el acceso público.

Crear una política de red

1. Seleccione Políticas en el panel de navegación de la izquierda y, a continuación, Crear política.
2. Asigne el nombre de esta política a los libros de políticas. Para la descripción, introduzca Política de redes para la colección de libros.
3. En la Regla 1, asígnele un nombre a la regla Acceso público a la colección de libros.
4. Para simplificar este tutorial, configuraremos el acceso público para la colección de libros. Para el tipo de acceso, seleccione Público.
5. Vamos a acceder a la colección desde los paneles. OpenSearch Para ello, debe configurar el acceso a la red para los paneles y el OpenSearch punto final; de lo contrario, los paneles no funcionarán.

   Para el tipo de recurso, habilite el acceso a los OpenSearch puntos finales y el acceso a los paneles. OpenSearch
6. En ambos cuadros de entrada, introduzca Nombre de la colección = libros. Esta configuración reduce el alcance de la política para que solo se aplique a una única colección (`books`). La regla debe tener un aspecto similar al siguiente:



7. Seleccione Crear.

Paso 4: crear una política de acceso a datos

No podrá acceder a los datos de su colección hasta que configure el acceso a los datos. Las políticas de acceso a los datos son independientes de la política basada en identidades de IAM que configuró en el paso 1. Permiten a los usuarios acceder a los datos reales de una colección.

En este tutorial, le brindamos a un solo usuario los permisos necesarios para indexar datos en la colección de libros.

Para crear una política de acceso a datos

1. En el panel de navegación de la izquierda, seleccione Políticas de acceso a datos y Crear política de acceso.
2. Asigne el nombre de esta política a los libros de políticas. Para la descripción, introduzca Política de acceso de datos para la colección de libros.
3. Seleccione JSON para el método de definición de políticas y pegue la siguiente política en el editor de JSON.

   Sustituya el ARN principal por el ARN de la cuenta que usará para iniciar sesión en los OpenSearch paneles e indexar datos.

```
[
```

```
   {
      "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"
      ]
   }
]
```

Esta política proporciona a un solo usuario los permisos mínimos necesarios para crear un índice en la colección de libros, indexar algunos datos y buscarlos.

4. Seleccione Crear.

Paso 5: crear una colección

Como ya configuró las políticas de cifrado y red, puede crear una colección coincidente y la configuración de seguridad se le aplicará automáticamente.

Para crear una colección sin servidor OpenSearch

1. Seleccione Colecciones en el panel de navegación de la izquierda y elija Crear colección.
2. Nombre los libros de la colección.
3. Para el tipo de colección, seleccione Buscar.
4. En Cifrado, OpenSearch Serverless le informa de que el nombre de la colección coincide con la política de `books-policy` cifrado.

5. En la configuración de acceso a la red, OpenSearch Serverless le informa de que el nombre de la colección coincide con la política de `books-policy` red.
6. Elija Siguiente.
7. En las opciones de la política de acceso a los datos, OpenSearch Serverless le informa de que el nombre de la colección coincide con la política de acceso a los `books-policy` datos.
8. Elija Siguiente.
9. Revise la configuración de la colección y seleccione Enviar. Las colecciones suelen tardar menos de un minuto en inicializarse.

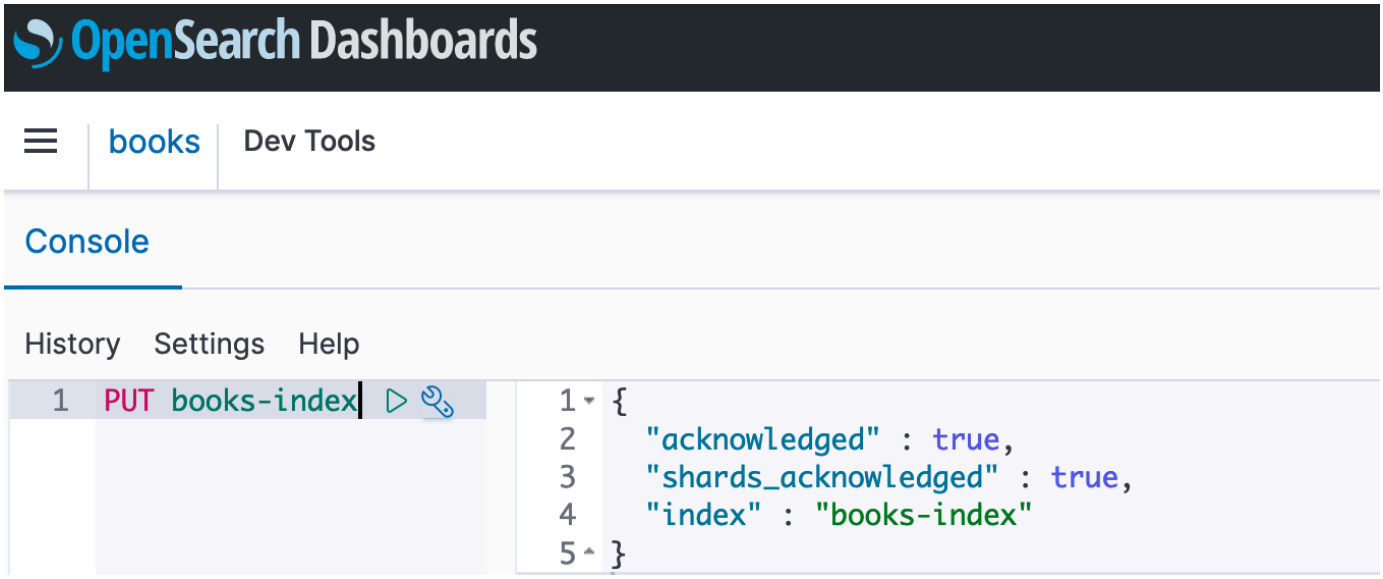
Paso 6: Cargar y buscar datos

Puede cargar datos a una colección OpenSearch sin servidor mediante Postman o curl. Para abreviar, en estos ejemplos se utilizan las herramientas de desarrollo de la consola de Dashboards. OpenSearch

Para indexar y buscar datos en la colección

1. Seleccione Colecciones en el panel de navegación de la izquierda y elija la colección de libros para abrir su página de detalles.
2. Elija la URL de los OpenSearch paneles de control para la colección. La URL toma el formato `https://collection-id.us-east-1.aoss.amazonaws.com/_dashboards`.
3. Inicie sesión en OpenSearch Dashboards con las claves de AWS acceso y secreta del director que especificó en su política de acceso a los datos.
4. En los OpenSearch paneles, abra el menú de navegación de la izquierda y seleccione Herramientas de desarrollo.
5. Para crear un índice único denominado books-index, ejecute el siguiente comando:

```
PUT books-index
```



6. Para indexar un solo documento en books-index, ejecute el siguiente comando:

```
PUT books-index/_doc/1
{
  "title": "The Shining",
  "author": "Stephen King",
  "year": 1977
}
```

7. Para buscar datos en los OpenSearch paneles, debe configurar al menos un patrón de índice. OpenSearch utiliza estos patrones para identificar los índices que desea analizar. Abra el menú principal de Dashboards, seleccione Gestión de pilas, luego seleccione Patrones de índice y, a continuación, seleccione Crear patrón de índice. Para este tutorial, introduzca books-index.
8. seleccione Siguiente paso y, a continuación, seleccione Crear patrón de índice. Una vez creado el patrón, puede ver los diversos campos de documento, como `author` y `title`.
9. Para comenzar a buscar los datos, vuelva a abrir el menú principal y elija Explorar o utilice la API de búsqueda.

Tutorial: Introducción a la seguridad en Amazon OpenSearch Serverless (CLI)

En este tutorial se explican los pasos descritos en el tutorial de introducción a la consola en materia de seguridad, pero se utiliza la consola de servicio AWS CLI en lugar de la consola de OpenSearch servicio.

En este tutorial, deberá completar los siguientes pasos:

1. Creación de una política de permisos de IAM
2. Anexo de la política de IAM a un rol de IAM
3. Crear una política de cifrado
4. Crear una política de red
5. Crear una recopilación
6. Configurar la política de acceso a datos
7. Recuperación del punto de conexión de la colección
8. Carga de los datos a su conexión
9. Búsqueda de datos en su colección

El objetivo de este tutorial es configurar una única colección OpenSearch sin servidor con una configuración bastante sencilla de cifrado, red y acceso a los datos. Por ejemplo, configuraremos el acceso a la red pública, el cifrado y una Clave administrada de AWS política de acceso a los datos simplificada que conceda permisos mínimos a un solo usuario.

En un escenario de producción, considere implementar una configuración más robusta, incluyendo autenticación SAML, una clave de cifrado personalizada y acceso VPC.

Para empezar con las políticas de seguridad en OpenSearch Serverless

1.
> **Note**
>
> Puede omitir este paso si ya utilizas una política más amplia basada en la identidad, como `Action":"aoss:*"` o `Action":"*"`. Sin embargo, en entornos de producción, le recomendamos que siga la entidad principal del privilegio mínimo y solo asigne los permisos mínimos necesarios para completar una tarea.

   Para empezar, cree una AWS Identity and Access Management política con los permisos mínimos necesarios para realizar los pasos de este tutorial. Le pondremos a la política el siguiente nombre `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\": \"*\"}]}"
```

Respuesta de ejemplo

```
{
    "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. Asocie `TutorialPolicy` al rol de IAM que indexará y buscará datos en la colección. Le pondremos al usuario el siguiente nombre `TutorialRole`:

```
aws iam attach-role-policy \
  --role-name TutorialRole \
  --policy-arn arn:aws:iam::123456789012:policy/TutorialPolicy
```

3. Antes de crear una colección, debe crear una política de cifrado que asigne una Clave propiedad de AWS a la colección de libros que creará en un paso posterior.

   Envíe la siguiente solicitud para crear una política de cifrado para la colección de libros:

```
aws opensearchserverless create-security-policy \
  --name books-policy \
  --type encryption --policy "{\"Rules\":[{\"ResourceType\":\"collection\",
\"Resource\":[\"collection\/books\"]}],\"AWSOwnedKey\":true}"
```

   Respuesta de ejemplo

```
{
```

```
    "securityPolicyDetail": {
        "type": "encryption",
        "name": "books-policy",
        "policyVersion": "MTY2OTI0MDAwNTk5MF8x",
        "policy": {
            "Rules": [
                {
                    "Resource": [
                        "collection/books"
                    ],
                    "ResourceType": "collection"
                }
            ],
            "AWSOwnedKey": true
        },
        "createdDate": 1669240005990,
        "lastModifiedDate": 1669240005990
    }
}
```

4. Cree una política de red que proporcione acceso público a la colección de libros:

```
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}]"
```

Respuesta de ejemplo

```
{
    "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 colección de libros:

```
aws opensearchserverless create-collection --name books --type SEARCH
```

Respuesta de ejemplo

```
{
    "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. Cree una política de acceso a datos que proporcione los permisos mínimos para indexar y buscar datos en la colección de libros. Reemplace entidad principal de ARN por el ARN de `TutorialRole` del paso 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\"]}]"
```

Respuesta de ejemplo

```
{
    "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
    }
```

```
}
```

`TutorialRole` ahora debería poder indexar y buscar documentos en la colección de libros.

7. Para realizar llamadas a la OpenSearch API, necesitas el punto final de recopilación. Envíe la siguiente solicitud para recuperar el parámetro `collectionEndpoint`:

```
aws opensearchserverless batch-get-collection --names books
```

Respuesta de ejemplo

```
{
    "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**
>
> No podrá ver el punto de conexión de la colección hasta que el estado de la colección cambie a `ACTIVE`. Puede que tenga que hacer varias llamadas para comprobar el estado hasta que la colección se haya creado de forma correcta.

8. Utilice una herramienta HTTP como Postman o curl para indexar los datos de la colección de libros. Crearemos un índice llamado books-index y agregaremos un solo documento.

Envíe la siguiente solicitud al punto de conexión de la colección que recuperó en el paso anterior, con las credenciales `TutorialRole`.

```
PUT https://8kw362bpwg4gx9b2f6e0.us-east-1.aoss.amazonaws.com/books-index/_doc/1
{
  "title": "The Shining",
  "author": "Stephen King",
  "year": 1977
}
```

Respuesta de ejemplo

```
{
  "_index" : "books-index",
  "_id" : "1",
  "_version" : 1,
  "result" : "created",
  "_shards" : {
    "total" : 0,
    "successful" : 0,
    "failed" : 0
  },
  "_seq_no" : 0,
  "_primary_term" : 0
}
```

9. Para empezar a buscar datos en tu colección, utilice la API de búsqueda. La siguiente consulta representa una búsqueda básica:

```
GET https://8kw362bpwg4gx9b2f6e0.us-east-1.aoss.amazonaws.com/books-index/_search
```

Respuesta de ejemplo

```
{
    "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 para Amazon OpenSearch Serverless

AWS Identity and Access Management (IAM) es una herramienta Servicio de AWS que ayuda al administrador a controlar de forma segura el acceso a los recursos. AWS Los administradores de IAM controlan quién puede autenticarse (iniciar sesión) y quién puede autorizarse (tener permisos) para usar OpenSearch los recursos sin servidor. La IAM es una opción Servicio de AWS que puede utilizar sin coste adicional.

Temas

- Políticas basadas en la identidad para Serverless OpenSearch
- Acciones políticas para Serverless OpenSearch
- Recursos de políticas para Serverless OpenSearch
- Claves de condición de la política para Amazon OpenSearch Serverless
- ABAC con Serverless OpenSearch
- Uso de credenciales temporales con Serverless OpenSearch

- Funciones vinculadas a servicios para Serverless OpenSearch
- Otros tipos de políticas
- Ejemplos de políticas basadas en la identidad para sistemas sin servidor OpenSearch
- Soporte de IAM Identity Center para Amazon Serverless OpenSearch

Políticas basadas en la identidad para Serverless OpenSearch

Compatibilidad con las políticas basadas en identidad: sí

Las políticas basadas en identidad son documentos de políticas de permisos JSON que puede asociar a una identidad, como un usuario de IAM, un grupo de usuarios o un rol. Estas políticas controlan qué acciones pueden realizar los usuarios y los roles, en qué recursos y en qué condiciones. Para obtener más información sobre cómo crear una política basada en la identidad, consulte Definición de permisos de IAM personalizados con políticas administradas por el cliente en la Guía del usuario de IAM.

Con las políticas basadas en identidades de IAM, puede especificar las acciones y los recursos permitidos o denegados, así como las condiciones en las que se permiten o deniegan las acciones. Para obtener más información sobre los elementos que puede utilizar en una política de JSON, consulte Referencia de los elementos de la política de JSON de IAM en la Guía del usuario de IAM.

Ejemplos de políticas basadas en la identidad para Serverless OpenSearch

Para ver ejemplos de políticas basadas en la identidad de OpenSearch Serverless, consulte. the section called “Ejemplos de políticas basadas en identidades”

Acciones políticas para Serverless OpenSearch

Compatibilidad con las acciones de políticas: sí

El elemento `Action` de una política JSON describe las acciones que puede utilizar para conceder o denegar el acceso en una política. Las acciones de política suelen tener el mismo nombre que la operación de AWS API asociada. Hay algunas excepciones, como acciones de solo permiso que no tienen una operación de API coincidente. También hay algunas operaciones que requieren varias acciones en una política. Estas acciones adicionales se denominan acciones dependientes.

Incluya acciones en una política para conceder permisos y así llevar a cabo la operación asociada.

Las acciones de política en OpenSearch Serverless utilizan el siguiente prefijo antes de la acción:

```
aoss
```

Para especificar varias acciones en una única instrucción, sepárelas con comas.

```
"Action": [
      "aoss:action1",
      "aoss:action2"
         ]
```

Puede especificar varias acciones utilizando caracteres comodín (*). Por ejemplo, para especificar todas las acciones que comiencen con la palabra `Describe`, incluya la siguiente acción:

```
"Action": "aoss:List*"
```

Para ver ejemplos de políticas basadas en la identidad de OpenSearch Serverless, consulte. Ejemplos de políticas basadas en la identidad para Serverless OpenSearch

Recursos de políticas para Serverless OpenSearch

Compatibilidad con los recursos de políticas: sí

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Resource` de la política JSON especifica el objeto u objetos a los que se aplica la acción. Como práctica recomendada, especifique un recurso utilizando el Nombre de recurso de Amazon (ARN). En el caso de las acciones que no admiten permisos por recurso, utilice un carácter comodín (*) para indicar que la instrucción se aplica a todos los recursos.

```
"Resource": "*"
```

Claves de condición de la política para Amazon OpenSearch Serverless

Compatibilidad con claves de condición de políticas específicas del servicio: sí

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Condition` especifica cuándo se ejecutan las instrucciones en función de criterios definidos. Puede crear expresiones condicionales que utilizan operadores de condición, tales como

igual o menor que, para que la condición de la política coincida con los valores de la solicitud. Para ver todas las claves de condición AWS globales, consulte las claves de contexto de condición AWS globales en la Guía del usuario de IAM.

Además del control de acceso basado en atributos (ABAC), OpenSearch Serverless admite las siguientes claves de condición:

- `aoss:collection`
- `aoss:CollectionId`
- `aoss:index`

Puede utilizar estas claves de condición incluso al proporcionar políticas de acceso y de seguridad. Por ejemplo:

```
[
   {
      "Effect":"Allow",
      "Action":[
         "aoss:CreateAccessPolicy",
         "aoss:CreateSecurityPolicy"
      ],
      "Resource":"*",
      "Condition":{
         "StringLike":{
            "aoss:collection":"log"
         }
      }
   }
]
```

En este ejemplo, la condición se aplica a las políticas que contienen reglas que coinciden tanto con el nombre como con el patrón de una colección. Las condiciones funcionan de la el siguiente manera:

- `StringEquals`: se aplica a políticas con reglas que contienen la cadena de exacta de recursos “log” (es decir,`collection/log`).
- `StringLike`: se aplica a políticas con reglas que contienen la cadena de recursos que incluye la cadena “log” (es decir, `collection/log` pero también `collection/logs-application` o `collection/applogs123`).

> **Note**
>
> Las claves de condiciones de colección no se aplican a nivel de índice. Por ejemplo, en la política anterior, la condición no se aplicaría a una política de acceso o seguridad que contenga la cadena de recursos`index/logs-application/*`.

Para ver una lista de claves de condición de OpenSearch Serverless, consulte Claves de condición de Amazon OpenSearch Serverless en la Referencia de autorización de servicio. Para saber con qué acciones y recursos puede utilizar una clave de condición, consulte Acciones definidas por Amazon OpenSearch Serverless.

ABAC con Serverless OpenSearch

Admite ABAC (etiquetas en las políticas): sí

El control de acceso basado en atributos (ABAC) es una estrategia de autorización que define permisos en función de atributos denominados etiquetas. Puede adjuntar etiquetas a las entidades y AWS los recursos de IAM y, a continuación, diseñar políticas de ABAC para permitir las operaciones cuando la etiqueta del principal coincida con la etiqueta del recurso.

Para controlar el acceso en función de etiquetas, debe proporcionar información de las etiquetas en el elemento de condición de una política utilizando las claves de condición `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` o `aws:TagKeys`.

Si un servicio admite las tres claves de condición para cada tipo de recurso, el valor es Sí para el servicio. Si un servicio admite las tres claves de condición solo para algunos tipos de recursos, el valor es Parcial.

Para obtener más información sobre ABAC, consulte Definición de permisos con la autorización de ABAC en la Guía del usuario de IAM. Para ver un tutorial con los pasos para configurar ABAC, consulte Uso del control de acceso basado en atributos (ABAC) en la Guía del usuario de IAM.

Para obtener más información sobre el etiquetado de recursos OpenSearch sin servidor, consulte. the section called “Etiquetado de colecciones”

Uso de credenciales temporales con Serverless OpenSearch

Compatibilidad con credenciales temporales: sí

Las credenciales temporales proporcionan acceso a AWS los recursos a corto plazo y se crean automáticamente cuando se utiliza la federación o se cambia de rol. AWS recomienda generar credenciales temporales de forma dinámica en lugar de utilizar claves de acceso a largo plazo. Para obtener más información, consulte Credenciales de seguridad temporales en IAM y Servicios de AWS que funcionan con IAM en la Guía del usuario de IAM.

Funciones vinculadas a servicios para Serverless OpenSearch

Compatible con roles vinculados al servicio: sí

Un rol vinculado a un servicio es un tipo de rol de servicio que está vinculado a un. Servicio de AWS El servicio puede asumir el rol para realizar una acción en su nombre. Los roles vinculados al servicio aparecen en usted Cuenta de AWS y son propiedad del servicio. Un administrador de IAM puede ver, pero no editar, los permisos de los roles vinculados a servicios.

Para obtener más información sobre la creación y administración de funciones vinculadas a servicios OpenSearch sin servidor, consulte. the section called “Rol de creación de colecciones”

Otros tipos de políticas

AWS admite tipos de políticas adicionales y menos comunes. Estos tipos de políticas pueden establecer el máximo de permisos que los tipos de políticas más frecuentes le conceden.

- Políticas de control de servicios (SCPs): SCPs son políticas de JSON que especifican los permisos máximos para una organización o unidad organizativa (OU). AWS Organizations AWS Organizations es un servicio para agrupar y administrar de forma centralizada varias AWS cuentas que son propiedad de su empresa. Si habilitas todas las funciones de una organización, puedes aplicar políticas de control de servicios (SCPs) a una o a todas tus cuentas. El SCP limita los permisos de las entidades en las cuentas de los miembros, incluido cada usuario raíz de la AWS cuenta. Para obtener más información sobre Organizations SCPs, consulte las políticas de control de servicios en la Guía del AWS Organizations usuario.
- Políticas de control de recursos (RCPs): RCPs son políticas de JSON que puedes usar para establecer los permisos máximos disponibles para los recursos de tus cuentas sin actualizar las políticas de IAM asociadas a cada recurso que poseas. La RCP limita los permisos de los recursos en las cuentas de miembros y puede afectar a los permisos efectivos de las identidades, incluidos los usuarios raíz de la cuenta de AWS , independientemente de si pertenecen a su organización. Para obtener más información sobre Organizations e RCPs incluir una lista de AWS los servicios compatibles RCPs, consulte Políticas de control de recursos (RCPs) en la Guía del AWS Organizations usuario.

Ejemplos de políticas basadas en la identidad para sistemas sin servidor OpenSearch

De forma predeterminada, los usuarios y los roles no tienen permiso para crear o modificar OpenSearch recursos de Serverless. Un administrador de IAM puede crear políticas de IAM para conceder permisos a los usuarios para realizar acciones en los recursos que necesitan.

Para obtener información acerca de cómo crear una política basada en identidades de IAM mediante el uso de estos documentos de políticas JSON de ejemplo, consulte Creación de políticas de IAM (consola) en la Guía del usuario de IAM.

Para obtener más información sobre las acciones y los tipos de recursos definidos por Amazon OpenSearch Serverless, incluido el ARNs formato de cada uno de los tipos de recursos, consulte Acciones, recursos y claves de condición de Amazon OpenSearch Serverless en la Referencia de autorización de servicios.

Temas

- Prácticas recomendadas sobre las políticas
- Uso de Serverless en la consola OpenSearch
- Administrar colecciones OpenSearch sin servidor
- Visualización de colecciones OpenSearch sin servidor
- Uso de operaciones OpenSearch de API
- ABAC para operaciones de API OpenSearch

Prácticas recomendadas sobre las políticas

Las políticas basadas en identidades son muy eficaces. Determinan si alguien puede crear, eliminar o acceder a los recursos de OpenSearch Serverless de su cuenta. Estas acciones pueden generar costos adicionales para su Cuenta de AWS. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidades:

Las políticas basadas en la identidad determinan si alguien puede crear recursos OpenSearch sin servidor de su cuenta, acceder a ellos o eliminarlos. Estas acciones pueden generar costos adicionales para su Cuenta de AWS. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidades:

- Comience con las políticas AWS administradas y avance hacia los permisos con privilegios mínimos: para empezar a conceder permisos a sus usuarios y cargas de trabajo, utilice las políticas AWS administradas que otorgan permisos para muchos casos de uso comunes. Están

disponibles en su. Cuenta de AWS Le recomendamos que reduzca aún más los permisos definiendo políticas administradas por el AWS cliente que sean específicas para sus casos de uso. Con el fin de obtener más información, consulte las políticas administradas por AWS o las políticas administradas por AWS para funciones de tarea en la Guía de usuario de IAM.

- Aplique permisos de privilegio mínimo: cuando establezca permisos con políticas de IAM, conceda solo los permisos necesarios para realizar una tarea. Para ello, debe definir las acciones que se pueden llevar a cabo en determinados recursos en condiciones específicas, también conocidos como permisos de privilegios mínimos. Con el fin de obtener más información sobre el uso de IAM para aplicar permisos, consulte Políticas y permisos en IAM en la Guía del usuario de IAM.
- Utilice condiciones en las políticas de IAM para restringir aún más el acceso: puede agregar una condición a sus políticas para limitar el acceso a las acciones y los recursos. Por ejemplo, puede escribir una condición de políticas para especificar que todas las solicitudes deben enviarse utilizando SSL. También puedes usar condiciones para conceder el acceso a las acciones del servicio si se utilizan a través de una acción específica Servicio de AWS, por ejemplo CloudFormation. Para obtener más información, consulte Elementos de la política de JSON de IAM: Condición en la Guía del usuario de IAM.
- Utiliza el analizador de acceso de IAM para validar las políticas de IAM con el fin de garantizar la seguridad y funcionalidad de los permisos: el analizador de acceso de IAM valida políticas nuevas y existentes para que respeten el lenguaje (JSON) de las políticas de IAM y las prácticas recomendadas de IAM. El analizador de acceso de IAM proporciona más de 100 verificaciones de políticas y recomendaciones procesables para ayudar a crear políticas seguras y funcionales. Para más información, consulte Validación de políticas con el Analizador de acceso de IAM en la Guía del usuario de IAM.
- Requerir autenticación multifactor (MFA): si tiene un escenario que requiere usuarios de IAM o un usuario raíz en Cuenta de AWS su cuenta, active la MFA para mayor seguridad. Para exigir la MFA cuando se invoquen las operaciones de la API, añada condiciones de MFA a sus políticas. Para más información, consulte Acceso seguro a la API con MFA en la Guía del usuario de IAM.

Para obtener más información sobre las prácticas recomendadas de IAM, consulte Prácticas recomendadas de seguridad en IAM en la Guía del usuario de IAM.

Uso de Serverless en la consola OpenSearch

Para acceder a OpenSearch Serverless desde la consola de OpenSearch servicio, debe tener un conjunto mínimo de permisos. Estos permisos deben permitirle enumerar y ver detalles sobre los recursos de OpenSearch Serverless de su AWS cuenta. Si se crea una política basada en

identidades que es más restrictiva que los permisos necesarios mínimos, la consola no funcionará del modo esperado para las entidades (tales como roles de IAM) que tengan esa política.

No es necesario que concedas permisos mínimos de consola a los usuarios que solo realizan llamadas a la API AWS CLI o a la AWS API. En su lugar, permite acceso únicamente a las acciones que coincidan con la operación de API que intenta realizar.

La siguiente política permite a un usuario acceder a OpenSearch Serverless desde la consola OpenSearch de servicio:

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"
            ]
        }
    ]
}
```

Administrar colecciones OpenSearch sin servidor

Esta política es un ejemplo de una política de «administración de colecciones» que permite a un usuario gestionar y administrar las colecciones de Amazon OpenSearch Serverless. El usuario puede crear, ver y eliminar colecciones.

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"
        }
    ]
}
```

Visualización de colecciones OpenSearch sin servidor

Este ejemplo de política permite a un usuario ver los detalles de todas las colecciones de Amazon OpenSearch Serverless de su cuenta. El usuario no puede modificar las colecciones ni las políticas de seguridad asociadas.

JSON

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

```
            "Action": [
                "aoss:ListAccessPolicies",
                "aoss:ListCollections",
                "aoss:ListSecurityPolicies",
                "aoss:ListTagsForResource",
                "aoss:BatchGetCollection"
            ],
            "Effect": "Allow"
        }
    ]
}
```

Uso de operaciones OpenSearch de API

Las operaciones de la API del plano de datos consisten en las funciones que se utilizan en OpenSearch Serverless para obtener el valor del servicio en tiempo real. Las operaciones de la API del plano de control consisten en las funciones que se utilizan para configurar el entorno.

Para acceder al plano de datos APIs y a los OpenSearch paneles de Amazon OpenSearch Serverless desde el navegador, debe añadir dos permisos de IAM para los recursos de recopilación. Estos permisos son `aoss:APIAccessAll` y `aoss:DashboardsAccessAll`.

> **Note**
>
> A partir del 10 de mayo de 2023, OpenSearch Serverless requerirá estos dos nuevos permisos de IAM para los recursos de recopilación. El `aoss:APIAccessAll` permiso permite el acceso al plano de datos y el `aoss:DashboardsAccessAll` permiso permite acceder a los OpenSearch paneles desde el navegador. Si no se agregan los dos nuevos permisos de IAM, se produce un error 403.

Este ejemplo de política permite a un usuario acceder al plano de datos APIs de una colección específica de su cuenta y acceder a los OpenSearch paneles de control de todas las colecciones de su cuenta.

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"
        }
    ]
}
```

Ambos `aoss:APIAccessAll` `aoss:DashboardsAccessAll` otorgan permisos de IAM completos a los recursos de la colección, mientras que el permiso de paneles también proporciona OpenSearch acceso a los paneles. Cada permiso funciona de forma independiente, por lo que una denegación explícita de `aoss:APIAccessAll` no bloquea el acceso `aoss:DashboardsAccessAll` a los recursos, incluidas las herramientas para desarrolladores. Lo mismo ocurre con una denegación. `aoss:DashboardsAccessAll` OpenSearch Serverless admite las siguientes claves de condición globales:

- `aws:CalledVia`
- `aws:CalledViaAWSService`
- `aws:CalledViaFirst`
- `aws:CalledViaLast`
- `aws:CurrentTime`
- `aws:EpochTime`
- `aws:PrincipalAccount`
- `aws:PrincipalArn`
- `aws:PrincipalIsAWSService`
- `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

El siguiente es un ejemplo de cómo usar aws:SourceIp en el bloque de condición de la política de IAM de su principal para llamadas al plano de datos:

```
"Condition": {
    "IpAddress": {
         "aws:SourceIp": "203.0.113.0"
    }
}
```

El siguiente es un ejemplo de cómo usar aws:SourceVpc en el bloque de condición de la política de IAM de su principal para llamadas al plano de datos:

```
"Condition": {
    "StringEquals": {
        "aws:SourceVpc": "vpc-0fdd2445d8EXAMPLE"
    }
}
```

Además, se ofrece soporte para las siguientes claves específicas de OpenSearch Serverless:

- aoss:CollectionId
- aoss:collection

El siguiente es un ejemplo de cómo usar `aoss:collection` en el bloque de condición de la política de IAM de su principal para llamadas al plano de datos:

```
"Condition": {
    "StringLike": {
         "aoss:collection": "log-*"
    }
}
```

ABAC para operaciones de API OpenSearch

Las políticas basadas en la identidad le permiten usar etiquetas para controlar el acceso al plano de datos de Amazon OpenSearch Serverless. APIs La siguiente política es un ejemplo para permitir que los directores adjuntos accedan al plano de datos APIs si la colección tiene la etiqueta: `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 siguiente política es un ejemplo para denegar a los directores adjuntos el acceso al plano de datos APIs y a los paneles si la colección tiene la etiqueta: `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 no admite claves `RequestTag` de condición `TagKeys` globales para el plano de datos. APIs

Soporte de IAM Identity Center para Amazon Serverless OpenSearch

Soporte de IAM Identity Center para Amazon Serverless OpenSearch

Puede utilizar los elementos principales del IAM Identity Center (usuarios y grupos) para acceder a los datos de Amazon OpenSearch Serverless a través de Amazon Applications. OpenSearch Para habilitar la compatibilidad con IAM Identity Center para Amazon OpenSearch Serverless, necesitará habilitar el uso de IAM Identity Center. Para obtener más información sobre cómo hacerlo, consulte ¿Qué es Centro de identidades de IAM?

> **Note**
>
> Para acceder a las colecciones de Amazon OpenSearch Serverless mediante usuarios o grupos del Centro de Identidad de IAM, debe utilizar la función de OpenSearch interfaz de usuario (aplicaciones). No se admite el acceso directo a los paneles de control OpenSearch sin servidor con las credenciales del IAM Identity Center. Para obtener más información, consulte Introducción a la OpenSearch interfaz de usuario.

Una vez creada la instancia del IAM Identity Center, el administrador de la cuenta del cliente debe crear una aplicación del IAM Identity Center para el servicio Amazon OpenSearch Serverless. Esto se puede hacer llamando al:. CreateSecurityConfig El administrador de la cuenta del cliente puede especificar qué atributos se usarán para autorizar la solicitud. Los atributos predeterminados que se usan son `UserId` y `GroupId`..

La integración del Centro de identidad de IAM para Amazon OpenSearch Serverless utiliza los siguientes permisos del Centro de identidad de AWS IAM (IAM):

- `aoss:CreateSecurityConfig`: crear un proveedor de Centro de identidades de IAM
- `aoss:ListSecurityConfig`: enumerar todos los proveedores de Centro de identidades de IAM en la cuenta actual.
- `aoss:GetSecurityConfig`: ver información del proveedor de Centro de identidades de IAM.
- `aoss:UpdateSecurityConfig`: modificar una configuración determinada de Centro de identidades de IAM
- `aoss:DeleteSecurityConfig`: eliminación de un proveedor de confianza de Centro de identidades de IAM.

La siguiente política de acceso basada en la identidad se puede utilizar para gestionar todas las configuraciones del centro de identidad de IAM:

JSON

```
{
"Version": "2012-10-17",
    "Statement": [
        {
"Action": [
                "aoss:CreateSecurityConfig",
                "aoss:DeleteSecurityConfig",
                "aoss:GetSecurityConfig",
                "aoss:UpdateSecurityConfig",
                "aoss:ListSecurityConfigs"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
```

```
}
```

> **Note**
>
> El elemento `Resource` debe ser un comodín.

Crear un proveedor de Centro de identidades de IAM (consola)

Puede crear un proveedor de centros de identidad de IAM para habilitar la autenticación con la aplicación. OpenSearch Para habilitar la autenticación del IAM Identity Center para los OpenSearch paneles, lleve a cabo los siguientes pasos:

1. Inicia sesión en la consola OpenSearch de Amazon Service.
2. En el panel de navegación izquierdo, expanda sin servidor y elija Autenticación.
3. Seleccione Autenticación de Centro de identidades de IAM.
4. Seleccione Editar.
5. Marque la casilla junto a Autenticarse con Centro de identidades de IAM.
6. Seleccione la clave de atributo de usuario y grupo del menú desplegable. Los atributos de usuario se usarán para autorizar a los usuarios según `UserName`, `UserId` y `Email`. Los atributos de grupo se usarán para autenticar a los usuarios según `GroupName` y `GroupId`.
7. Seleccione la instancia de Centro de identidades de IAM.
8. Seleccione Guardar.

Crear un proveedor de Centro de identidades de IAM (AWS CLI)

Para crear un proveedor de un centro de identidad de IAM mediante AWS Command Line Interface (AWS CLI), utilice el siguiente 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"
}'
```

Después de habilitar un Centro de identidades de IAM, los clientes solo pueden modificar los atributos de usuario y grupo.

```
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"
}'
```

Para ver el proveedor del centro de identidad de IAM mediante el AWS Command Line Interface, utilice el siguiente comando:

```
aws opensearchserverless list-security-configs --type iamidentitycenter
```

Eliminar un proveedor de Centro de identidades de IAM

El IAM Identity Center ofrece dos instancias de proveedores, una para su cuenta de organización y otra para su cuenta de miembro. Si necesita cambiar su instancia del Centro de identidades de IAM, debe eliminar su configuración de seguridad a través de la API de `DeleteSecurityConfig` y crear una nueva configuración de seguridad usando la nueva instancia del Centro de identidades de IAM. El siguiente comando se puede usar para eliminar un proveedor de Centro de identidades de IAM:

```
aws opensearchserverless delete-security-config \
--region us-east-1 \
--id <id_from_list_security_configs>
```

Conceder acceso de Centro de identidades de IAM a los datos de la colección

Después de habilitar su proveedor del Centro de identidades de IAM, puede actualizar la política de acceso a los datos de la colección para incluir entidades principales del Centro de identidades de IAM. Las entidades principales de Centro de identidades de IAM deben actualizarse en el siguiente formato:

```
[
   {
"Rules":[
       ...
      ],
      "Principal":[
         "iamidentitycenter/<iamidentitycenter-instance-id>/user/<UserName>",
         "iamidentitycenter/<iamidentitycenter-instance-id>/group/<GroupId>"
      ]
   }
]
```

> **Note**
>
> Amazon OpenSearch Serverless solo admite una instancia del centro de identidad de IAM para todas las colecciones de clientes y puede admitir hasta 100 grupos para un solo usuario. Si intenta usar más de la cantidad de instancias permitidas, experimentará inconsistencias con el procesamiento de autorización de la política de acceso a datos y recibirá un mensaje de error `403`.

Puede concederles acceso a colecciones, índices o ambos. Si desea que diferentes usuarios tengan distintos permisos, deberá crear múltiples reglas. Para ver una lista de los permisos disponibles, consulta Identity and Access Management in Amazon OpenSearch Service. Para obtener información sobre cómo formatear una política de acceso, consulte Conceder acceso a identidades SAML a los datos de la colección.

Cifrado en Amazon OpenSearch Serverless

Cifrado en reposo

Cada colección de Amazon OpenSearch Serverless que cree está protegida con el cifrado de los datos en reposo, una característica de seguridad que ayuda a evitar el acceso no autorizado a sus datos. El cifrado en reposo utiliza AWS Key Management Service (AWS KMS) para almacenar y administrar sus claves de cifrado. Utiliza el algoritmo estándar de cifrado avanzado con claves de 256 bits (AES-256) para realizar el cifrado.

Temas

- Políticas de cifrado

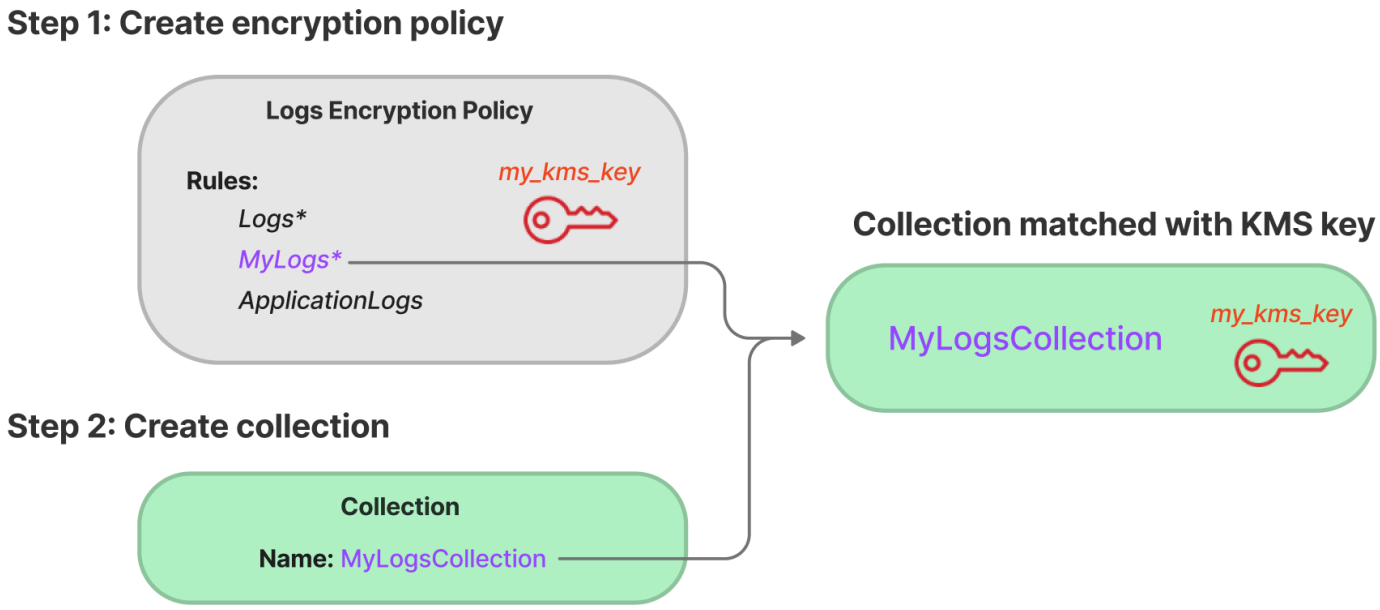
- Consideraciones
- Permisos necesarios
- Política de claves para una clave administrada por el cliente
- Cómo utiliza Serverless las subvenciones en OpenSearch AWS KMS
- Creación de políticas de cifrado (consola)
- Creación de políticas de cifrado (AWS CLI)
- Visualización de políticas de cifrado
- Actualización de políticas de cifrado
- Eliminación de políticas de cifrado

Políticas de cifrado

Con las políticas de cifrado, puede administrar numerosas colecciones a escala mediante la asignación automática de una clave de cifrado a las colecciones más recientes que coincidan con un nombre o un patrón específicos.

Al crear una política de cifrado, puede especificar un prefijo, que es una regla de coincidencia basada en caracteres comodín, como `MyCollection*`, o escribir un nombre de colección único. Después, cuando se cree una colección que coincida con ese patrón de nombre o prefijo, se le asignarán de forma automática la política y la clave de KMS correspondientes.

Al crear una colección, puede especificar una AWS KMS clave de dos maneras: mediante políticas de seguridad o directamente en la `CreateCollection` solicitud. Si proporciona una AWS KMS clave como parte de la `CreateCollection` solicitud, esta tendrá prioridad sobre cualquier política de seguridad coincidente. Con este enfoque, tiene la flexibilidad de anular la configuración de cifrado basada en políticas para colecciones específicas cuando sea necesario.



Las políticas de cifrado contienen los siguientes elementos:

- `Rules`: una o más reglas de coincidencia de colecciones, cada una con los siguientes subelementos:
  - `ResourceType`: en la actualidad, la única opción es “colección”. Las políticas de cifrado solo se aplican a los recursos de colecciones.
  - `Resource`: uno o más nombres o patrones de colección a los que se aplicará la política, en el formato `collection/<collection name|pattern>`.
- `AWSOwnedKey`: si se utiliza una Clave propiedad de AWS o no.
- `KmsARN`: si configura el valor `AWSOwnedKey` en falso, especifique el nombre de recurso de Amazon (ARN) de la clave KMS con la que se va a cifrar las colecciones asociadas. Si incluye este parámetro, OpenSearch Serverless lo ignora. `AWSOwnedKey`

El siguiente ejemplo de política asignará una clave administrada por el cliente a cualquier colección futura denominada `autopartsinventory`, así como a las colecciones que comiencen por el término “ventas”:

```
{
   "Rules":[
      {
         "ResourceType":"collection",
         "Resource":[
```

```
            "collection/autopartsinventory",
            "collection/sales*"
         ]
      }
   ],
   "AWSOwnedKey":false,
   "KmsARN":"arn:aws:kms:us-east-1:123456789012:key/93fd6da4-a317-4c17-
bfe9-382b5d988b36"
}
```

Incluso si una política coincide con el nombre de una colección, puede optar por anular esta asignación automática durante la creación de la colección si el patrón de recursos contiene un comodín (*). Si decide anular la asignación automática de claves, OpenSearch Serverless crea una política de cifrado para usted denominada auto-< ***collection-name*** > y la adjunta a la colección. En un principio, la política solo se aplica a una única colección, pero se puede modificar para que incluya colecciones adicionales.

Si modifica las reglas de políticas para que dejen de coincidir con una colección, la clave KMS asociada continuará asignada a esa colección. La colección siempre permanece cifrada con su clave de cifrado inicial. Si desea cambiar la clave de cifrado de una colección, debe volver a crear la colección.

Si las reglas de varias políticas coinciden con una colección, se utiliza la regla más específica. Por ejemplo, si una política contiene una regla para `collection/log*` y otra para `collection/logSpecial`, se utiliza la clave de cifrado de la segunda política porque es más específica.

No puede usar un nombre o un prefijo en una política si ya existe en otra política. OpenSearch Serverless muestra un error si intenta configurar patrones de recursos idénticos en distintas políticas de cifrado.

Consideraciones

Cuando configure el cifrado de las colecciones, tenga en cuenta lo siguiente:

- El cifrado en reposo es obligatorio para todas las colecciones sin servidor.
- Puede elegir entre utilizar una clave administrada por el cliente o una Clave propiedad de AWS. Si elige una clave administrada por el cliente, le recomendamos habilitar la rotación automática de claves.
- No se puede cambiar la clave de cifrado de una colección después de su creación. Elija cuidadosamente cuál AWS KMS desea usar la primera vez que configure una colección.

- Una colección solo puede coincidir con una única política de cifrado.
- Las colecciones con claves KMS únicas no pueden compartir OpenSearch Compute Units (OCUs) con otras colecciones. Cada colección con una clave única requiere sus propias 4 OCUs.
- Si actualiza la clave KMS de una política de cifrado, el cambio no afectará a las colecciones coincidentes existentes con claves KMS ya asignadas.
- OpenSearch Serverless no comprueba explícitamente los permisos de usuario en las claves administradas por el cliente. Si un usuario tiene permisos para acceder a una colección mediante una política de acceso a los datos, podrá incorporar y consultar los datos cifrados con la clave asociada.

Permisos necesarios

El cifrado en reposo para OpenSearch Serverless utiliza los siguientes permisos AWS Identity and Access Management (IAM). Puede especificar las condiciones de IAM para restringir a los usuarios a colecciones específicas.

- `aoss:CreateSecurityPolicy`: cree una política de cifrado.
- `aoss:ListSecurityPolicies` enumere todas las políticas y colecciones de cifrado a las que están adjuntas.
- `aoss:GetSecurityPolicy`: consulte los detalles de una política de cifrado específica.
- `aoss:UpdateSecurityPolicy`: modifique una política de cifrado.
- `aoss:DeleteSecurityPolicy`: elimine una política de cifrado.

El siguiente ejemplo de política de acceso basada en la identidad proporciona los permisos mínimos necesarios para que un usuario administre las políticas de cifrado con el patrón de recursos `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":"*"
      }
   ]
}
```

Política de claves para una clave administrada por el cliente

Si selecciona una clave administrada por el cliente para proteger una colección, OpenSearch Serverless obtiene permiso para usar la clave KMS en nombre del responsable que realiza la selección. Ese principal, un usuario o un rol, debe tener los permisos en la clave de KMS que requiere OpenSearch Serverless. Puede proporcionar estos permisos en una política de claves o en una política de IAM.

OpenSearch Serverless realiza `GenerateDataKey` llamadas a la API de `Decrypt` KMS durante las operaciones de mantenimiento, como el escalado automático y las actualizaciones de software. Es posible que observe que estas llamadas están fuera de sus patrones de tráfico habituales. Estas llamadas forman parte de las operaciones normales del servicio y no indican que haya tráfico de usuarios activos.

OpenSearch Serverless lanza una `KMSKeyInaccessibleException` cuando no puede acceder a la clave KMS que cifra los datos en reposo. Esto ocurre al deshabilitar o eliminar la clave KMS, o al revocar las concesiones que permiten a OpenSearch Serverless usar la clave.

Como mínimo, OpenSearch Serverless requiere los siguientes permisos en una clave administrada por el cliente:

- kms: DescribeKey
- km: CreateGrant

Por ejemplo:

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 concesión con los permisos kms: GenerateDataKey y kms:Decrypt.

Para más información, consulte Uso de políticas de claves en AWS KMS en la Guía del desarrollador de AWS Key Management Service .

Cómo utiliza Serverless las subvenciones en OpenSearch AWS KMS

OpenSearch Serverless requiere una subvención para poder utilizar una clave gestionada por el cliente.

Cuando creas una política de cifrado en tu cuenta con una clave nueva, OpenSearch Serverless crea una subvención en tu nombre enviando una CreateGrantsolicitud a. AWS KMS Las concesiones in AWS KMS se utilizan para permitir que OpenSearch Serverless acceda a una clave de KMS de la cuenta de un cliente.

OpenSearch Serverless requiere la concesión para utilizar la clave gestionada por el cliente en las siguientes operaciones internas:

- Envíe DescribeKeysolicitudes AWS KMS a para comprobar que el identificador de clave simétrico gestionado por el cliente proporcionado es válido.
- Envíe GenerateDataKeysolicitudes a la clave KMS para crear claves de datos con las que cifrar objetos.
- Envíe solicitudes de descifrado AWS KMS a para descifrar las claves de datos cifradas para que puedan usarse para cifrar sus datos.

Puede revocar el acceso a la concesión o eliminar el acceso del servicio a la clave administrada por el cliente en cualquier momento. Si lo hace, OpenSearch Serverless no podrá acceder a ninguno de los datos cifrados por la clave gestionada por el cliente, lo que afectará a todas las operaciones que dependen de esos datos y provocará `AccessDeniedException` errores y fallas en los flujos de trabajo asíncronos.

OpenSearch Serverless retira las concesiones en un flujo de trabajo asíncrono cuando una clave determinada administrada por el cliente no está asociada a ninguna política o colección de seguridad.

Creación de políticas de cifrado (consola)

En una política de cifrado, se especifica una clave KMS y una serie de patrones de colección a los que se aplicará la política. A cualquier colección nueva que coincida con uno de los patrones definidos en la política se le asignará la clave KMS correspondiente al crear la colección. Recomendamos que cree las políticas de cifrado antes de empezar a crear las colecciones.

Para crear una política de cifrado sin servidor OpenSearch

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, amplíe Sin servidor y seleccione Políticas de cifrado.
3. Seleccione Crear política de cifrado.
4. Escriba un nombre y una descripción para la política.
5. En Recursos, ingrese uno o más patrones de recursos para esta política de cifrado. Todas las colecciones recién creadas en la región y Cuenta de AWS actuales que coincidan con uno de los patrones se asignan de forma automática a esta política. Por ejemplo, si introduce `ApplicationLogs` (sin comodín) y luego crea una colección con ese nombre, la política y la clave KMS correspondientes se asignarán a esa colección.

   También puede proporcionar un prefijo como `Logs*`, que asigna la política a cualquier colección nueva cuyo nombre comience por `Logs`. Mediante el uso de comodines, puede administrar la configuración de cifrado de varias colecciones a escala.
6. En Cifrado, seleccione la clave KMS que se debe utilizar.
7. Seleccione Crear.

Siguiente paso: crear colecciones

Después de configurar una o más políticas de cifrado, puede empezar a crear colecciones que coincidan con las reglas definidas en esas políticas. Para obtener instrucciones, consulte the section called “Creación de colecciones”.

En el paso de cifrado de la creación de la colección, OpenSearch Serverless le informa de que el nombre que ha introducido coincide con el patrón definido en una política de cifrado y asigna

automáticamente la clave KMS correspondiente a la colección. Si el patrón de recursos contiene un comodín (*), puede anular la coincidencia y seleccionar su propia clave.

Creación de políticas de cifrado (AWS CLI)

Para crear una política de cifrado mediante las operaciones de la API OpenSearch sin servidor, debe especificar los patrones de recursos y una clave de cifrado en formato JSON. La CreateSecurityPolicysolicitud acepta tanto políticas en línea como archivos.json.

Las políticas de cifrado tienen el siguiente formato. El archivo de muestra `my-policy.json` coincidirá con cualquier colección futura con el nombre `autopartsinventory`, así como con cualquier colección cuyo nombre comience 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"
}
```

Para usar una clave propiedad del servicio, configure la `AWSOwnedKey` como `true`:

```
{
   "Rules":[
      {
         "ResourceType":"collection",
         "Resource":[
            "collection/autopartsinventory",
            "collection/sales*"
         ]
      }
   ],
   "AWSOwnedKey":true
}
```

La siguiente solicitud crea la política de cifrado:

```
aws opensearchserverless create-security-policy \
    --name sales-inventory \
    --type encryption \
    --policy file://my-policy.json
```

A continuación, utilice la operación de la CreateCollectionAPI para crear una o más colecciones que coincidan con uno de los patrones de recursos.

Visualización de políticas de cifrado

Antes de crear una colección, puede que desee obtener una vista previa de las políticas de cifrado existentes en su cuenta para ver cuál tiene un patrón de recursos que coincide con el nombre de su colección. La siguiente ListSecurityPoliciessolicitud muestra todas las políticas de cifrado de tu cuenta:

```
aws opensearchserverless list-security-policies --type encryption
```

La solicitud devuelve información sobre todas las políticas de cifrado configuradas. Utilice el contenido del elemento `policy` para ver las reglas de patrón que se definen en la política:

```
{
   "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"
      }
   ]
}
```

Para ver información detallada sobre una política específica, incluida la clave KMS, usa el GetSecurityPolicycomando.

Actualización de políticas de cifrado

Si actualiza la clave KMS en una política de cifrado, el cambio solo se aplica a las colecciones recién creadas que coincidan con el nombre o patrón configurado. No afecta a las colecciones existentes que ya tienen claves KMS asignadas.

Lo mismo sucede con las reglas de coincidencia de las políticas. Si agrega, modifica o elimina una regla, el cambio solo se aplica a las colecciones recién creadas. Las colecciones existentes no pierden la clave KMS asignada si modifica las reglas de una política para que ya no coincida con el nombre de una colección.

Para actualizar una política de cifrado en la consola OpenSearch sin servidor, elija Políticas de cifrado, seleccione la política que desee modificar y elija Editar. Realice los cambios y elija Guardar.

Para actualizar una política de cifrado mediante la API OpenSearch sin servidor, utilice la UpdateSecurityPolicyoperación. La siguiente solicitud actualiza una política de cifrado con un nuevo documento JSON de política:

```
aws opensearchserverless update-security-policy \
    --name sales-inventory \
    --type encryption \
    --policy-version 2 \
    --policy file://my-new-policy.json
```

Eliminación de políticas de cifrado

Cuando se elimina una política de cifrado, las colecciones que estén utilizando en ese momento la clave KMS definida en la política no se ven afectadas. Para eliminar una política en la consola OpenSearch sin servidor, selecciónela y elija Eliminar.

También puede utilizar la DeleteSecurityPolicyoperación:

```
aws opensearchserverless delete-security-policy --name my-policy --type encryption
```

Cifrado en tránsito

En OpenSearch Serverless, todas las rutas de una colección se cifran en tránsito mediante Transport Layer Security 1.2 (TLS) y un cifrado AES-256 estándar del sector. El acceso a todos los paneles APIs y a los paneles de Opensearch también se realiza a través de TLS 1.2. TLS es un conjunto

de protocolos criptográficos estándares del sector que se utilizan para cifrar la información que se intercambia a través de la red.

Acceso a la red para Amazon OpenSearch Serverless

La configuración de red de una colección de Amazon OpenSearch Serverless determina si se puede acceder a la colección a través de Internet desde redes públicas o si se debe acceder a ella de forma privada.

El acceso privado puede aplicarse a uno de los siguientes lugares o ambos:

- OpenSearch Terminales de VPC gestionados sin servidor
- Compatible Servicios de AWS , como Amazon Bedrock

Puede configurar el acceso a la red por separado para el OpenSearchpunto final de una colección y el punto final de OpenSearch Dashboards correspondiente.

El acceso a la red es el mecanismo de aislamiento que permite el acceso desde diferentes redes de origen. Por ejemplo, si el punto final de los OpenSearch paneles de una colección es de acceso público, pero el punto final de la OpenSearch API no, el usuario solo podrá acceder a los datos de la recopilación a través de los paneles cuando se conecte desde una red pública. Si intenta llamarlos OpenSearch APIs directamente desde una red pública, se bloqueará. La configuración de red se puede utilizar para estas permutaciones de origen a tipo de recurso. Amazon OpenSearch Serverless es compatible con ambos tipos IPv4 de IPv6 conectividad.

Temas

- Políticas de red
- Consideraciones
- Permisos necesarios para configurar políticas de red
- Prioridad política
- Creación de políticas de (consola)
- Creación de políticas de (AWS CLI)
- Visualización de políticas de red
- Actualización de las políticas de red
- Eliminar las políticas de red

Políticas de red

Las políticas de red le permiten administrar muchas colecciones a gran escala al asignar de forma automática la configuración de acceso a la red a las colecciones que cumplen con las reglas definidas en la política.

En una política de red, se especifican una serie de reglas. Estas reglas definen los permisos de acceso a los puntos de enlace de recopilación y a los puntos de enlace de OpenSearch Dashboards. Cada regla consta de un tipo de acceso (público o privado) y un tipo de recurso (punto final de colección o OpenSearch panel de control). Para cada tipo de recurso (`collection` y `dashboard`), especifique una serie de reglas que definen a qué colecciones se aplicará la política.

En este ejemplo de política, la primera regla especifica el acceso al punto de conexión de VPC tanto en el punto de conexión de la colección como al punto de conexión de Dashboards para todas las colecciones que comiencen por el término `marketing*`. También especifica el acceso a Amazon Bedrock.

> **Note**
>
> El acceso privado a Servicios de AWS Amazon Bedrock solo se aplica al punto final de la colección, no al OpenSearch punto final de OpenSearch Dashboards. Incluso si es así`dashboard`, Servicios de AWS no `ResourceType` se le puede conceder acceso a OpenSearch los paneles de control.

La segunda regla especifica el acceso público a la colección `finance`, pero solo para el punto de conexión de la colección (sin acceso a los paneles).

```
[
   {
      "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
   }
]
```

Esta política solo proporciona acceso público a los OpenSearch paneles de control para las colecciones que comiencen por «finanzas». Cualquier intento de acceder directamente a la OpenSearch API fallará.

```
[
  {
    "Description": "Dashboards access",
    "Rules": [
      {
        "ResourceType": "dashboard",
        "Resource": [
          "collection/finance*"
        ]
      }
    ],
    "AllowFromPublic": true
```

```
  }
]
```

Las políticas de red pueden aplicarse a las colecciones existentes, así como a las colecciones futuras. Por ejemplo, puede crear una colección y, a continuación, crear una política de red con una regla que coincida con el nombre de la colección. No es necesario crear políticas de red para crear colecciones.

Consideraciones

Al configurar el acceso a la red para sus colecciones, tenga en cuenta lo siguiente:

- Si planea configurar el acceso al punto de enlace de VPC para una colección, primero debe crear al OpenSearch menos un punto de enlace de VPC administrado sin servidor.
- El acceso privado Servicios de AWS solo se aplica al punto final de la colección, no al OpenSearch punto final de Dashboards. OpenSearch Incluso si es así`dashboard`, Servicios de AWS no `ResourceType` se le puede conceder acceso a los OpenSearch paneles de control.
- Si se puede acceder a una colección desde redes públicas, también se puede acceder a ella desde todos los puntos finales de VPC OpenSearch gestionados por Serverless y todos ellos. Servicios de AWS
- Se pueden aplicar varias políticas de red a una sola colección. Para obtener más información, consulte the section called “Prioridad política”.

Permisos necesarios para configurar políticas de red

El acceso a la red para OpenSearch Serverless utiliza los siguientes permisos (IAM). AWS Identity and Access Management Puede especificar condiciones de IAM para restringir a los usuarios a las políticas de red asociadas con colecciones específicas.

- `aoss:CreateSecurityPolicy`: cree una política de acceso a la red.
- `aoss:ListSecurityPolicies`: enumere todas las políticas de red de la cuenta actual.
- `aoss:GetSecurityPolicy`: vea una especificación de la política de acceso a la red.
- `aoss:UpdateSecurityPolicy`: modifique una política de acceso a la red determinada y cambiar el ID de VPC o la designación de acceso público.
- `aoss:DeleteSecurityPolicy`: elimine una política de acceso a la red (después de separarla de todas las colecciones).

La siguiente política de acceso basada en identidades permite al usuario ver todas las políticas de red y actualizarlas según el patrón de recursos`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**
>
> Además, OpenSearch Serverless requiere los `aoss:DashboardsAccessAll` permisos `aoss:APIAccessAll` y los permisos para los recursos de recopilación. Para obtener más información, consulte the section called “Uso de operaciones OpenSearch de API”.

Prioridad política

Puede haber situaciones en las que las reglas de la política de red se superpongan, dentro de las políticas o entre ellas. Cuando esto ocurre, una regla que especifica el acceso público anula la regla que especifica el acceso privado para cualquier colección que sea común a ambas reglas.

Por ejemplo, en la siguiente política, ambas reglas asignan acceso a la red de la colección `finance`, pero una regla especifica el acceso a la VPC y la otra especifica el acceso público. En esta situación, el acceso público anula el acceso a la VPC únicamente para la recaudación de fondos (porque existe en ambas reglas), por lo que se podrá acceder a la colección de fondos desde las redes públicas. La colección de ventas tendrá acceso a VPC desde el punto de conexión especificado.

```
[
   {
      "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
   }
]
```

Si se aplican varios puntos de conexión de VPC de diferentes reglas a una colección, las reglas son aditivas y se podrá acceder a la colección desde todos los puntos de conexión especificados. Si lo configura `AllowFromPublic` `true` pero también proporciona uno `SourceVPCEs` o más`SourceServices`, OpenSearch Serverless ignora los puntos finales de la VPC y los identificadores de servicio, y las colecciones asociadas tendrán acceso público.

Creación de políticas de (consola)

Las políticas de red pueden aplicarse a las colecciones existentes, así como a las colecciones futuras. Se recomienda crear políticas de red antes de empezar a crear colecciones.

Para crear una política de red sin servidor OpenSearch

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, expanda Sin servidor y seleccione Políticas de red.
3. Seleccione Crear políticas de red.
4. Escriba un nombre y una descripción para la política.
5. Proporcione una o más reglas. Estas reglas definen los permisos de acceso para sus colecciones OpenSearch sin servidor y sus puntos finales de OpenSearch Dashboards.

   Cada regla contiene los siguientes elementos:

Element	Description (Descripción)
Nombre de la regla	Nombre que describa el contenido de la regla. Por ejemplo, “Acceso a VPC para el equipo de marketing”.
Tipo de acceso	Seleccione el acceso público o privado. A continuación, seleccione una de las siguiente s opciones o ambas: • Puntos de enlace de VPC para el acceso: especifique uno o más puntos de enlace de VPC administrados sin servidor (puntos de enlace de VPC administrados)Open Search .

Element	Description (Descripción)
	• Servicio de AWS acceso privado: seleccione uno o Servicios de AWS más compatibles.
Tipo de recurso	Seleccione si desea proporcionar acceso a los OpenSearch puntos finales (lo que permite realizar llamadas a la OpenSearch API), a los OpenSearch paneles (lo que permite el acceso a las visualizaciones y a la interfaz de usuario de los OpenSearch complementos) o a ambos. **Note** Servicio de AWS El acceso privado solo se aplica al punto final de la colección, no al OpenSearch punto final de Dashboards. OpenSearch Incluso si selecciona los OpenSearch paneles de control, solo Servicios de AWS se le puede conceder el acceso al punto final.

Para cada tipo de recurso que seleccione, puede elegir las colecciones existentes a las que aplicar la configuración de políticas y and/or crear uno o más patrones de recursos. Los patrones de recursos constan de un prefijo y un comodín (*) y definen a qué colecciones se aplicará la configuración de la política.

Por ejemplo, si incluye un patrón denominado `Marketing*`, a cualquier colección nueva o existente cuyo nombre comience por “Marketing” se le aplicará automáticamente la configuración de red de esta política. Un único comodín (`*`) aplica la política a todas las colecciones actuales y future.

Además, puede especificar el nombre de una colección futura sin caracteres comodín, como`Finance`. OpenSearch Serverless aplicará la configuración de la política a cualquier colección recién creada con ese nombre exacto.

6. Cuando esté satisfecho, seleccione Crear.

Creación de políticas de (AWS CLI)

Para crear una política de red mediante las operaciones de la API OpenSearch sin servidor, debe especificar las reglas en formato JSON. La CreateSecurityPolicysolicitud acepta políticas en línea y archivos.json. Todas las colecciones y patrones deben tomar la forma `collection/<collection name|pattern>`.

> **Note**
>
> El tipo de recurso `dashboards` solo permite el acceso a los OpenSearch paneles, pero para que los OpenSearch paneles funcionen, también debes permitir el acceso a la colección desde las mismas fuentes. Consulte la segunda política a continuación para ver un ejemplo.

Para especificar el acceso privado, incluya uno o los siguientes elementos:

- `SourceVPCEs`— Especifique uno o más puntos finales de VPC OpenSearch gestionados sin servidor.
- `SourceServices`— Especifique el identificador de uno o más de los compatibles. Servicios de AWS Actualmente, se admiten los siguientes identificadores de servicio:
  - `bedrock.amazonaws.com`: Amazon Bedrock

El siguiente ejemplo de política de red proporciona acceso privado, a una punto de conexión de VPC y Amazon Bedrock, a los puntos de conexión de las colecciones solo para las colecciones que comienzan con el prefijo `log*`. Los usuarios autenticados no pueden iniciar sesión en los OpenSearch paneles; solo pueden acceder al punto final de la colección mediante programación.

```
[
   {
      "Description":"Private access for log collections",
      "Rules":[
```

```
         {
            "ResourceType":"collection",
            "Resource":[
               "collection/log*"
            ]
         }
      ],
      "AllowFromPublic":false,
      "SourceVPCEs":[
         "vpce-050f79086ee71ac05"
      ],
      "SourceServices":[
         "bedrock.amazonaws.com"
      ],
   }
]
```

La siguiente política proporciona acceso público al OpenSearch punto final y a los OpenSearch paneles de control de una sola colección denominada. `finance` Si la colección no existe, la configuración de red se aplicará a la colección siempre y cuando se cree.

```
[
   {
      "Description":"Public access for finance collection",
      "Rules":[
         {
            "ResourceType":"dashboard",
            "Resource":[
               "collection/finance"
            ]
         },
         {
            "ResourceType":"collection",
            "Resource":[
               "collection/finance"
            ]
         }
      ],
      "AllowFromPublic":true
   }
]
```

La siguiente solicitud crea la política de red anterior:

```
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}]"
```

Para proporcionar la política en un archivo JSON, utilice el formato`--policy file://my-policy.json`

Visualización de políticas de red

Antes de crear una colección, puede que desee obtener una vista previa de las políticas de red existentes en su cuenta para ver cuáles tienen un patrón de recursos que coincide con el nombre de su colección. La siguiente ListSecurityPoliciessolicitud muestra todas las políticas de red de su cuenta:

```
aws opensearchserverless list-security-policies --type network
```

La solicitud devuelve información sobre todas las políticas de red configuradas. Para ver las reglas de patrón definidas en una política específica, busque la información de la política en el contenido del elemento `securityPolicySummaries` en la respuesta. Tenga en cuenta el `name` `type` final de esta política y utilice estas propiedades en una GetSecurityPolicysolicitud para recibir una respuesta con los siguientes detalles de la política:

```
{
    "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
        }
    ]
}
```

Para ver información detallada sobre una política específica, utilice el GetSecurityPolicycomando.

Actualización de las políticas de red

Al modificar los puntos de conexión de VPC o la designación de acceso público de una red, todas las colecciones asociadas se ven afectadas. Para actualizar una política de red en la consola OpenSearch sin servidor, expanda Políticas de red, seleccione la política que desee modificar y elija Editar. Realice los cambios y elija Guardar.

Para actualizar una política de red mediante la API OpenSearch sin servidor, utilice el UpdateSecurityPolicycomando. Debe incluir un control de versiones de política en la solicitud. Puede recuperar el control de versiones de la política mediante los comandos `ListSecurityPolicies` o `GetSecurityPolicy`. Incluir la versión más reciente de la política garantiza que no se anule inadvertidamente un cambio realizado por otra persona.

La siguiente solicitud actualiza una política de red con un nuevo documento JSON de política:

```
aws opensearchserverless update-security-policy \
    --name sales-inventory \
    --type network \
    --policy-version MTY2MzY5MTY1MDA3Ml8x \
    --policy file://my-new-policy.json
```

Eliminar las políticas de red

Para poder eliminar una política de red, debe desconectarla de todas las colecciones. Para eliminar una política en la consola OpenSearch sin servidor, selecciónela y elija Eliminar.

También puede usar el DeleteSecurityPolicycomando:

```
aws opensearchserverless delete-security-policy --name my-policy --type network
```

Conformidad con FIPS en Amazon Serverless OpenSearch

Amazon OpenSearch Serverless es compatible con los estándares federales de procesamiento de información (FIPS) 140-2, que son estándares gubernamentales de EE. UU. y Canadá que especifican los requisitos de seguridad para los módulos criptográficos que protegen la información confidencial. Cuando se conecta a puntos de conexión compatibles con FIPS con OpenSearch

Serverless, las operaciones criptográficas se realizan mediante bibliotecas criptográficas validadas por FIPS.

OpenSearch Los puntos finales FIPS sin servidor están disponibles en los lugares en los que se admite FIPS. Regiones de AWS Estos puntos de conexión utilizan TLS 1.2 o una versión posterior y algoritmos criptográficos validados FIPS para todas las comunicaciones. Para obtener más información, consulte Cumplimiento FIPS en la Guía del usuario de AWS Verified Access.

Temas

- Uso de puntos finales FIPS con Serverless OpenSearch
- Utilice los puntos finales de FIPS con AWS SDKs
- Configurar grupos de seguridad para puntos de conexión de VPC
- Uso del punto de conexión de VPC compatible con FIPS
- Verificar el cumplimiento de FIPS
- Solución de problemas de conectividad de puntos de conexión de FIPS en zonas alojadas privadas

Uso de puntos finales FIPS con Serverless OpenSearch

Regiones de AWS Cuando se admite FIPS, se puede acceder a las colecciones OpenSearch sin servidor a través de puntos de conexión estándar y compatibles con FIPS. Para obtener más información, consulte Cumplimiento FIPS en la Guía del usuario de AWS Verified Access.

En los siguientes ejemplos, sustituya *`collection_id`* y por su identificador de colección y el *`Región de AWS`* suyo. Región de AWS

- Puntos de conexión estándar: **`https://`*`collection_id`*`.`*`Región de AWS`*`.aoss.amazonaws.com`**.
- Punto de conexión compatible con FIPS: **`https://`*`collection_id`*`.`*`Región de AWS`*`.aoss-fips.amazonaws.com`**.

Del mismo modo, se puede acceder a los OpenSearch paneles de control a través de puntos de conexión estándar y compatibles con el FIPS:

- Punto de conexión estándar de Dashboards: **`https://`*`collection_id`*`.`*`Región de AWS`*`.aoss.amazonaws.com/_dashboards`**.

- Punto de conexión de Dashboards compatible con FIPS: **`https://collection_id.Región de AWS.aoss-fips.amazonaws.com/_dashboards`**.

> **Note**
>
> En las regiones habilitadas para FIPS, tanto los puntos de conexión estándar como los compatibles con FIPS proporcionan criptografía compatible con FIPS. Los puntos de conexión específicos de FIPS lo ayudan a cumplir con los requisitos de cumplimiento que exigen específicamente el uso de puntos de conexión con FIPS en el nombre.

Utilice los puntos finales de FIPS con AWS SDKs

Cuando lo utilice AWS SDKs, puede especificar el punto final de FIPS al crear el cliente. En el siguiente ejemplo, sustituya *`collection_id`* y *`Región de AWS`* por su identificador de colección y el suyo Región de AWS.

```
# Python SDK example
from opensearchpy import OpenSearch, RequestsHttpConnection, AWSV4SignerAuth
import boto3
host = '"https://collection_id.Región de 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
)
```

Configurar grupos de seguridad para puntos de conexión de VPC

Para garantizar una comunicación adecuada con su punto final de Amazon VPC (VPC) compatible con FIPS, cree o modifique un grupo de seguridad para permitir el tráfico HTTPS entrante (puerto TCP 443) desde los recursos de su VPC que necesitan acceder a Serverless. OpenSearch Luego,

asocie este grupo de seguridad con su punto de conexión de VPC durante la creación o modificando el punto de conexión después de la creación. Para obtener más información, consulte Crear un grupo de seguridad en la Guía del usuario de Amazon VPC.

Uso del punto de conexión de VPC compatible con FIPS

Después de crear el punto final de la VPC compatible con FIPS, puede usarlo para OpenSearch acceder a Serverless desde los recursos de su VPC. Para usar el punto de conexión en operaciones de API, configure su SDK para usar el punto de conexión regional FIPS según se describe en la sección Uso de puntos finales FIPS con Serverless OpenSearch. Para acceder a los OpenSearch paneles, utilice la URL de paneles específica de la colección, que se enrutará automáticamente por el punto final de la VPC compatible con FIPS cuando se acceda a ellos desde su VPC. Para obtener más información, consulte Uso de OpenSearch cuadros de mando con Amazon Service OpenSearch.

Verificar el cumplimiento de FIPS

Para comprobar que tus conexiones a Serverless utilizan criptografía compatible con FIPS, OpenSearch úsala para supervisar las llamadas de API realizadas a Serverless. AWS CloudTrail OpenSearch Comprueba que el `eventSource` campo de los registros muestre las llamadas a la API CloudTrail . `aoss-fips.amazonaws.com`

Para acceder a los OpenSearch paneles de control, puede utilizar las herramientas de desarrollo del navegador para inspeccionar los detalles de la conexión TLS y comprobar que se utilizan conjuntos de cifrado compatibles con el FIPS.

Solución de problemas de conectividad de puntos de conexión de FIPS en zonas alojadas privadas

Los puntos de enlace FIPS funcionan con las colecciones de Amazon OpenSearch Serverless que tienen acceso público. Para las nuevas colecciones de VPC que utilizan puntos de conexión de VPC recién creados, los puntos de conexión de FIPS funcionan según lo esperado. Para otras colecciones de VPC, es posible que deba realizar una configuración manual para garantizar que los puntos de conexión de FIPS funcionen correctamente.

Para configurar zonas alojadas privadas de FIPS en Amazon Route 53

1. Abra la consola de Route 53 en. https://console.aws.amazon.com/route53/
2. Revise sus zonas alojadas:

a. Localice las zonas alojadas en las que se encuentran Regiones de AWS sus colecciones.

b. Compruebe los patrones de nomenclatura de las zonas alojadas:

- Formato que no es de FIPS: *region*`.aoss.amazonaws.com`.
- Formato que es de FIPS: *region*`.aoss-fips.amazonaws.com`.

c. Confirme que el Tipo de todas sus zonas alojadas esté configurado como Zona alojada privada.

3. Si falta la zona alojada privada de FIPS:

a. Seleccione la zona alojada privada no FIPS correspondiente.

b. Copia la VPCs información asociada. Por ejemplo: `vpc-1234567890abcdef0 | us-east-2`.

c. Busque el registro de dominio con comodín. Por ejemplo: `*.us-east-2.aoss.amazonaws.com`.

d. Copie la información de Valor/Dirigir el tráfico a. Por ejemplo: `uoc1c1qsw7poexampleewjeno1pte3rw.3ym756xh7yj.aoss.searchservices.aws`.

4. Cree la zona alojada privada de FIPS:

a. Cree una nueva zona alojada privada con el formato de FIPS. Por ejemplo: `us-east-2.aoss-fips.amazonaws.com`.

b. VPCsEn Asociado, introduzca la información de la VPC que copió de la zona alojada privada que no es de FIPS.

5. Agregue una nueva regla con las siguientes opciones de configuración:

a. Nombre del registro: *

b. Tipo de registro: CNAME

c. Valor: ingrese la información de Valor/Dirigir el tráfico a que copió anteriormente.

Problemas comunes

Si tiene problemas de conectividad con los puntos de conexión de VPC compatibles con FIPS, consulte la siguiente información para ayudar a resolverlos.

- Errores de resolución de DNS: no puede resolver el nombre de dominio del punto de conexión de FIPS en su VPC

- Tiempos de espera de conexión: se agota el tiempo de espera de sus solicitudes al punto de conexión de FIPS
- Errores de acceso denegado: la autenticación o la autorización fallan cuando se utilizan puntos de conexión de FIPS
- Faltan registros de zonas alojadas privadas para colecciones exclusivas de VPC

Para solucionar problemas de conectividad de puntos de conexión de FIPS

1. Compruebe la configuración de su zona alojada privada:

   a. Confirme que existe una zona alojada privada para el dominio del punto de conexión de FIPS (`*.region.aoss-fips.amazonaws.com`).

   b. Compruebe que la zona alojada privada esté asociada a la VPC correcta.

      Para obtener más información, consulte Zonas alojadas privadas en la Guía para desarrolladores de Amazon Route 53 y Administración de nombres de DNS en la Guía de AWS PrivateLink .

2. Pruebe de la resolución de DNS:

   a. Conéctese a una instancia de EC2 en su VPC.

   b. Use el siguiente comando:

      ```
      nslookup collection-id.region.aoss-fips.amazonaws.com
      ```

   c. Compruebe que la respuesta incluya la dirección IP privada del punto de conexión de VPC.

      Para obtener más información, consulte Políticas de puntos de conexión y Atributos de DNS en la Guía del usuario de Amazon VPC.

3. Compruebe la configuración del grupo de seguridad:

   a. Compruebe que el grupo de seguridad adjunto al punto de conexión de VPC permita el tráfico HTTPS (puerto 443) desde sus recursos.

   b. Confirme que los grupos de seguridad de sus recursos permitan el tráfico saliente al punto de conexión de VPC.

   Para obtener más información, consulte Políticas de puntos de conexión en la Guía de AWS PrivateLink y Grupos de seguridad en la Guía del usuario de Amazon VPC.

4. Revise la configuración de la ACL de red:

   - Compruebe que la red ACLs permita el tráfico entre sus recursos y el punto final de la VPC.

     Para obtener más información, consulte Red ACLs en la Guía del usuario de Amazon VPC.

5. Revise su política de puntos de conexión:

   - Compruebe que la política de puntos finales de la VPC permita las acciones necesarias en sus recursos OpenSearch sin servidor.

     Para obtener más información, consulte Permisos de puntos de conexión de VPC necesarios y Políticas de puntos de conexión en la Guía de AWS PrivateLink .

> **Tip**
>
> Si utilizas resolutores de DNS personalizados en tu VPC, configúralos para que reenvíen las `*.amazonaws.com` solicitudes de dominios a AWS los servidores.

Control de acceso a datos para Amazon OpenSearch Serverless

Con el control de acceso a los datos de Amazon OpenSearch Serverless, puede permitir que los usuarios accedan a colecciones e índices, independientemente de su mecanismo de acceso o fuente de red. Puede proporcionar acceso a roles de IAM e identidades SAML.

Los permisos de acceso se administran mediante las políticas de acceso a los datos, que se aplican a las colecciones y a los recursos de índice. Las políticas de acceso a datos le ayudan a administrar las colecciones a escala mediante la asignación automática de permisos de acceso a las colecciones e índices que coinciden con un patrón específico. Se pueden aplicar varias políticas de acceso a los datos a un solo recurso. Tenga en cuenta que debe tener una política de acceso a los datos para su colección para poder acceder a la URL de su OpenSearch panel de control.

Temas

- Políticas de acceso a datos frente a políticas de IAM
- Permisos de IAM necesarios para configurar las políticas de acceso a los datos
- Sintaxis de la política
- Permisos de política compatibles

- Ejemplos de conjuntos de datos en paneles OpenSearch
- Creación de políticas de acceso a los datos (consola)
- Creación de políticas de acceso a los datos (AWS CLI)
- Visualización de políticas de acceso a los datos
- Actualización de las políticas de acceso a los datos
- Eliminación de políticas de acceso a los datos
- Acceso a los datos entre cuentas

Políticas de acceso a datos frente a políticas de IAM

Las políticas de acceso a los datos son lógicamente independientes de las políticas AWS Identity and Access Management (de IAM). Los permisos de IAM controlan el acceso a las operaciones de la API sin servidor, como `CreateCollection` y `ListAccessPolicies`. Las políticas de acceso a los datos controlan el acceso a las OpenSearch operaciones compatibles con OpenSearch Serverless, como o. `PUT <index>` `GET _cat/indices`

Los permisos de IAM que controlan el acceso a las operaciones de la API de la política de acceso a los datos, como `aoss:CreateAccessPolicy` y `aoss:GetAccessPolicy` (que se describen en la siguiente sección), no afectan al permiso especificado en una política de acceso a los datos.

Por ejemplo, supongamos que una política de IAM niega a un usuario crear políticas de acceso a los datos para `collection-a`, pero le permite crear políticas de acceso a datos para todas las colecciones (`*`):

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Deny",
            "Action": [
                "aoss:CreateAccessPolicy"
            ],
            "Resource": "*",
            "Condition": {
                "StringLike": {
```

```
                    "aoss:collection": "collection-a"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "aoss:CreateAccessPolicy"
            ],
            "Resource": "*"
        }
    ]
}
```

Si el usuario crea una política de acceso a los datos que permite ciertos permisos a todas las colecciones (`collection/*` o `index/*/*`), la política se aplicará a todas las colecciones, incluida la colección A.

> **Important**
>
> Contar con permisos en el marco de una política de acceso a los datos no es suficiente para acceder a los datos de su colección OpenSearch Serverless. A una entidad principal también se le debe conceder acceso a los permisos de IAM `aoss:APIAccessAll` y `aoss:DashboardsAccessAll`. Ambos permisos otorgan acceso total a los recursos de la recopilación, mientras que el permiso de paneles también proporciona acceso a los OpenSearch paneles. Si una entidad principal no tiene estos dos permisos de IAM, recibirá errores 403 cuando intente enviar solicitudes a la colección. Para obtener más información, consulte the section called “Uso de operaciones OpenSearch de API”.

Permisos de IAM necesarios para configurar las políticas de acceso a los datos

El control de acceso a los datos para OpenSearch Serverless utiliza los siguientes permisos de IAM. Puede especificar las condiciones de IAM para restringir a los usuarios a nombres de políticas de acceso específicos.

- `aoss:CreateAccessPolicy`: cree una política de acceso.
- `aoss:ListAccessPolicies`: enumere todas las políticas de acceso.
- `aoss:GetAccessPolicy`: consulte los detalles sobre una política de acceso específica.

- `aoss:UpdateAccessPolicy`: modifique una política de acceso.
- `aoss:DeleteAccessPolicy`: elimine una política de acceso.

La siguiente política de acceso basada en la identidad permite al usuario ver todas las políticas de acceso y actualizar las políticas que contienen el patrón de recursos `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**
>
> Además, OpenSearch Serverless requiere los `aoss:DashboardsAccessAll` permisos `aoss:APIAccessAll` y los permisos para los recursos de recopilación. Para obtener más información, consulte the section called “Uso de operaciones OpenSearch de API”.

Sintaxis de la política

Una política de acceso a datos incluye un conjunto de reglas, cada una con los siguientes elementos:

Element	Description (Descripción)
`ResourceType`	El tipo de recurso (colección o índice) al que se le aplican los permisos. Los permisos de alias y plantillas se encuentran en el nivel de la colección , mientras que los permisos para crear, modificar y buscar datos están en el nivel del índice. Para más información, consulte Permisos de políticas compatibles.
`Resource`	Una lista de and/or patrones de nombres de recursos. Los patrones son prefijos seguidos de un comodín (*), que permiten que los permisos asociados se apliquen a varios recursos. • Las colecciones adoptan el formato `collection/` *`<name|pattern>`* . • Los índices toman el formato `index/`*`<collection-name|pattern>`* */`*`<index-name|pattern/>`* .
`Permission`	Una lista de permisos a otorgar para los recursos especificados. Para obtener una lista completa de las operaciones de la API y de los permisos que permiten, consulte the section called “Operaciones y permisos de OpenSearch API compatibles”.
`Principal`	Una lista de una o más entidades principales a las que conceder acceso. Los principales pueden ser identidades de rol de IAM ARNs o de SAML. Estas entidades principales deben estar dentro de la Cuenta de AWS actual. Las políticas de acceso a los datos no admiten directamente el acceso entre cuentas, pero puedes incluir en tu política una función que un usuario de otra cuenta Cuenta de AWS pueda asumir en la cuenta propietaria de la colección

Element	Description (Descripción)
	. Para obtener más información, consulte the section called “Acceso a los datos entre cuentas”.

La siguiente política de ejemplo otorga permisos de alias y plantillas a la colección llamada `autopartsinventory`, así como a cualquier colección que comience por el prefijo `sales*`. También otorga permisos de lectura y escritura a todos los índices dentro de la colección `autopartsinventory` y a todos los índices de la colección `salesorders` que comiencen por el prefijo `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"
        ]
    }
]
```

No se puede denegar el acceso de forma explícita dentro de una política. Por lo tanto, todos los permisos de la política son aditivos. Por ejemplo, si una política concede el permiso `aoss:ReadDocument` a un usuario y otra `aoss:WriteDocument`, el usuario tendrá ambos permisos. Si una tercera política concede al mismo usuario el permiso `aoss:*`, el usuario puede realizar todas las acciones del índice asociado; los permisos más restrictivos no anulan los menos restrictivos.

Permisos de política compatibles

Las políticas de acceso a los datos admiten los siguientes permisos. Para ver las operaciones OpenSearch de API que permite cada permiso, consulte. the section called “Operaciones y permisos de OpenSearch API compatibles”

Permisos de colección

- `aoss:CreateCollectionItems`
- `aoss:DeleteCollectionItems`
- `aoss:UpdateCollectionItems`
- `aoss:DescribeCollectionItems`
- `aoss:*`

Permisos de índice

- `aoss:ReadDocument`
- `aoss:WriteDocument`
- `aoss:CreateIndex`
- `aoss:DeleteIndex`
- `aoss:UpdateIndex`
- `aoss:DescribeIndex`
- `aoss:*`

Ejemplos de conjuntos de datos en paneles OpenSearch

OpenSearch Los paneles proporcionan conjuntos de datos de muestra que incluyen visualizaciones, paneles y otras herramientas para ayudarlo a explorar los paneles antes de agregar sus propios datos. Para crear índices a partir de estos datos de ejemplo, necesita una política de acceso a los datos que otorgue permisos al conjunto de datos con el que desee trabajar. La siguiente política usa un comodín (*) para proporcionar permisos a los tres conjuntos de datos de ejemplo.

```
[
  {
    "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>"
    ]
  }
]
```

Creación de políticas de acceso a los datos (consola)

Puede crear una política de acceso a los datos con el editor visual o en el formato JSON. A cualquier colección nueva que coincida con uno de los patrones definidos en la política se le asignarán los permisos correspondientes al crear la colección.

Para crear una política de acceso a los datos sin servidor OpenSearch

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, expanda sin servidor y, en Seguridad, elija Políticas de acceso a datos.
3. Seleccione Crear política de acceso.

4. Escriba un nombre y una descripción para la política.
5. Proporcione un nombre para la primera regla de la política. Por ejemplo, “Acceso a la colección de registros”.
6. Seleccione Agregar entidades principales, y luego uno o varios roles de IAM o usuarios y grupos de SAML para proporcionarles acceso a los datos.

> **Note**
>
> Para seleccionar las entidades principales desde los menús desplegables, debe tener los permisos `iam:ListUsers` y `iam:ListRoles` (para las entidades principales de IAM) y el permiso `aoss:ListSecurityConfigs` (para las identidades de SAML).

7. Seleccione Otorgar y seleccione los permisos de alias, plantillas e índices que desea conceder a las entidades principales asociadas. Para obtener una lista completa de los permisos y el acceso que permiten, consulte the section called “Operaciones y permisos de OpenSearch API compatibles”.
8. (Opcional) Configure reglas adicionales para la política.
9. Seleccione Crear. Puede transcurrir alrededor de un minuto de retardo entre el momento en que se crea la política y el momento en el que se aplican los permisos. Si tarda más de 5 minutos, póngase en contacto con Soporte.

> **Important**
>
> Si la política solo incluye permisos de índices (y no permisos de colecciones), es posible que siga apareciendo un mensaje sobre las colecciones coincidentes que indique lo siguiente: `Collection cannot be accessed yet. Configure data access policies so that users can access the data within this collection`. Puede omitir esta advertencia. Las entidades principales autorizadas pueden seguir realizando las operaciones relacionadas con índices que tengan asignadas en la colección.

Creación de políticas de acceso a los datos (AWS CLI)

Para crear una política de acceso a los datos mediante la API OpenSearch sin servidor, utilice el `CreateAccessPolicy` comando. El comando acepta políticas insertadas y archivos .json. Las políticas insertadas deben codificarse como una cadena de escape de JSON.

La siguiente solicitud crea una política de acceso a los datos:

```
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\"]}]"
```

Para proporcionar la política dentro de un archivo .json, utilice el formato `--policy file://my-policy.json`.

Los principales incluidos en la política ahora pueden usar OpenSearch las operaciones a las que se les concedió acceso.

Visualización de políticas de acceso a los datos

Antes de crear una colección, puede que desee obtener una vista previa de las políticas de acceso a los datos existentes en la cuenta para ver cuáles tienen un patrón de recursos que coincide con el nombre de su colección. En la siguiente ListAccessPoliciessolicitud, se enumeran todas las políticas de acceso a los datos de tu cuenta:

```
aws opensearchserverless list-access-policies --type data
```

La solicitud devuelve información sobre todas las políticas de acceso a los datos configuradas. Para ver las reglas de patrón definidas en una política específica, busque la información de la política en el contenido del elemento `accessPolicySummaries` en la respuesta. Tenga en cuenta el `name` `type` final de esta política y utilice estas propiedades en una GetAccessPolicysolicitud para recibir una respuesta con los siguientes detalles de la política:

```
{
    "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
        }
    ]
}
```

Puede incluir filtros de recursos para limitar los resultados a políticas que contengan colecciones o índices específicos:

```
aws opensearchserverless list-access-policies --type data --resource
 "index/autopartsinventory/*"
```

Para ver los detalles de una política específica, utilice el GetAccessPolicycomando.

Actualización de las políticas de acceso a los datos

Al actualizar una política de acceso a los datos, todas las colecciones asociadas se ven afectadas. Para actualizar una política de acceso a datos en la consola OpenSearch sin servidor, elija Control de acceso a datos, seleccione la política que desee modificar y elija Editar. Realice los cambios y elija Guardar.

Para actualizar una política de acceso a datos mediante la API OpenSearch sin servidor, envía una `UpdateAccessPolicy` solicitud. Debe incluir una versión de la política, que puede recuperar mediante los comandos `ListAccessPolicies` or `GetAccessPolicy`. Incluir la versión más reciente de la política garantiza que no se anule inadvertidamente un cambio realizado por otra persona.

La siguiente UpdateAccessPolicysolicitud actualiza una política de acceso a datos con un nuevo documento JSON de política:

```
aws opensearchserverless update-access-policy \
    --name sales-inventory \
    --type data \
    --policy-version MTY2NDA1NDE4MDg1OF8x \
```

```
    --policy file://my-new-policy.json
```

Pueden transcurrir unos minutos de retardo entre el momento en que se actualiza la política y el momento en que se aplican los nuevos permisos.

Eliminación de políticas de acceso a los datos

Al eliminar una política de acceso a los datos, todas las colecciones asociadas pierden el acceso definido en la política. Asegúrese de que sus usuarios de IAM y SAML tengan el acceso adecuado a la colección antes de eliminar una política. Para eliminar una política en la consola OpenSearch sin servidor, selecciónela y elija Eliminar.

También puede usar el DeleteAccessPolicycomando:

```
aws opensearchserverless delete-access-policy --name my-policy --type data
```

Acceso a los datos entre cuentas

Si bien no puede crear una política de acceso a los datos con identidades o colecciones entre cuentas, sí puede configurar el acceso entre cuentas con la opción de asumir un rol. Por ejemplo, si la *account-a* es propietaria de una colección a la que la *account-b* necesita acceso, el usuario de la *account-b* puede asumir un rol en la *account-a*. El rol debe tener los permisos de IAM `aoss:APIAccessAll` y `aoss:DashboardsAccessAll` y estar incluido en la política de acceso a los datos de la *account-a*.

Acceso al plano de datos a través de AWS PrivateLink

Amazon OpenSearch Serverless admite dos tipos de AWS PrivateLink conexiones para las operaciones del plano de control y del plano de datos. Las operaciones del plano de control incluyen la creación y eliminación de colecciones y la administración de las políticas de acceso. Las operaciones del plano de datos sirven para indexar y consultar los datos de una colección. En esta página se describen los puntos finales de VPC del plano de datos. Para obtener información sobre los AWS PrivateLink puntos finales del plano de control, consulte. the section called “Punto final de VPC del plano de control”

Puede usarlo AWS PrivateLink para crear una conexión privada entre su VPC y Amazon OpenSearch Serverless. Puede acceder a OpenSearch Serverless como si estuviera en su VPC, sin el uso de una puerta de enlace a Internet, un dispositivo NAT, una conexión VPN o una conexión. Direct Connect Las instancias de su VPC no necesitan direcciones IP públicas para acceder

OpenSearch a Serverless. Para obtener más información sobre el acceso a la red de VPC, consulte Patrones de conectividad de red para Amazon OpenSearch Serverless.

Esta conexión privada se establece mediante la creación de un punto de conexión de interfaz alimentado por AWS PrivateLink. Creamos una interfaz de red de punto de conexión en cada subred especificada para el punto de conexión de interfaz. Se trata de interfaces de red administradas por el solicitante que sirven como punto de entrada para el tráfico destinado a Serverless. OpenSearch

Para obtener más información, consulte Acceso a Servicios de AWS a través de AWS PrivateLink en la Guía de AWS PrivateLink .

Temas

- Resolución de DNS de los puntos de conexión de colección
- VPCs y políticas de acceso a la red
- VPCs y políticas de puntos finales
- Consideraciones
- Permisos necesarios
- Cree un punto final de interfaz para Serverless OpenSearch
- Configuración de VPC compartida para Amazon Serverless OpenSearch

Resolución de DNS de los puntos de conexión de colección

Al crear un punto final de VPC del plano de datos a través de la consola OpenSearch sin servidor, el servicio crea una nueva zona alojada Amazon Route 53 privada y la adjunta a la VPC. Esta zona alojada privada consta de un registro para convertir el registro DNS comodín de las colecciones OpenSearch sin servidor (`*.us-east-1.aoss.amazonaws.com`) en las direcciones de interfaz utilizadas para el punto final. Solo necesita un punto final de VPC OpenSearch sin servidor en una VPC para acceder a todas y cada una de las colecciones y paneles de control de cada una. Región de AWS Cada VPC con un punto final para OpenSearch Serverless tiene su propia zona alojada privada adjunta.

El punto final de la interfaz OpenSearch sin servidor también crea un registro DNS comodín público de Route 53 para todas las colecciones de la región. El nombre DNS se resuelve en las direcciones IP públicas OpenSearch sin servidor. Los clientes VPCs que no tengan un punto final de VPC OpenSearch sin servidor o los clientes de redes públicas pueden usar la resolución pública de Route 53 y acceder a las colecciones y los paneles con esas direcciones IP. El tipo de dirección IP

(IPv4 IPv6, o Dualstack) del punto final de la VPC se determina en función de las subredes que se proporcionan al crear un punto final de interfaz para Serverless. OpenSearch

> **Note**
>
> OpenSearch Serverless crea una zona alojada privada `<region>.opensearch.amazonaws.com` (`) de Amazon Route 53 adicional para OpenSearch una resolución de dominio de servicio. Puede actualizar su punto final de IPv4 VPC existente a Dualstack mediante el update-vpc-endpointcomando de. AWS CLI

La dirección de resolver de DNS de una VPC determinada es la segunda dirección IP del CIDR de la VPC. Todos los clientes de la VPC deben usar ese resolver para obtener la dirección de punto de conexión de VPC para cualquier colección. El solucionador utiliza una zona alojada privada creada por Serverless. OpenSearch Basta con usar ese resolverr para todas las colecciones de cualquier cuenta. También es posible utilizar el resolver de VPC para algunos puntos de conexión de colección y el resolver público para otros, aunque normalmente no es necesario.

VPCs y políticas de acceso a la red

Para conceder permisos de red OpenSearch APIs y paneles de control a sus colecciones, puede utilizar las políticas de acceso a la red OpenSearch sin servidor. Puede controlar este acceso a la red desde sus puntos de conexión de VPC o desde la Internet pública. Como su política de red solo controla los permisos de tráfico, también debe configurar una política de acceso a los datos que especifique los permisos para operar con los datos de una colección y sus índices. Piense en un punto final de VPC OpenSearch sin servidor como un punto de acceso al servicio, una política de acceso a la red como el punto de acceso a nivel de red a las colecciones y los paneles de control, y una política de acceso a los datos como el punto de acceso para un control de acceso detallado para cualquier operación con los datos de la recopilación.

Como puede especificar varios puntos de enlace de VPC IDs en una política de red, le recomendamos que cree un punto de enlace de VPC para cada VPC que necesite acceder a una colección. VPCsPueden pertenecer a AWS cuentas distintas de la cuenta propietaria de la política de red y recopilación OpenSearch sin servidor. No le recomendamos que cree una VPC-to-VPC solución de interconexión u otra solución de proxy entre dos cuentas para que la VPC de una cuenta pueda utilizar el punto de enlace de VPC de otra cuenta. Esto es menos seguro y rentable que cada VPC que tenga su propio punto de conexión. El administrador de la otra VPC, que ha configurado el acceso al punto final de esa VPC en la política de red, no podrá ver fácilmente la primera VPC.

VPCs y políticas de puntos finales

Amazon OpenSearch Serverless admite políticas de puntos finales para VPCs. Una política de punto final es una política de IAM basada en recursos que se adjunta a un punto final de VPC para controlar qué entidades AWS principales pueden usar el punto final para acceder a su servicio. AWS Para obtener más información, consulte Controlar el acceso a puntos de conexión de VPC con políticas de punto de conexión.

Para usar una política de punto de conexión, primero debe crear un punto de conexión de interfaz. Puedes crear un punto final de interfaz mediante la consola sin servidor o la API OpenSearch sin servidor. OpenSearch Después de crear el punto de conexión de interfaz, tendrá que añadir la política del punto de conexión al punto de conexión. Para obtener más información, consulte the section called “Creación de un punto de conexión de interfaz”.

> **Note**
>
> No puede definir una política de puntos finales directamente en la consola de OpenSearch servicio.

Una política de punto de conexión no anula ni reemplaza otras políticas basadas en identidades, políticas basadas en recursos, políticas de red o políticas de acceso a datos que haya configurado. Para obtener información sobre cómo actualizar la política de puntos de conexión, consulte Controlar el acceso a puntos de conexión de VPC con políticas de punto de conexión.

De forma predeterminada, una política de punto de conexión concede acceso completo al punto de conexión de VPC.

```
{
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": "*",
            "Action": "*",
            "Resource": "*"
        }
    ]
}
```

Si bien la política de puntos de conexión de VPC predeterminada concede acceso total a los puntos de conexión, puede configurar una política de puntos de conexión de VPC para permitir el acceso a roles y usuarios específicos. Para eso, vea el siguiente ejemplo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": [
                    "123456789012",
                    "987654321098"
                ]
            },
            "Action": "*",
            "Resource": "*"
        }
    ]
}
```

Puede especificar una colección OpenSearch sin servidor para incluirla como elemento condicional en su política de puntos finales de VPC. Para eso, vea el siguiente ejemplo:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": "*",
            "Action": "*",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aoss:collection": [
                        "coll-abc"
```

```
                    ]
                }
            }
        }
    ]
}
```

El soporte para `aoss:CollectionId` está disponible.

```
Condition": {
         "StringEquals": {
               "aoss:CollectionId": "collection-id"
          }
}
```

Puede usar las identidades de SAML en la política de puntos de conexión de VPC para determinar el acceso a esos puntos. Debe usar un comodín (*) en la sección de la entidad principal de la política de puntos de conexión de VPC. Para eso, vea el siguiente ejemplo:

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"
                    ]
                }
            }
        }
    ]
```

```
}
```

Además, puede configurar su política de puntos de conexión para que incluya una política de entidad principal de SAML específica. Para eso, vea lo siguiente:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": "*",
            "Action": "*",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:PrincipalTag/Department": [
                        "Engineering"]
                    }
                }
            }
        ]
    }
```

Para obtener más información sobre el uso de la autenticación SAML con Amazon OpenSearch Serverless, consulte Autenticación SAML para Amazon Serverless. OpenSearch

También puede incluir usuarios de IAM y SAML en la misma política de puntos de conexión de VPC. Para eso, vea el siguiente ejemplo:

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": "*"
        }
    ]
}
```

También puede acceder a una colección de Amazon OpenSearch Serverless desde Amazon EC2 a través de los puntos de enlace de la interfaz de la VPC. Para obtener más información, consulte Acceder a una colección OpenSearch sin servidor desde Amazon EC2 (mediante puntos de enlace de la interfaz de VPC).

Consideraciones

Antes de configurar un punto final de interfaz para OpenSearch Serverless, tenga en cuenta lo siguiente:

- OpenSearch Serverless permite realizar llamadas a todas las operaciones de OpenSearch API compatibles (no a las operaciones de API de configuración) a través del punto final de la interfaz.
- Después de crear un punto final de interfaz para OpenSearch Serverless, aún debe incluirlo en las políticas de acceso a la red para que pueda acceder a las colecciones sin servidor.

- De forma predeterminada, se permite el acceso total a OpenSearch Serverless a través del punto final de la interfaz. Puede asociar un grupo de seguridad a las interfaces de red de los puntos finales para controlar el tráfico a OpenSearch Serverless a través del punto final de la interfaz.
- Una sola unidad Cuenta de AWS puede tener un máximo de 50 puntos finales de VPC OpenSearch sin servidor.
- Si habilita el acceso público a Internet a la API o los paneles de control de su colección en una política de red, cualquier VPC y la Internet pública pueden acceder a su colección.
- Si está en las instalaciones y fuera de la VPC, no puede usar directamente una resolución de DNS para la resolución del punto final de la VPC OpenSearch sin servidor. Si necesita acceso a una VPN, la VPC necesita un resolver de proxy de DNS para que lo usen los clientes externos. Route 53 ofrece una opción de punto de conexión de entrada que puede usar para resolver consultas de DNS a su VPC desde su red en las instalaciones u otra VPC.
- El servicio administra la zona alojada privada que OpenSearch Serverless crea y adjunta a la VPC, pero aparece en sus Amazon Route 53 recursos y se factura a su cuenta.
- Para otras consideraciones, consulte Consideraciones en la Guía de AWS PrivateLink .

Permisos necesarios

El acceso a la VPC para OpenSearch Serverless utiliza los siguientes permisos AWS Identity and Access Management (IAM). Puede especificar las condiciones de IAM para restringir a los usuarios a colecciones específicas.

- `aoss:CreateVpcEndpoint`: cree un punto de conexión de VPC.
- `aoss:ListVpcEndpoints`: enumere todos los puntos de conexión de VPC.
- `aoss:BatchGetVpcEndpoint`: consulte los detalles sobre un subconjunto de puntos de conexión de VPC.
- `aoss:UpdateVpcEndpoint`: modifique un punto de conexión de VPC.
- `aoss:DeleteVpcEndpoint`: elimine un punto de conexión de VPC.

Además, necesita los siguientes permisos de Amazon EC2 y Route 53 para crear un punto de conexión de 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`

Cree un punto final de interfaz para Serverless OpenSearch

Puede crear un punto final de interfaz para OpenSearch Serverless mediante la consola o la API OpenSearch Serverless.

Para crear un punto final de interfaz para una colección sin servidor OpenSearch

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, expanda Sin servidor y seleccione Puntos de conexión de VPC.
3. Seleccione Crear punto de conexión de VPC.
4. Proporcione un nombre para el punto de conexión.
5. En el caso de la VPC, selecciona la VPC desde la que accederás a Serverless. OpenSearch
6. En el caso de las subredes, selecciona una subred desde la que accederás a Serverless. OpenSearch

    - La dirección IP y el tipo de DNS del punto de conexión se basan en el tipo de subred.
        - Dualstack: si todas las subredes tienen ambos rangos de direcciones IPv4 IPv6

- IPv6: Si todas las subredes son solo subredes IPv6
- IPv4: Si todas las subredes tienen rangos de direcciones IPv4

7. En Grupos de seguridad, seleccione los grupos de seguridad que deban asociarse a las interfaces de red del punto de conexión. Este es un paso fundamental en el que se limitan los puertos, los protocolos y las fuentes del tráfico entrante que se autoriza a entrar en el punto final. Asegúrese de que las reglas del grupo de seguridad permitan que los recursos que utilizarán el punto final de la VPC se comuniquen con OpenSearch Serverless para comunicarse con la interfaz de red del punto final.
8. Seleccione Crear punto de conexión.

Para crear un punto final de VPC mediante la API OpenSearch sin servidor, utilice el comando. `CreateVpcEndpoint`

> **Note**
>
> Después de crear un punto de conexión, registre su ID (por ejemplo, `vpce-abc123def4EXAMPLE`. Para proporcionar el acceso del punto de conexión a sus colecciones, debe incluir este ID en una o más políticas de acceso a la red.

Tras crear un punto de conexión de interfaz, debe proporcionarles acceso a las colecciones mediante políticas de acceso a la red. Para obtener más información, consulte the section called “Acceso a la red”.

Configuración de VPC compartida para Amazon Serverless OpenSearch

Puede utilizar Amazon Virtual Private Cloud (VPC) para compartir subredes de VPC con otros Cuentas de AWS miembros de su organización, así como compartir la infraestructura de red, como una VPN, entre varios recursos. Cuentas de AWS

Actualmente, Amazon OpenSearch Serverless no admite la creación de una AWS PrivateLink conexión en una VPC compartida a menos que usted sea propietario de esa VPC. AWS PrivateLink tampoco admite el uso compartido de conexiones entre ellas. Cuentas de AWS

Sin embargo, en función de la arquitectura flexible y modular de OpenSearch Serverless, aún puede configurar una VPC compartida. Esto se debe a que la infraestructura de red OpenSearch sin servidor es independiente de la infraestructura de recopilación (OpenSearch servicio) individual. Por lo tanto, puede crear un AWS PrivateLink VPCe punto de conexión para una cuenta en la que

esté ubicada una VPC y, a continuación, usar un VPCe ID en la política de red de otras cuentas para restringir el tráfico y evitar que provenga únicamente de esa VPC compartida.

Los siguientes procedimientos se refieren a una cuenta propietaria y una cuenta consumidora.

Un propietario de la cuenta actúa como una cuenta de redes común en la que se configura una VPC y se comparte con otras cuentas. Las cuentas de consumidor son aquellas cuentas que crean y mantienen sus colecciones OpenSearch sin servidor en la VPC que comparte con ellos la cuenta del propietario.

Requisitos previos

Asegúrese de que se cumplan los siguientes requisitos antes de configurar la VPC compartida:

- El propietario de la cuenta debe haber configurado previamente una VPC, subredes, tabla de enrutamiento y otros recursos necesarios en nube privada virtual de Amazon. Para obtener más información, consulte la Guía del usuario de Amazon VPC.
- El propietario de la cuenta y las cuentas consumidoras deben pertenecer a la misma organización en AWS Organizations. Para obtener más información, consulte la Guía del usuario de AWS Organizations.

Para configurar una VPC compartida en una cuenta de account/common red propietaria.

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. Siga los pasos de the section called “Creación de un punto de conexión de interfaz”. Mientras realiza el procedimiento, haga las siguientes selecciones:
   - Seleccione una VPC y subredes que estén compartidas con las cuentas consumidoras de su organización.
3. Tras crear el punto de conexión, anota el VPCe ID que se genera y entrégalo a los administradores que van a realizar la tarea de configuración de las cuentas de los consumidores.

   VPCe IDs están en el formato`vpce-abc123def4EXAMPLE`.

Para configurar una VPC compartida en una cuenta consumidora:

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.

2. Use la información de the section called “Administración de colecciones” para crear una colección, si aún no tiene una.
3. Use la información de the section called “Creación de políticas de (consola)” para crear una política de red. A medida que avanza, realice las siguientes selecciones.

   > **Note**
   >
   > También puede actualizar una política de red existente para este fin.

   a. En Tipo de acceso, seleccione VPC (recomendado).
   b. Para acceder a los puntos finales de VPC, elija el VPCe ID que le proporcionó la cuenta propietaria, en el formato. `vpce-abc123def4EXAMPLE`
   c. En el área Tipo de recurso, haga lo siguiente:

      - Seleccione la casilla Habilitar el acceso al OpenSearch punto final y, a continuación, seleccione el nombre de la colección o el patrón de recopilación que se utilizará para habilitar el acceso desde esa VPC compartida.
      - Seleccione la casilla Habilitar el acceso al OpenSearch panel de control y, a continuación, seleccione el nombre o el patrón de la colección que se utilizará para habilitar el acceso desde esa VPC compartida.
4. Para una nueva política, seleccione Crear. Para una política existente, seleccione Actualizar.

Acceso al plano de control a través de AWS PrivateLink

Amazon OpenSearch Serverless admite dos tipos de AWS PrivateLink conexiones para las operaciones del plano de control y del plano de datos. Las operaciones del plano de control incluyen la creación y eliminación de colecciones y la administración de las políticas de acceso. Las operaciones del plano de datos sirven para indexar y consultar los datos de una colección. Esta página describe el punto final del plano de control. AWS PrivateLink Para obtener información sobre los puntos finales de VPC del plano de datos, consulte. the section called “Punto final de VPC del plano de datos”

Creación de un punto final del plano de control AWS PrivateLink

Puede mejorar la seguridad de su VPC configurando OpenSearch Serverless para que utilice un punto final de VPC de interfaz. Los puntos finales de la interfaz funcionan con. AWS PrivateLink Esta

tecnología le permite acceder de forma privada a OpenSearch Serverless APIs sin una puerta de enlace a Internet, un dispositivo NAT, una conexión VPN o una conexión AWS Direct Connect.

Para obtener más información sobre AWS PrivateLink los puntos de enlace de VPC, consulte los puntos de enlace de VPC en la Guía del usuario de Amazon VPC.

Consideraciones

- Los puntos finales de VPC solo se admiten en la misma región.
- Los puntos de conexión de VPC solo admiten DNS proporcionadas por Amazon a través de Amazon Route 53.
- Los puntos de enlace de VPC admiten políticas de puntos finales para controlar el acceso a colecciones OpenSearch , políticas y. VpcEndpoints
- OpenSearch Serverless solo admite puntos finales de interfaz. Los puntos de conexión de puerta de enlace no se admiten.

Creación del punto final de la VPC

Para crear el punto de enlace de VPC del plano de control para Amazon OpenSearch Serverless, utilice el procedimiento Acceso a un AWS servicio mediante un punto de enlace de VPC de interfaz de la Guía para desarrolladores de Amazon VPC. Cree el siguiente punto final:

- `com.amazonaws.region.aoss`

Para crear un punto final de VPC del plano de control mediante la consola

1. Abra la consola de Amazon VPC en https://console.aws.amazon.com/vpc/.
2. En el panel de navegación, elija Puntos de conexión.
3. Elija Crear punto de conexión.
4. En Categoría de servicios, elija Servicios de AWS.
5. Para Servicios, elija`com.amazonaws.region.aoss`. Por ejemplo, `com.amazonaws.us-east-1.aoss`.
6. En VPC, elija la VPC en la que desea crear el punto de conexión.
7. En Subredes, seleccione las subredes (zonas de disponibilidad) en las que va a crear las interfaces de red de punto de conexión.

8. En el caso de los grupos de seguridad, elija los grupos de seguridad que desee asociar a las interfaces de red de los puntos finales. Asegúrese de que se permita HTTPS (puerto 443).
9. En Política, selecciona Acceso total para permitir todas las operaciones o selecciona Personalizado para adjuntar una política personalizada.
10. Seleccione Crear punto de conexión.

Creación de una política de punto de conexión

Puede adjuntar una política de punto final a su punto de enlace de VPC que controle el acceso a Amazon OpenSearch Serverless. La política especifica la siguiente información:

- La entidad principal que puede realizar acciones.
- Las acciones que se pueden realizar.
- Los recursos en los que se pueden llevar a cabo las acciones.

Para más información, consulte Control del acceso a los servicios con puntos de enlace de la VPC en la Guía del usuario de Amazon VPC.

Example Política de puntos finales de VPC para Serverless OpenSearch

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": "*",
      "Action": [
        "aoss:ListCollections",
        "aoss:BatchGetCollection"
      ],
      "Resource": "*"
    }
  ]
}
```

Example Política restrictiva que solo permite operaciones de listas

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

```
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": "*",
      "Action": "aoss:ListCollections",
      "Resource": "*"
    }
  ]
}
```

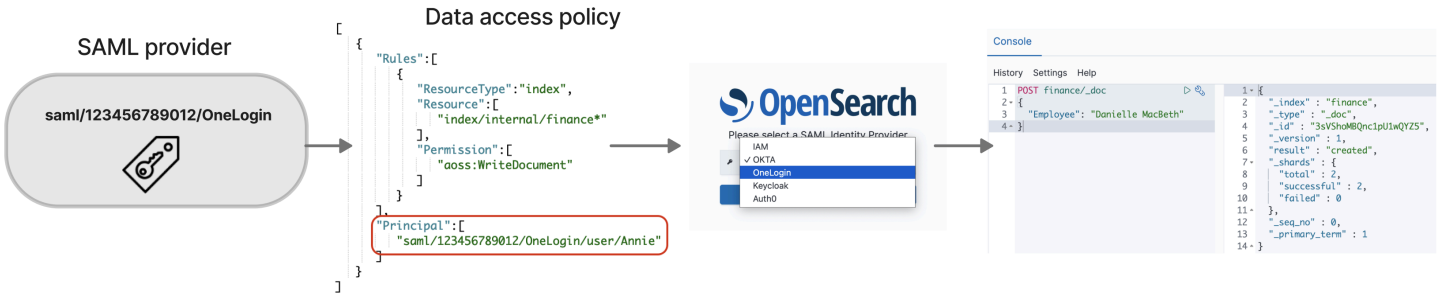
Autenticación SAML para Amazon Serverless OpenSearch

Con la autenticación SAML para Amazon OpenSearch Serverless, puede usar su proveedor de identidad actual para ofrecer el inicio de sesión único (SSO) para los puntos de enlace de los OpenSearch paneles de control de las colecciones sin servidor.

La autenticación SAML le permite utilizar proveedores de identidad externos para iniciar sesión en Dashboards con el fin de indexar y buscar datos. OpenSearch OpenSearch Serverless es compatible con los proveedores que utilizan el estándar SAML 2.0, como IAM Identity Center, Okta, Keycloak, Active Directory Federation Services (AD FS) y Auth0. Puede configurar el Centro de identidades de IAM para sincronizar usuarios y grupos de otras fuentes de identidad, como Okta OneLogin y Microsoft Entra ID. Para obtener una lista de los orígenes de identidades que admite IAM Identity Center y los pasos para configurarlos, consulte Getting started tutorials en la Guía del usuario de IAM Identity Center.

> **Note**
>
> La autenticación SAML solo sirve para acceder a los OpenSearch paneles de control a través de un navegador web. Los usuarios autenticados solo pueden realizar solicitudes a las operaciones de la OpenSearch API a través de las herramientas de desarrollo de los paneles de control. OpenSearch Sus credenciales de SAML no le permiten realizar solicitudes HTTP directas a las operaciones de la OpenSearch API.

Para configurar la autenticación SAML, primero debe configurar un proveedor de identidades (IdP) SAML. A continuación, se incluyen uno o más usuarios de ese IdP en una política de acceso a datos. Esta política le otorga ciertos permisos para los and/or índices de colecciones. A continuación, un usuario puede iniciar sesión en los OpenSearch paneles de control y realizar las acciones que se permiten en la política de acceso a los datos.



Temas

- Consideraciones
- Permisos necesarios
- Crear proveedores SAML de (consola)
- Acceder a los paneles OpenSearch
- Concesión de acceso a las identidades de SAML a los datos de la colección
- Creación de proveedores SAML (AWS CLI)
- Ver proveedores SAML
- Actualización de proveedores SAML
- Eliminar proveedores SAML

Consideraciones

A la hora de configurar la autenticación SAML, tenga en cuenta lo siguiente:

- No se admiten las solicitudes firmadas y cifradas.
- No se admiten las aserciones cifradas.
- No se admite la autenticación ni el cierre de sesión iniciada por el IDP.
- Las políticas de control de servicios (SCP) no se aplicarán ni evaluarán en el caso de identidades que no sean de IAM (como SAML en OpenSearch Amazon Serverless y SAML y la autorización básica de usuario interna para Amazon Service). OpenSearch

Permisos necesarios

La autenticación SAML para OpenSearch Serverless utiliza los siguientes permisos (IAM): AWS Identity and Access Management

- `aoss:CreateSecurityConfig`: cree un proveedor SAML.
- `aoss:ListSecurityConfig`: enumere todos los proveedores de SAML en la cuenta actual.
- `aoss:GetSecurityConfig`: vea la información del proveedor de SAML.
- `aoss:UpdateSecurityConfig`: modifique la configuración de proveedor de SAML determinada, incluidos los metadatos XML.
- `aoss:DeleteSecurityConfig`: elimine un proveedor SAML.

La siguiente política de acceso basada en la identidad permite al usuario administrar todas las configuraciones de IdP:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
                "aoss:CreateSecurityConfig",
                "aoss:DeleteSecurityConfig",
                "aoss:GetSecurityConfig",
                "aoss:UpdateSecurityConfig",
                "aoss:ListSecurityConfigs"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

Tenga en cuenta que el elemento `Resource` debe ser un comodín.

Crear proveedores SAML de (consola)

En estos pasos se explica cómo crear proveedores de SAML. Esto permite la autenticación SAML con la autenticación iniciada por el proveedor de servicios (SP) para los paneles. OpenSearch No se admite la autenticación IdP.

Para habilitar la autenticación SAML en los paneles OpenSearch

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, expanda Sin servidor y seleccione Autenticación SAML.
3. Seleccione Agregar proveedor SAML.
4. Escriba un nombre y una descripción para el proveedor.

   > **Note**
   >
   > El nombre que especifique es de acceso público y aparecerá en un menú desplegable cuando los usuarios inicien sesión en los OpenSearch paneles. Asegúrese de que el nombre sea fácilmente reconocible y no revele información confidencial sobre su proveedor de identidad.

5. En Configurar IdP, copie la URL del servicio al consumidor de aserciones (ACS).
6. Utilice la URL de ACS que acaba de copiar para configurar su proveedor de identidades. La terminología y los pasos varían según el proveedor. Consulte la documentación de su proveedor.

   En Okta, por ejemplo, se crea una “aplicación web SAML 2.0” y se especifica la URL de ACS como URL de inicio de sesión única, URL de destinatario y URL de destino. En el caso de Auth0, debe especificarlo en Allowed Callback. URLs
7. Indique la restricción de audiencia si su IdP tiene un campo para ello. La restricción de audiencia es un valor dentro de la aserción SAML que especifica a quién va dirigida esta. Con OpenSearch Serverless, puede hacer lo siguiente. Asegúrese de reemplazar el *`content`* siguiente ejemplo de código por su propio Cuenta de AWS ID:

   1. Use la restricción de audiencia predeterminada `:opensearch:`*`111122223333`*.
   2. (Opcional) configure una restricción de audiencia personalizada mediante AWS CLI. Para obtener más información, consulte Creación de proveedores SAML (AWS CLI).

   El nombre del campo de restricción de audiencia varía según el proveedor. Para Okta, es Audiencia URI (ID de entidad del SP). En el caso de IAM Identity Center, es la audiencia de Application SAML.

8. Si utiliza IAM Identity Center, también debe especificar el siguiente asignación de atributos: `Subject=${user:name}`, con un formato de `unspecified`.
9. Después de configurar el proveedor de identidades, genera un archivo de metadatos de IdP. Este archivo XML contiene información sobre el proveedor, como un certificado TLS, puntos de conexión de inicio de sesión único y el ID de entidad del proveedor de identidad.

   Copie el texto del archivo de metadatos del IdP y péguelo en Proporcionar metadatos desde el campo de IdP. Alternativamente, seleccione Importar desde archivo XML y cargue el archivo. El archivo de metadatos debe tener un aspecto similar al siguiente:

```
<?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. Mantenga vacío el campo Atributo ID de usuario personalizado para utilizar el elemento `NameID` de la aserción SAML para el nombre de usuario. Si su aserción no utiliza este elemento estándar y, en su lugar, incluye el nombre de usuario como un atributo personalizado, especifique ese atributo aquí. Los atributos distinguen entre mayúsculas y minúsculas. Solo se admite un único atributo de usuario.

    El siguiente ejemplo muestra un atributo de anulación para `NameID` en la aserción 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. (Opcional) Especifique un atributo personalizado en el campo Atributos del grupo, como `role` o `group`. Solo se admite un único atributo de grupo. No hay ningún atributo de grupo predeterminado. Si no especifica ninguna, sus políticas de acceso a datos solo pueden contener entidades principales de usuario.

    El siguiente ejemplo muestra un atributo de grupo en la aserción 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. De forma predeterminada, OpenSearch Dashboards cierra la sesión de los usuarios después de 24 horas. Puede configurar este valor en cualquier número entre 1 y 12 horas (15 y 720 minutos) especificando el tiempo de espera de los OpenSearch paneles. Si intenta establecer el tiempo de espera igual o inferior a 15 minutos, la sesión se restablecerá a una hora.
13. Seleccione Crear proveedor.

Acceder a los paneles OpenSearch

Tras configurar un proveedor de SAML, todos los usuarios y grupos asociados a ese proveedor pueden navegar hasta el punto de conexión de los OpenSearch paneles. La URL de Dashboards tiene el formato *`collection-endpoint`*`/_dashboards/` para todas las colecciones.

Si tienes SAML activado, al seleccionar el enlace que aparece en la Consola de administración de AWS página de selección de IdP, donde puedes iniciar sesión con tus credenciales de SAML. Primero, use el menú desplegable para seleccionar un proveedor de identidades:



Inicie sesión con sus credenciales de usuario de IdP.

Si no tienes el SAML activado, al seleccionar el enlace que aparece en él podrás Consola de administración de AWS iniciar sesión como usuario o rol de IAM, sin opción de usar SAML.

Concesión de acceso a las identidades de SAML a los datos de la colección

Después de crear un proveedor de SAML, aún tendrá que concederles acceso a los datos de sus colecciones a los usuarios y grupos subyacentes. Para concederles el acceso deberá hacerlo a través de políticas de acceso a datos. Hasta que no les proporcione acceso a los usuarios, estos no podrán leer, escribir ni eliminar ningún dato de sus colecciones.

Para conceder el acceso, cree una política de acceso a los datos y especifique su and/or grupo IDs de usuarios de SAML en la siguiente declaración: `Principal`

```
[
```

```
   {
      "Rules":[
       ...
      ],
      "Principal":[
         "saml/987654321098/myprovider/user/Shaheen",
         "saml/987654321098/myprovider/group/finance"
      ]
   }
]
```

Puede concederles acceso a colecciones, índices o ambos. Si desea que diferentes usuarios tengan permisos diferentes, cree varias reglas. Para obtener una lista de los permisos disponibles, consulte Permisos de políticas compatibles. Para obtener información sobre cómo dar formato a una política de acceso, consulte Sintaxis de políticas.

Creación de proveedores SAML (AWS CLI)

Para crear un proveedor de SAML mediante la API OpenSearch sin servidor, envía una solicitud: CreateSecurityConfig

```
aws opensearchserverless create-security-config \
    --name myprovider \
    --type saml \
    --saml-options file://saml-auth0.json
```

Especifique `saml-options`, incluido el XML de los metadatos, como un mapa clave-valor dentro de un archivo .json. El XML de los metadatos debe estar codificado como una cadena de 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**
>
> (Opcional) configure una restricción de audiencia personalizada mediante AWS CLI. Para obtener más información, consulte Creación de proveedores SAML (AWS CLI).

Ver proveedores SAML

La siguiente ListSecurityConfigssolicitud muestra una lista de todos los proveedores de SAML de tu cuenta:

```
aws opensearchserverless list-security-configs --type saml
```

La solicitud devuelve información sobre todos los proveedores de SAML existentes, incluidos los metadatos completos del IdP que genera su proveedor de identidades:

```
{
   "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"
         }
      }
   ]
}
```

Para ver los detalles de un proveedor específico, incluida la `configVersion` para actualizaciones futuras, envíe una solicitud `GetSecurityConfig`.

Actualización de proveedores SAML

Para actualizar un proveedor de SAML mediante la consola OpenSearch sin servidor, elige la autenticación SAML, selecciona tu proveedor de identidad y selecciona Editar. Puede modificar todos los campos, incluidos los metadatos y los atributos personalizados.

Para actualizar un proveedor a través de la API OpenSearch sin servidor, envía una UpdateSecurityConfigsolicitud e incluye el identificador de la política que se va a actualizar. También debe incluir una versión de configuración, que puede recuperar mediante los comandos `ListSecurityConfigs` o `GetSecurityConfig`. Incluir la versión más reciente garantiza que no se anule inadvertidamente un cambio realizado por otra persona.

La siguiente solicitud actualiza las opciones de SAML de un proveedor:

```
aws opensearchserverless update-security-config \
    --id saml/123456789012/myprovider \
    --type saml \
    --saml-options file://saml-auth0.json \
    --config-version MTY2NDA1MjY4NDQ5M18x
```

Especifique las opciones de configuración de SAML como un mapa clave-valor dentro de un archivo .json.

> **Important**
>
> Las actualizaciones de las opciones de SAML no son graduales. Si al realizar una actualización no especifica el valor para un parámetro en el objeto `SAMLOptions`, los valores existentes se sustituirán por valores vacíos. Por ejemplo, si la configuración actual contiene un valor para `userAttribute` y luego hace una actualización y no incluye este valor, el valor se elimina de la configuración. Asegúrese de saber cuáles son los valores existentes antes de realizar una actualización llamando a la operación `GetSecurityConfig`.

Eliminar proveedores SAML

Al eliminar un proveedor de SAML, las referencias de los usuarios y grupos asociados en las políticas de acceso a datos dejan de funcionar. Para evitar confusiones, le recomendamos que

elimine todas las referencias al punto de conexión en sus políticas de acceso antes de eliminar el punto de conexión.

Para eliminar un proveedor de SAML mediante la consola OpenSearch sin servidor, elija Autenticación, seleccione el proveedor y elija Eliminar.

Para eliminar un proveedor a través de la API OpenSearch sin servidor, envía una solicitud: DeleteSecurityConfig

```
aws opensearchserverless delete-security-config --id saml/123456789012/myprovider
```

Validación de conformidad para Amazon OpenSearch Serverless

Los auditores externos evalúan la seguridad y el cumplimiento de Amazon OpenSearch Serverless como parte de varios programas de AWS cumplimiento. Entre estos, se incluyen SOC, PCI e HIPAA.

Para saber si un programa de conformidad Servicio de AWS se encuentra dentro del ámbito de aplicación de programas de conformidad específicos, consulte Servicios de AWS Alcance por programa de conformidad Servicios de AWS de conformidad y elija el programa de conformidad que le interese. Para obtener información general, consulte Programas de AWS cumplimiento > Programas AWS .

Puede descargar informes de auditoría de terceros utilizando AWS Artifact. Para obtener más información, consulte Descarga de informes en AWS Artifact .

Su responsabilidad de cumplimiento al Servicios de AWS utilizarlos viene determinada por la confidencialidad de sus datos, los objetivos de cumplimiento de su empresa y las leyes y reglamentos aplicables. Para obtener más información sobre su responsabilidad de conformidad al utilizarlos Servicios de AWS, consulte AWS la documentación de seguridad.

Etiquetado de colecciones de Amazon OpenSearch Serverless

Las etiquetas le permiten asignar información arbitraria a una colección de Amazon OpenSearch Serverless para que pueda categorizar y filtrar esa información. Una etiqueta es una etiqueta de metadatos que usted asigna o que AWS asigna a un recurso. AWS

Cada etiqueta consta de una clave y un valor. En el caso de etiquetas que usted asigna, debe definir la clave y el valor. Por ejemplo, puede definir la clave como `stage` y el valor de un recurso como `test`.

Con las etiquetas, puede identificar y organizar sus AWS recursos. Muchos AWS servicios admiten el etiquetado, por lo que puede asignar la misma etiqueta a los recursos de diferentes servicios para indicar que los recursos están relacionados. Por ejemplo, puedes asignar la misma etiqueta a una colección OpenSearch sin servidor que a un dominio de Amazon OpenSearch Service.

En OpenSearch Serverless, el recurso principal es una colección. Puede usar la consola de OpenSearch servicio AWS CLI, las operaciones de la API OpenSearch sin servidor o la AWS SDKs para agregar, administrar y eliminar etiquetas de una colección.

Permisos necesarios

OpenSearch Serverless utiliza los siguientes permisos AWS Identity and Access Management Access Analyzer (IAM) para etiquetar colecciones:

- `aoss:TagResource`
- `aoss:ListTagsForResource`
- `aoss:UntagResource`

Etiquetado de colecciones (consola)

La consola es la forma más sencilla para etiquetar un dominio.

Para crear una etiqueta (consola)

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. Expanda Serverless (Sin servidor) en el panel de navegación de la izquierda y seleccione Collections (Colecciones).
3. Seleccione la colección a la que desee agregar etiquetas y vaya a la pestaña Tags (Etiquetas).
4. Seleccione Administrar y Add new tag (Agregar nueva etiqueta).
5. Introduzca una clave de etiqueta y un valor opcional.
6. Seleccione guardar.

Para eliminar una etiqueta, siga los mismos pasos y elija Quitar en la página Administrar etiquetas.

Para obtener más información sobre cómo utilizar la consola para trabajar con etiquetas, consulte Tag Editor en la Guía de introducción a la consola de administración de AWS .

Etiquetado de colecciones (AWS CLI)

Para etiquetar una colección mediante el AWS CLI, envía una TagResourcesolicitud:

```
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
```

Vea las etiquetas existentes de una colección con el ListTagsForResourcecomando:

```
aws opensearchserverless list-tags-for-resource
  --resource-arn arn:aws:aoss:us-east-1:123456789012:collection/my-collection
```

Elimine las etiquetas de una colección mediante el UntagResourcecomando:

```
aws opensearchserverless untag-resource
  --resource-arn arn:aws:aoss:us-east-1:123456789012:collection/my-collection
  --tag-keys service
```

Operaciones y complementos compatibles en Amazon OpenSearch Serverless

Amazon OpenSearch Serverless admite diversos OpenSearch complementos, así como un subconjunto de las operaciones de la API de indexación, búsqueda y metadatos disponibles en. OpenSearch Puede incluir los permisos en la columna izquierda de la tabla dentro de las políticas de acceso a datos para limitar el acceso a determinadas operaciones.

Temas

- Operaciones y permisos de OpenSearch API compatibles
- OpenSearch Plugins compatibles

Operaciones y permisos de OpenSearch API compatibles

En la siguiente tabla, se enumeran las operaciones de API que admite OpenSearch Serverless, junto con sus correspondientes permisos de política de acceso a los datos:

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
`aoss:CreateIndex`	PUT <index>	Crea índices. Para obtener más información, consulte la sección Crear un índice. **Note** Este permiso también se aplica a la creación de índices con los datos de muestra en los OpenSearch paneles.
`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> • HEAD <index>	Describe los índices. Para obtener más información, consulte los siguientes recursos: • Obtener índice • Obtener un mapeo • Obtener la configuración • Index exists • CAT indices (la respuesta no incluye los campos `health` ni `status`).
`aoss:WriteDocument`	• DELETE <index>/_doc/<id> • POST <index>/_bulk	Redacta y actualiza documentos. Para obtener más información, consulte los recursos de siguientes:

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
	• POST <index>/_doc/ <id> (solo para tipos de colecciones de búsquedas) • POST <index>/_doc • POST <index>/_update/<id> (solo para tipos de colecciones de búsquedas) • POST _bulk • PUT <index>/_create/<id> (solo para tipos de colecciones de búsquedas) • PUT <index>/_doc/<id> (solo para tipos de colecciones de búsquedas)	• Bulk • Índice de datos **Note** Algunas operaciones solo están permitidas para colecciones de tipos SEARCH. Para obtener más información, consulte the section called “Elección de un tipo de colección”.

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
`aoss:ReadDocument`	• GET <index>/_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/point_in_time • POST <index>/_explain/ <id> • POST <index>/_count • POST <index>/_field_caps	Lee los documentos. Para obtener más información, consulte los siguientes recursos: • Realice el análisis de texto • Obtener el documento • Recuento • Consulte DSL • Evaluación de clasificación • Analice la API • Explain • Punto en el tiempo • SQL y PPL

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
	• POST <index>/_rank_eval • POST <index>/_search • POST _analyze • POST _field_caps • POST _search • DELETE /_search/point_in_time/_all • DELETE /_search/point_in_time	
`aoss:DeleteIndex`	DELETE <target>	Elimina índices. Para obtener más información, consulte Eliminar índices.
`aoss:UpdateIndex`	• POST _mapping • POST <index>/_mapping/ • POST <index>/_mappings/ • POST <index>/_setting • POST <index>/_setting • POST _setting • POST _settings • PUT _mapping • PUT <index>/_mapping • PUT <index>/_mapping • PUT <index>/_setting • PUT <index>/_settings • PUT _setting • PUT _settings	Actualiza las configuraciones del índice. Para obtener más información, consulte los siguientes recursos: • Mapeo • Actualizar las configuraciones

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
`aoss:CreateCollectionItems`	• POST _aliases • POST /_plugins/_flow_framework/workflow • * POST /_plugins/_flow_framework/workflow/<workflow_id>/_provision • PUT _ingest/pipeline/<pipeline-id> • PUT _search/pipeline/<pipeline-id>	• Crear alias, canalizaciones y plantillas de índices. Para obtener más información, consulte Crear alias. • * Aprovisionar o volver a aprovisionar plantillas. Los servicios ML Commons Client y OpenSearch Serverless administran las políticas dependientes.

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
`aoss:DescribeCollectionItems`	• GET <index>/_alias/ <alias> • GET _alias • GET _alias/ <alias> • GET _cat/alias • 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 • GET /_plugins/_flow_framework/workflow/<workflow-id> • GET /_plugins/_flow_framework/workflow/_search • GET /_plugins/_flow_framework/workflow/<workflow-id>/_status • OBTENGA /_plugins/_flow_ / _search framework/workflow/state • GET /_plugins/_flow_framework/workflow/_steps • GET /_plugins/_flow_framework/workflow/_step?workflow_step=<step_name> • GET _search/pipeline/<pipeline-id>	Describe cómo trabajar con alias, plantillas de índices y marcos y canalizaciones. Para obtener más información, consulte los siguientes recursos: • Administrar alias • Plantillas del índice

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
	• HEAD _alias/ <alias> • HEAD _component_template/ <component-template> • HEAD _index_template/ <name> • HEAD <index>/_alias/ <alias> • POST _ingest/pipeline/_simulate • POST /_plugins/_flow_framework/workflow/_search • PUBLICA /_plugins/_flow_ /_search framework/workflow/state	
`aoss:UpdateCollectionItems`	• POST <index>/_alias/ <alias> • POST <index>/_aliases/ <alias> • POST _component_template/ <component-template> • POST _index_template/ <index-template> • * POST /_plugins/_flow_framework/workflow/<workflow_id>/_deprovision • PUT <index>/_alias/ <alias> • PUT <index>/_aliases/ <alias> • PUT _component_template/ <component-template> • PUT _index_template/ <index-template> • PUT /_plugins/_flow_framework/workflow/<workflow_id>	Actualice los alias, las plantillas de índice y las plantillas de framework. Para obtener más informaci ón, consulte los siguientes recursos: • Alias de índice • Plantillas del índice *. La API para desaprovi sionar plantillas. Los servicios ML Commons Client y Serverless gestionan las políticas dependientes. OpenSearch

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
`aoss:DeleteCollectionItems`	• DELETE <index>/_alias/ <alias> • DELETE _component_template/ <component-template> • DELETE _index_template/ <index-template> • DELETE <index>/_aliases/ <alias> • DELETE _search/pipeline/< pipeline-id> • DELETE _ingest/pipeline/< pipeline-id> • DELETE /_plugins/_flow_framework/workflow/<workflow_id>	Elimina alias, plantillas de índices y marcos, y canalizaciones. Para obtener más información, consulte los siguientes recursos: • Eliminar alias • Elimina una plantilla

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
`aoss:DescribeMLResource`	• GET /_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 • OBTENGA /_plugins/_/ml/profile/tasks<task_id> • POST /_plugins/_ml/models/_search • POST /_plugins/_ml/model_groups/_search • POST /_plugins/_ml/connectors/_search	Describe GET y la búsqueda APIs para recuperar información sobre modelos y conectores.
`aoss:CreateMLResource`	• POST /_plugins/_ml/models/_register • POST /_plugins/_ml/model_groups/_register • POST /_plugins/_ml/connectors/_create	Proporciona permiso para crear recursos de ML.

Permiso de política de acceso a datos	OpenSearch Operaciones de API	Descripción y advertencias
`aoss:UpdateMLResource`	• PUT /_plugins/_ml/models/<model_id> • POST /_plugins/_ml/models/<model_id>/_deploy • POST /_plugins/_ml/models/<model_id>/_undeploy • PUT /_plugins/_ml/model_groups/<model_group_id> • PUT /_plugins/_ml/connectors/<connector_id>	Otorga permiso para actualizar recursos de aprendizaje automático existentes.
`aoss:DeleteMLResource`	• DELETE /_plugins/_ml/models/<model_id> • DELETE _plugins/_ml/model_groups/<model_group_id> • DELETE /_plugins/_ml/connectors/<connector_id> • DELETE /_plugins/_ml/tasks/<task_id>	Proporciona permiso para eliminar recursos de aprendizaje automático.
`aoss:ExecuteMLResource`	• POST /_plugins/_ml/models/<model_id>/_predict	Proporciona permiso para ejecutar modelos.

OpenSearch Plugins compatibles

OpenSearch Las colecciones sin servidor vienen preempaquetadas con los siguientes complementos de la OpenSearch comunidad. La ejecución sin servidor implementa y administra automáticamente los complementos.

Complementos de análisis

- ICU Analysis
- Japanese (kuromoji) Analysis

- Análisis coreano (Nori)
- Phonetic Analysis
- Smart Chinese Analysis
- Stempel Polish Analysis
- Ukrainian Analysis

Complementos de Mapper

- Mapper Size
- Mapper Murmur3
- Texto anotado de Mapper

Complementos de la creación de scripts

- Painless
- Expression
- Mustache

Además, OpenSearch Serverless incluye todos los complementos que se envían como módulos.

Supervisión de Amazon OpenSearch Serverless

La supervisión es una parte importante del mantenimiento de la fiabilidad, la disponibilidad y el rendimiento de Amazon OpenSearch Serverless y sus demás AWS soluciones. AWS proporciona las siguientes herramientas de monitoreo para monitorear OpenSearch Serverless, informar cuando algo anda mal y tomar medidas automáticas cuando sea apropiado:

- Amazon CloudWatch monitorea tus AWS recursos y las aplicaciones en las que AWS ejecutas en tiempo real. Puede recopilar métricas y realizar un seguimiento de las métricas, crear paneles personalizados y definir alarmas que le advierten o que toman medidas cuando una métrica determinada alcanza el umbral que se especifique.

  Por ejemplo, puede CloudWatch hacer un seguimiento del uso de la CPU u otras métricas de sus instancias de Amazon EC2 y lanzar automáticamente nuevas instancias cuando sea necesario. Para obtener más información, consulta la Guía del CloudWatch usuario de Amazon.

- AWS CloudTrail captura las llamadas a la API y los eventos relacionados realizados por su cuenta de Cuenta de AWS o en su nombre. Entrega los archivos de registro al bucket de Amazon S3 que se especifique. Puede identificar qué usuarios y cuentas llamaron AWS, la dirección IP de origen desde la que se realizaron las llamadas y cuándo se produjeron. Para obtener más información, consulte la Guía del usuario de AWS CloudTrail.
- Amazon EventBridge ofrece una transmisión casi en tiempo real de los eventos del sistema que describen los cambios en sus dominios OpenSearch de servicio. Puede crear reglas que vigilen ciertos eventos y activen acciones automatizadas en otros Servicios de AWS cuando se produzcan estos eventos. Para obtener más información, consulta la Guía del EventBridge usuario de Amazon.

Supervisión OpenSearch sin servidor con Amazon CloudWatch

Puede supervisar Amazon OpenSearch Serverless mediante Amazon Serverless CloudWatch, que recopila datos sin procesar y los procesa para convertirlos en métricas legibles prácticamente en tiempo real. Estas estadísticas se mantienen durante 15 meses, de forma que pueda obtener acceso a información histórica y disponer de una mejor perspectiva sobre el desempeño de su aplicación web o servicio.

También puede establecer alarmas que vigilen determinados umbrales y enviar notificaciones o realizar acciones cuando se cumplan dichos umbrales. Para obtener más información, consulta la Guía del CloudWatch usuario de Amazon.

OpenSearch Serverless informa de las siguientes métricas en el `AWS/AOSS` espacio de nombres.

Métrica	Description (Descripción)
`ActiveCollection`	Indica si una colección está activa. Un valor de 1 significa que la colección está en estado `ACTIVE`. Este valor se emite al crear correctamente una colección y permanece como 1 hasta que esta se elimine. La métrica no puede tener un valor de 0. Estadísticas pertinentes: máximo Dimensiones: `ClientId`, `CollectionId` , `CollectionName`

Métrica	Description (Descripción)
	Frecuencia: 60 segundos
`DeletedDocuments`	El número total de documentos eliminados. Estadísticas pertinentes: promedio, suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frecuencia: 60 segundos
`IndexingOCU`	El número de unidades de OpenSearch cómputo (OCUs) utilizadas para ingerir los datos de la recopilación. Esta métrica se aplica según el nivel de la cuenta. Represent a el uso solo para las colecciones que no forman parte de ningún grupo de colecciones. Estadísticas pertinentes: suma Dimensiones: `ClientId` Frecuencia: 60 segundos
`IndexingOCU`	El número de unidades de OpenSearch cómputo (OCUs) utilizadas para ingerir los datos de la recopilación. Esta métrica se aplica a nivel de grupo de recopilación. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionGroupId` , `CollectionGroupName` Frecuencia: 60 segundos

Métrica	Description (Descripción)
`IngestionDataRate`	La velocidad de indexación en GiB por segundo de una colección o índice. Esta métrica se aplica únicamente a las solicitudes de indexación masiva. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frecuencia: 60 segundos
`IngestionDocumentErrors`	Número total de errores del documento durante el procesamiento de una colección o índice. Tras una solicitud de indexación masiva satisfactoria, los redactores procesan la solicitud y emiten errores para todos los documentos fallidos incluidos en la solicitud. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frecuencia: 60 segundos
`IngestionDocumentRate`	La velocidad por segundo a la que se incorporan los documentos a una colección o índice. Esta métrica se aplica únicamente a las solicitudes de indexación masiva. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frecuencia: 60 segundos

Métrica	Description (Descripción)
`IngestionRequestErrors`	El número total de errores en las solicitudes de indexació n masiva de una colección. OpenSearch Serverless emite esta métrica cuando una solicitud de indexación masiva falla por cualquier motivo, como un problema de autentica ción o disponibilidad. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` , `Collectio nName` Frecuencia: 60 segundos
`IngestionRequestLatency`	La latencia, en segundos, de las operaciones de escritura masiva en una colección. Estadísticas pertinentes: mínimo, máximo, promedio Dimensiones: `ClientId`, `CollectionId` , `Collectio nName` Frecuencia: 60 segundos
`IngestionRequestRate`	El número total de operaciones de escritura masiva recibidas por una colección. Estadísticas pertinentes: mínimo, máximo, promedio Dimensiones: `ClientId`, `CollectionId` , `Collectio nName` Frecuencia: 60 segundos

Métrica	Description (Descripción)
`IngestionRequestSuccess`	El número de operaciones de indexación de datos de una colección. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` Frecuencia: 60 segundos
`SearchableDocuments`	Número total de documentos susceptibles de búsqueda en una colección o índice. Estadísticas pertinentes: Suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frecuencia: 60 segundos
`SearchRequestErrors`	El número total de errores de consulta por minuto de una colección. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` Frecuencia: 60 segundos
`SearchRequestLatency`	El tiempo promedio, en milisegundos, que se tarda en completar una operación de búsqueda en una colección. Estadísticas pertinentes: mínimo, máximo, promedio Dimensiones: `ClientId`, `CollectionId` , `CollectionName` Frecuencia: 60 segundos

Métrica	Description (Descripción)
`Search0CU`	El número de unidades de OpenSearch cómputo (OCUs) utilizadas para buscar los datos de recopilación. Esta métrica se aplica según el nivel de la cuenta. Represent a el uso solo para las colecciones que no forman parte de ningún grupo de colecciones. Estadísticas pertinentes: suma Dimensiones: `ClientId` Frecuencia: 60 segundos
`Search0CU`	El número de unidades de OpenSearch cálculo (OCUs) utilizadas para buscar los datos de la recopilación. Esta métrica se aplica a nivel de grupo de recopilación. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionGroupId` , `CollectionGroupName` Frecuencia: 60 segundos
`SearchRequestRate`	El número total de solicitudes de búsqueda por minuto para una colección. Estadísticas pertinentes: promedio, máximo, suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` Frecuencia: 60 segundos

Métrica	Description (Descripción)
`StorageUsedInS3`	La cantidad, en bytes, de almacenamiento de Amazon S3 utilizada. OpenSearch La tecnología Serverless almacena los datos indexados en Amazon S3. Debe seleccionar el periodo en un minuto para obtener un valor preciso. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` , `IndexId`, `IndexName` Frecuencia: 60 segundos
`VectorIndexBuildAccelerationOCU`	El número de unidades de OpenSearch cómputo (OCUs) utilizadas para acelerar la indexación vectorial. Esta métrica se aplica a nivel de recopilación. Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` Frecuencia: 60 segundos
`2xx, 3xx, 4xx, 5xx`	El número de solicitudes para la colección que produjeron el código de respuesta HTTP especificado (2xx, 3xx, 4xx, 5xx). Estadísticas pertinentes: suma Dimensiones: `ClientId`, `CollectionId` , `CollectionName` Frecuencia: 60 segundos

Registrar las llamadas a la API OpenSearch sin servidor mediante AWS CloudTrail

Amazon OpenSearch Serverless está integrado con AWS CloudTrail un servicio que proporciona un registro de las acciones realizadas por un usuario, un rol o un AWS servicio en Serverless.

CloudTrail captura todas las llamadas a la API de OpenSearch Serverless como eventos. Las llamadas capturadas incluyen las llamadas de la sección Serverless de la consola de OpenSearch servicio y las llamadas en código a las operaciones de la API OpenSearch Serverless.

Si crea una ruta, puede habilitar la entrega continua de CloudTrail eventos a un bucket de Amazon S3, incluidos los eventos para OpenSearch Serverless. Si no configura una ruta, podrá ver los eventos más recientes en la CloudTrail consola, en el historial de eventos.

Con la información recopilada por CloudTrail, puede determinar la solicitud que se realizó a OpenSearch Serverless, la dirección IP desde la que se realizó la solicitud, quién la hizo, cuándo se realizó y detalles adicionales.

Para obtener más información CloudTrail, consulte la Guía del AWS CloudTrail usuario.

OpenSearch Información sobre sistemas sin servidor en CloudTrail

CloudTrail está habilitada en tu cuenta Cuenta de AWS al crear la cuenta. Cuando se produce una actividad en OpenSearch Serverless, esa actividad se registra en un CloudTrail evento junto con otros eventos de AWS servicio en el historial de eventos. Puede ver, buscar y descargar los eventos recientes en su Cuenta de AWS. Para obtener más información, consulte Visualización de eventos con el historial de CloudTrail eventos.

Para tener un registro continuo de tus eventos Cuenta de AWS, incluidos los eventos de OpenSearch Serverless, crea una ruta. Un rastro permite CloudTrail entregar archivos de registro a un bucket de Amazon S3. De forma predeterminada, cuando se crea un registro de seguimiento en la consola, el registro de seguimiento se aplica a todas las Regiones de AWS.

La ruta registra los eventos de todas las regiones de la AWS partición y envía los archivos de registro al bucket de Amazon S3 que especifique. Además, puede configurar otros AWS servicios para analizar más a fondo los datos de eventos recopilados en los CloudTrail registros y actuar en función de ellos. Para más información, consulte los siguientes temas:

- Introducción a la creación de registros de seguimiento

- CloudTrail servicios e integraciones compatibles
- Configuración de las notificaciones de Amazon SNS para CloudTrail
- Recibir archivos de CloudTrail registro de varias regiones y recibir archivos de CloudTrail registro de varias cuentas

Todas las acciones de OpenSearch Serverless se registran CloudTrail y se documentan en la referencia de la API OpenSearch Serverless. Por ejemplo, las llamadas a las `CreateCollection` `DeleteCollection` acciones y las llamadas generan entradas en los archivos de CloudTrail registro. `ListCollections`

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. La información de identidad lo ayuda a determinar:

- Si la solicitud se realizó con credenciales de usuario root o AWS Identity and Access Management (IAM).
- Si la solicitud se realizó con credenciales de seguridad temporales de un rol o fue un usuario federado.
- Si la solicitud la realizó otro AWS servicio.

Para obtener más información, consulte el elemento userIdentity de CloudTrail .

OpenSearch Eventos de datos sin servidor en CloudTrail

Los eventos de datos proporcionan información sobre las operaciones de recursos que se realizan en un recurso o dentro de él (por ejemplo, buscar o indexar en una colección OpenSearch sin servidor). Se denominan también operaciones del plano de datos. Los eventos de datos suelen ser actividades de gran volumen. De forma predeterminada, CloudTrail no registra los eventos de datos. El historial de CloudTrail eventos no registra los eventos de datos.

Se aplican cargos adicionales a los eventos de datos. Para obtener más información sobre CloudTrail los precios, consulta AWS CloudTrail Precios.

Puede registrar eventos de datos para los tipos de `AWS::AOSS::Collection` recursos mediante la CloudTrail consola o las operaciones de la CloudTrail API. AWS CLI Para obtener más información sobre cómo registrar los eventos de datos, consulte Registro de eventos de datos con la Consola de administración de AWS y Registro de eventos de datos con la AWS Command Line Interface en la Guía del usuario de AWS CloudTrail .

Puede configurar selectores de eventos avanzados para filtrar según los campos `eventName`, `readOnly` y `resources.ARN` y así registrar solo los eventos que son importantes para usted. Para obtener más información acerca de estos campos, consulte AdvancedFieldSelector en la Referencia de la API de AWS CloudTrail .

Descripción de las entradas de eventos de datos OpenSearch sin servidor

En el siguiente ejemplo:

- El campo `requestParameters` contiene detalles sobre la llamada a la API realizada a la colección. Incluye la ruta base de la petición (sin parámetros de consulta).
- El campo `responseElements` incluye un código de estado que indica el resultado de su petición al modificar recursos. Este código de estado lo ayuda a verificar si los cambios se procesaron correctamente o si requieren atención.
- OpenSearch Serverless registra los eventos CloudTrail de datos solo para las solicitudes que han completado correctamente la autenticación de 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"
    }
  ]
}
```

Descripción de las entradas de los eventos de OpenSearch administración sin servidor

Un rastro es una configuración que permite la entrega de eventos como archivos de registro a un bucket de Amazon S3 que usted especifique. CloudTrail Los archivos de registro contienen una o más entradas de registro.

Un evento representa una única solicitud desde cualquier origen. Incluye información sobre la acción solicitada, la fecha y la hora de la acción, los parámetros de la solicitud, etc. CloudTrail Los archivos de registro no son un registro ordenado de las llamadas a las API públicas, por lo que no aparecen en ningún orden específico.

En el siguiente ejemplo, se muestra una entrada de CloudTrail registro que demuestra la `CreateCollection` acción.


```
{
   "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"
   }
}
```


Supervisión de eventos OpenSearch sin servidor mediante Amazon EventBridge

Amazon OpenSearch Service se integra con Amazon EventBridge para notificarte determinados eventos que afectan a tus dominios. Los eventos de AWS los servicios se envían casi EventBridge en tiempo real. Los mismos eventos también se envían a Amazon CloudWatch Events, la predecesora de Amazon EventBridge. Puede crear reglas para indicar qué eventos le resultan de interés, así como qué acciones automatizadas se van a realizar cuando un evento cumple una de las reglas. Entre los ejemplos de acciones que puede activar automáticamente se incluyen los siguientes:

- Invocar una función AWS Lambda
- Invocar un Run Command de Amazon EC2
- Desviar el evento a Amazon Kinesis Data Streams
- Activación de una máquina de estados AWS Step Functions
- Notificar un tema de Amazon SNS o una cola de Amazon SQS

Para obtener más información, consulta Cómo empezar con Amazon EventBridge en la Guía del EventBridge usuario de Amazon.

Configuración de notificaciones

Puede utilizar las notificaciones AWS de usuario para recibir notificaciones cuando se produzca un evento OpenSearch sin servidor. Un evento es un indicador de un cambio en un entorno OpenSearch sin servidor, por ejemplo, cuando se alcanza el límite máximo de uso de la OCU. Amazon EventBridge recibe el evento y envía una notificación al Centro de Consola de administración de AWS notificaciones y a los canales de entrega que elijas. Recibirá una notificación cuando un evento coincida con una regla que especifique.

OpenSearch Eventos de unidades de cómputo (OCU)

OpenSearch Serverless envía los eventos EventBridge cuando se produce uno de los siguientes eventos relacionados con la OCU.

Uso de OCU cercano al límite máximo

OpenSearch Serverless envía este evento cuando el uso de la OCU de búsqueda o indexación alcanza el 75% del límite de capacidad. El uso de OCU se calcula en función del límite de capacidad configurado y del consumo actual de OCU.

Ejemplo

A continuación, se muestra un ejemplo de este evento (OCU de búsqueda):

```
{
  "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."
  }
```

```
}
```

A continuación, se muestra un ejemplo de este evento (OCU de indexación):

```
{
  "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."
  }
```

El uso de OCU ha alcanzado el límite máximo

OpenSearch Serverless envía este evento cuando el uso de la OCU de búsqueda o indexación alcanza el 100% del límite de capacidad. El uso de OCU se calcula en función del límite de capacidad configurado y del consumo actual de OCU.

Ejemplo

A continuación, se muestra un ejemplo de este evento (OCU de búsqueda):

```
{
  "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."
```

```
  }
}
```

A continuación, se muestra un ejemplo de este evento (OCU de indexación):

```
{
  "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."
  }
}
```

Creación y administración de dominios OpenSearch de Amazon Service

En este capítulo se describe cómo crear y gestionar los dominios OpenSearch de Amazon Service. Un dominio es el equivalente AWS aprovisionado de un clúster de código OpenSearch abierto. Al crear un dominio, especifica la configuración, los tipos de instancias, los recuentos de instancias y la asignación de almacenamiento. Para obtener más información sobre los clústeres de código abierto, consulte Crear un clúster en la OpenSearch documentación.

A diferencia de las breves instrucciones incluidas en el tutorial de introducción, en este capítulo se describen todas las opciones y se proporciona información de referencia relevante. Puede completar cada procedimiento siguiendo las instrucciones de la consola de OpenSearch servicio, el AWS Command Line Interface (AWS CLI) o el AWS SDKs.

Creación de dominios OpenSearch de servicio

En esta sección se describe cómo crear OpenSearch dominios de OpenSearch servicio mediante la consola de servicio o mediante el AWS CLI `create-domain` comando with.

Creación OpenSearch de dominios de servicio (consola)

Utilice el siguiente procedimiento para crear un dominio de OpenSearch servicio mediante la consola.

Para crear un dominio OpenSearch de servicio (consola)

1. Ve a Iniciar sesión en la consola https://aws.amazon.comy selecciona Iniciar sesión en ella.
2. En Analytics, selecciona Amazon OpenSearch Service.
3. Seleccione Create domain (Crear un dominio).
4. Para el Nombre de dominio, escriba un nombre de dominio. El nombre debe cumplir los siguientes criterios:

   - Único para su cuenta y Región de AWS
   - Comenzar por una letra minúscula.
   - Contener entre 3 y 28 caracteres.
   - Contiene únicamente letras minúsculas a-z, números de 0-9 y un guion (-).

5. Seleccione Creación estándar para el método de creación del dominio.
6. En Plantillas, seleccione la opción que mejor se ajuste a la finalidad del dominio:

   - Dominios de producción para cargas de trabajo que necesitan alta disponibilidad y rendimiento. Estos dominios utilizan Multi-AZ (con o sin modo de espera) y nodos principales dedicados para una mayor disponibilidad.
   - Desarrollo/pruebas para desarrollo o pruebas. Estos dominios pueden usar Multi-AZ (con o sin modo de espera) o una única zona de disponibilidad.

   > **Important**
   >
   > Los diversos tipos de implementación presentarán opciones distintas en las páginas siguientes. Estos pasos incluyen todas las opciones.

7. En cuanto a las Opciones de implementación, seleccione Dominio con modo de espera para configurar un dominio 3-AZ, con los nodos de una de las zonas reservados como don modo de espera. Esta opción aplica una serie de prácticas recomendadas, como un recuento de nodos de datos específico, un recuento de nodos maestros, un tipo de instancia, un recuento de réplicas y una configuración de actualización de software.
8. En Versión, elija la versión del OSS Elasticsearch anterior OpenSearch o la versión anterior que desee utilizar. Le recomendamos que elija la versión más reciente de. OpenSearch Para obtener más información, consulte the section called “Versiones compatibles de Elasticsearch y OpenSearch”.

   (Opcional) Si has elegido una OpenSearch versión para tu dominio, selecciona Habilitar el modo de compatibilidad para que su versión se OpenSearch muestre como 7.10, lo que permite que determinados clientes y complementos de Elasticsearch OSS que comprueben la versión antes de conectarse puedan seguir trabajando con el servicio.
9. En Instance type (Tipo de instancia), elija un tipo de instancia para los nodos de datos. Para obtener más información, consulte the section called “Tipos de instancias admitidas”.

   > **Note**
   >
   > No todas las zonas de disponibilidad admiten todos los tipos de instancias. Si elige Multi-AZ con o sin modo de espera, recomendamos elegir tipos de instancias de la generación actual, como R5 o I3.

10. Para el Número de nodos, elija el número de nodos de datos.

    Para ver los valores máximos, consulta Cuotas de dominios e OpenSearch instancias del servicio. Los clústeres de un solo nodo son adecuados para desarrollo y pruebas, pero no deben utilizarse para cargas de trabajo de producción. Para obtener más información, consulte the section called “Determinación del tamaño de dominios” y the section called “Configuración de un dominio multi-AZ”.

    > **Note**
    >
    > (Opcional) Los nodos coordinadores dedicados son compatibles con todas OpenSearch las versiones, ElasticSearch desde la 6.8 hasta la 7.10. Los nodos coordinadores dedicados están disponible para su uso con dominios que tienen un administrador de clústeres dedicado habilitado. Para habilitar los nodos coordinadores dedicados, seleccione el tipo y el recuento de instancias. Como práctica recomendada, debe mantener la familia de instancias de su nodo coordinador dedicado igual que la de sus nodos de datos (instancias basadas en Intel o instancias basadas en Graviton).

11. Para Tipo de almacenamiento, seleccione Amazon EBS. Los tipos de volumen disponibles en la lista dependen del tipo de instancia que haya elegido. Para obtener instrucciones sobre cómo crear dominios especialmente grandes, consulte the section called “Escala de petabytes”.
12. Para el almacenamiento EBS, configure los siguientes ajustes adicionales. Es posible que algunos ajustes no aparezcan en función del tipo de volumen que elija.

Opción	Description (Descripción)
Tipo de volumen de EBS	Elija entre Propósito general (SSD) - gp3 y Propósito general (SSD) - gp2, o la generación anterior IOPS aprovisionadas (SSD) y Magnético (estándar).
Tamaño de almacenamiento de EBS por nodo	Ingrese el tamaño del volumen de EBS que desea adjuntar a cada nodo de datos. Tamaño de volumen de EBS es por cada nodo. Puede calcular el tamaño total del clúster para el dominio de OpenSearch servicio multiplicando el número de nodos de datos por el tamaño del volumen de EBS. El tamaño mínimo y máximo de un volumen de EBS depende tanto del tipo de volumen de EBS

Opción	Description (Descripción)
	especificado como del tipo de instancia al que se ha adjuntado . Para obtener más información, consulte las restricciones de tamaño del volumen de EBS.
Provisioned IOPS (IOPS aprovisionadas)	Si ha seleccionado un tipo de volumen SSD de IOPS aprovisionadas, introduzca el número de I/O operaciones por segundo (IOPS) que puede soportar el volumen.

13. (Opcional) Si seleccionó un tipo de gp3 volumen, amplíe la configuración avanzada y especifique IOPS (hasta 16 000 por cada tamaño de volumen de 3 TiB aprovisionado por nodo de datos) y rendimiento (hasta 1000 por MiB/s cada tamaño de volumen de 3 TiB aprovisionado por nodo de datos) adicionales a los incluidos en el precio del almacenamiento, por un costo adicional. Para obtener más información, consulta los precios OpenSearch de Amazon Service.
14. (Opcional) Para habilitar el UltraWarm almacenamiento, selecciona Habilitar nodos UltraWarm de datos. Cada tipo de instancia tiene una cantidad máxima de almacenamiento que puede poner a disposición. Multiplique esa cantidad por el número de nodos de datos templados para obtener el almacenamiento templado total a disposición.
15. (Opcional) Para habilitar el almacenamiento en frío, seleccione Habilitar almacenamiento en frío. Debe habilitarlo UltraWarm para habilitar el almacenamiento en frío.
16. Si utiliza Multi-AZ con modo de espera, ya hay tres nodos maestros dedicados habilitados. Seleccione el tipo de nodos maestros que desee. Si eligió un dominio Multi-AZ sin modo de espera, seleccione Habilitar nodos maestros dedicados y elija el tipo y la cantidad de nodos maestros que desee. Los nodos maestros dedicados aumentan la estabilidad de los clústeres y son necesarios en los dominios que tienen un número de instancias mayor de 10. Recomendamos utilizar tres nodos maestros dedicados en los dominios de producción.

> **Note**
>
> Puede elegir tipos de instancias diferentes para los nodos maestros dedicados y los nodos de datos. Por ejemplo, para los nodos de datos, puede elegir instancias de uso general u optimizadas para el almacenamiento y, para los nodos maestros dedicados, instancias optimizadas para computación.

17. (Opcional) En el caso de los dominios que ejecutan Elasticsearch 5.3 OpenSearch o versiones posteriores, la configuración de Snapshot es irrelevante. Para obtener más información sobre las instantáneas automatizadas, consulte the section called “Crear instantáneas de índice”.
18. Si desea utilizar un punto de conexión personalizado en lugar del estándar de `https://search-mydomain-1a2a3a4a5a6a7a8a9a0a9a8a7a.us-east-1.es.amazonaws.com`, seleccione Habilitar punto de enlace personalizado y proporcione un nombre y un certificado. Para obtener más información, consulte the section called “Creación de un punto de conexión personalizado”.
19. En Red, seleccione Acceso a la VPC o Acceso público. Si elige Acceso público, vaya al siguiente paso. Si elige Acceso a la VPC, asegúrese de que se cumplen los requisitos previos y, a continuación, haga lo siguiente:

Opción	Description (Descripción)
VPC	Seleccione el ID de la nube privada virtual (VPC) que desee utilizar. La VPC y el dominio deben estar en el mismo Región de AWS lugar y debe seleccionar una VPC con la tenencia configurada como Predeterminada. OpenSearch El servicio aún no admite el uso de un VPCs arrendamiento dedicado.
Subred	Seleccione una subred. Si habilitó Multi-AZ, debe elegir dos o tres subredes. OpenSearch El servicio colocará un punto final de VPC e interfaces de red elásticas en las subredes. Debe reservar suficientes direcciones IP para las interfaces de red en la subred o subredes. Para obtener más información, consulte Reserva de direcciones IP en una subred de una VPC.
Grupos de seguridad	Elija uno o más grupos de seguridad de VPC que permitan que la aplicación requerida llegue al dominio del OpenSearch servicio en los puertos (80 o 443) y protocolos (HTTP o HTTPS) expuestos por el dominio. Para obtener más información, consulte the section called “Compatibilidad con VPC”.
IAM Role	Mantenga el rol predeterminado. OpenSearch El servicio usa esta función predefinida (también conocida como función vinculada a un servicio) para acceder a la VPC y colocar un punto final de la VPC y las

Opción	Description (Descripción)
	interfaces de red en la subred de la VPC. Para obtener más información, consulte Roles vinculados a servicios para el acceso de VPC.
Tipo de dirección IP	Elija una pila doble o un tipo de dirección IP IPv4 . La pila doble le permite compartir los recursos del dominio IPv4 y los tipos de IPv6 direcciones, y es la opción recomendada. Si configura el tipo de dirección IP en doble pila, no podrá cambiarlo más adelante.

20. Habilite o deshabilite el control de acceso detallado:

    - Si desea utilizar IAM para la gestión de usuarios, seleccione Establecer ARN de IAM como usuario maestro y especifique el ARN para un rol de IAM.
    - Si desea utilizar la base de datos de usuarios interna, seleccione Crear usuario maestro y especifique un nombre de usuario y una contraseña.

    Sea cual sea la opción que elija, el usuario maestro puede acceder a todos los índices del clúster y a todos OpenSearch APIs ellos. Para obtener información sobre qué opción elegir, consulte the section called “Conceptos clave”.

    Si deshabilita el control de acceso detallado, puede controlar el acceso a su dominio al colocarlo dentro de una VPC, al aplicar una política de acceso restrictiva o puede utilizar ambas opciones. Debe habilitar el node-to-node cifrado y el cifrado en reposo para utilizar un control de acceso detallado.

    > **Note**
    >
    > Se recomienda encarecidamente habilitar un control de acceso detallado para proteger los datos del dominio. El control de acceso detallado proporciona seguridad en los niveles de clúster, índice, documento y campo.

21. (Opcional) Si quiere usar la autenticación SAML para los OpenSearch paneles, elija Habilitar la autenticación SAML y configure las opciones de SAML para el dominio. Para obtener instrucciones, consulte the section called “Autenticación SAML para paneles OpenSearch”.
22. (Opcional) Si desea utilizar la autenticación de Amazon Cognito para los OpenSearch paneles, elija Activar la autenticación de Amazon Cognito. A continuación, elija el grupo de usuarios y el grupo de identidades de Amazon Cognito que desee utilizar para la autenticación de

OpenSearch Dashboards. Para obtener orientación sobre cómo crear estos recursos, consulte the section called “Autenticación de Amazon Cognito para paneles OpenSearch ”.

23. (Opcional) Si desea utilizar la autenticación del Centro de Identidad de IAM (IDC) para conectar su fuente de identidad existente y ofrecer a sus AWS aplicaciones una vista común de sus usuarios, elija Habilitar el acceso a la API autenticado con el Centro de Identidad de IAM. Para obtener más información, consulte Descripción general de la propagación de identidades confiables en la Guía del usuario de Centro de identidades de IAM.
24. (Opcional) En la sección Característica avanzadas, deje seleccionadas Habilitar generación de consultas en lenguaje natural y características de Amazon Q Developer, si desea utilizar estas funciones.

    a. Seleccione Activar S3 Vectors como opción de motor para mejorar las opciones de búsqueda vectorial. Para obtener más información, consulte Capacidades de búsqueda avanzada con un motor vectorial Amazon S3.

    b. Seleccione Activar la aceleración de la GPU para obtener opciones de búsqueda vectoriales mejoradas. Para obtener más información, consulte Aceleración de GPU para indexación vectorial.

25. En Política de acceso, seleccione una política de acceso o configure una de las que posee. Si decide crear una política personalizada, puede configurarla usted mismo o importar una de otro dominio. Para obtener más información, consulte the section called “Gestión de identidad y acceso”.

    > **Note**
    >
    > Si habilitó el acceso de la VPC, no puede utilizar políticas basadas en IP. En su lugar, puede utilizar grupos de seguridad para controlar qué direcciones IP pueden tener acceso al dominio. Para obtener más información, consulte the section called “Acerca de las políticas de acceso en los dominios de VPC”.

26. (Opcional) Para determinar que todas las solicitudes al dominio lleguen a través de HTTPS, seleccione Require HTTPS for all traffic to the domain (Solicitar HTTPS para todo el tráfico del dominio). Para habilitar el node-to-node cifrado, selecciona el ode-to-node cifrado N. Para obtener más información, consulte the section called “Node-to-node cifrado”. Para habilitar el cifrado de los datos en reposo, seleccione Habilitar cifrado de datos en reposo. Estas opciones están preseleccionadas si elige la opción de implementación de Multi-AZ con modo de espera.

27. (Opcional) Seleccione AWS Usar una clave propia para que el OpenSearch Servicio cree una clave de AWS KMS cifrado en su nombre (o use la que ya creó). De lo contrario, elija su propia clave de KMS. Para obtener más información, consulte the section called “Cifrado en reposo”.
28. En los períodos de menor actividad, selecciona una hora de inicio para programar las actualizaciones del software del servicio y las optimizaciones de Auto-Tune que requieran una implementación. blue/green Las actualizaciones en ventanas de menor actividad ayudan a minimizar la carga de los nodos maestros dedicados de un clúster durante los períodos de alto tráfico.
29. En el caso de Auto-Tune, decide si quieres que el OpenSearch Servicio sugiera cambios de configuración relacionados con la memoria en tu dominio para mejorar la velocidad y la estabilidad. Para obtener más información, consulte the section called “Ajuste automático”.

    (Opcional) Seleccione Ventana de menor actividad para programar una ventana periódica para que el ajuste automático actualice el dominio.
30. (Opcional) Seleccione Actualización automática de software para activar las actualizaciones automáticas de software.
31. (Opcional) Agregue etiquetas para describir el dominio para que pueda categorizar y filtrar esa información. Para obtener más información, consulte the section called “Etiquetado de dominios”.
32. (Opcional) Expanda y configure Advanced cluster settings (Configuración avanzada del clúster). Para obtener un resumen de estas opciones, consulte the section called “Configuración avanzada de clústeres”.
33. Seleccione Crear.

Crear dominios de OpenSearch servicio ()AWS CLI

En lugar de crear un dominio de OpenSearch servicio mediante la consola, puede utilizar la AWS CLI. Para conocer la sintaxis, consulte Amazon OpenSearch Service en la referencia de comandos de la AWS CLI a.

Comandos de ejemplo

En este primer ejemplo, se muestra la siguiente configuración del dominio de OpenSearch servicio:

- Crea un dominio OpenSearch de servicio denominado mylogs con la OpenSearch versión 1.2
- Rellena el dominio con dos instancias del tipo `r6g.large.search`.

- Utiliza un volumen de EBS de uso general (SSD) `gp3` de 100 GiB para el almacenamiento de cada nodo de datos
- Permite el acceso anónimo, pero solo desde una única dirección 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"]}}}]}'
```

El siguiente ejemplo muestra la siguiente configuración del dominio OpenSearch de servicio:

- Crea un dominio OpenSearch de servicio denominado mylogs con la versión 7.10 de Elasticsearch
- Rellena el dominio con seis instancias del tipo `r6g.large.search`.
- Utiliza un volumen de EBS de uso general (SSD) `gp2` de 100 GiB para el almacenamiento de cada nodo de datos
- Restringe el acceso al servicio a un solo usuario, identificado por el ID del usuario: 555555555555 Cuenta de AWS
- Distribuye las instancias entre tres zonas de disponibilidad.

```
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/*" } ] }'
```

El siguiente ejemplo muestra la siguiente configuración del dominio de OpenSearch servicio:

- Crea un dominio OpenSearch de servicio denominado mylogs con la OpenSearch versión 1.0

- Rellena el dominio con diez instancias del tipo `r6g.xlarge.search`.
- Rellena el dominio con tres instancias del tipo `r6g.large.search` que actuarán como nodos maestros dedicados.
- Usa un volumen de EBS de IOPS aprovisionadas de 100 GiB para el almacenamiento, configurado con una velocidad de referencia de 1000 IOPS por cada nodo de datos
- Restringe el acceso a un único usuario y a un único subrecurso, la 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**
>
> Si intenta crear un dominio de OpenSearch servicio y ya existe un dominio con el mismo nombre, la CLI no informa de ningún error. En su lugar, muestra los detalles sobre el dominio existente.

Creación OpenSearch de dominios de servicio (AWS SDKs)

Las AWS SDKs (excepto Android e iOS SDKs) admiten todas las acciones definidas en la referencia de la API de Amazon OpenSearch Service, incluidas`CreateDomain`. Para ver el código de muestra, consulte the section called “Usando el AWS SDKs”. Para obtener más información sobre la instalación y el uso de AWS SDKs, consulte los kits de desarrollo de AWS software.

Creación OpenSearch de dominios de servicio (AWS CloudFormation)

OpenSearch El servicio está integrado con AWS CloudFormation un servicio que le ayuda a modelar y configurar sus AWS recursos para que pueda dedicar menos tiempo a crear y administrar sus recursos e infraestructura. Crea una plantilla que describe el OpenSearch dominio que desea crear

y CloudFormation aprovisiona y configura el dominio por usted. Para obtener más información, incluidos ejemplos de plantillas JSON y YAML para OpenSearch dominios, consulta la referencia de tipos de recursos de Amazon OpenSearch Service en la Guía del AWS CloudFormation usuario.

Configurar políticas de acceso

Amazon OpenSearch Service ofrece varias formas de configurar el acceso a tus dominios OpenSearch de servicio. Para obtener más información, consulte the section called “Gestión de identidad y acceso” y the section called “Control de acceso detallado”.

La consola proporciona políticas de acceso preconfiguradas que puede personalizar según las necesidades específicas de su dominio. También puede importar políticas de acceso de otros dominios OpenSearch de servicio. Para obtener información sobre cómo interactúan estas políticas de acceso con el acceso de las VPC, consulte the section called “Acerca de las políticas de acceso en los dominios de VPC”.

Para configurar políticas de acceso (consola)

1. Vaya a https://aws.amazon.com, y seleccione Sign In to the Console (Inicie sesión en la consola).
2. En Analytics, selecciona Amazon OpenSearch Service.
3. En el panel de navegación, en Domains (Dominios), elija el dominio que desea actualizar.
4. Seleccione Acciones y Editar la configuración de seguridad.
5. Edite la política de acceso JSON o importe una opción preconfigurada.
6. Seleccione Save changes (Guardar cambios).

Migración al OpenSearch servicio con Migration Assistant

Migration Assistant for Amazon OpenSearch Service es una solución integral que simplifica el proceso de migración desde clústeres o Elasticsearch autogestionados a Service. OpenSearch OpenSearch Este conjunto de herramientas aborda la complejidad operativa de las migraciones, al tiempo que garantiza la integridad de los datos y valida el rendimiento después de la migración.

Descripción general de

Ya sea que esté configurando una proof-of-concept carga de trabajo de producción AWS, realizando la transición o actualizándola a las OpenSearch versiones más recientes, Migration Assistant

proporciona step-by-step orientación, mejores prácticas y herramientas para aprovechar todo el potencial de las migraciones. OpenSearch

El Asistente de Migración ofrece los siguientes beneficios clave:

- Migración de metadatos: migra los metadatos del clúster, incluidos la configuración de índices, los mapeos de tipos, las plantillas de índices y los alias.
- Migración de datos: migra los datos existentes de los clústeres heredados a los dominios de servicio OpenSearch
- Gestión del tráfico en tiempo real: intercepta y redirige el tráfico en tiempo real de los clústeres autogestionados a los dominios de OpenSearch servicio con una latencia mínima
- Replicación de tráfico: replica el tráfico de producción en los clústeres de destino para validar la precisión y el rendimiento.
- Pruebas de rendimiento: simula tráfico del mundo real capturando y reproduciendo patrones de solicitudes para refinar el rendimiento del sistema.
- Disponibilidad global: se despliega en las AWS regiones más comunes para lograr un alcance y una escalabilidad globales

El Asistente de Migración admite la migración desde las versiones 6.x y 7.x, y 1.x y 2.x de Elasticsearch. OpenSearch Para obtener más información, consulte Rutas de migración compatibles.

> **Note**
>
> Tenga en cuenta la siguiente información adicional sobre el Asistente de Migración.
>
> - La herramienta admite migraciones con varios saltos (por ejemplo, la migración de Elasticsearch 5.x a Service 3.x en un solo salto). OpenSearch
> - Puede revertir una migración.
> - Para algunos casos de uso, la herramienta requiere poco o ningún tiempo de inactividad.
> - La herramienta proporciona relleno de datos de alto rendimiento sin afectar al clúster de origen.

Escenarios de migración

El Asistente de Migración está diseñado para manejar los siguientes escenarios de migración:

Migración de metadatos

Migra los metadatos del clúster, como la configuración de los índices, los alias y las plantillas, del clúster de origen al dominio de servicio de destino. OpenSearch

Migración de relleno de datos

Migra los datos existentes o históricos de un clúster de origen a un dominio de OpenSearch servicio de destino, lo que garantiza que todos los datos importantes se conserven durante la transición.

Migración de tráfico en vivo

Replica el tráfico continuo en tiempo real desde el clúster de origen al dominio de OpenSearch servicio de destino, lo que le permite mantener la disponibilidad del servicio durante la migración.

> **Important**
>
> Las estrategias de migración no son universalmente aplicables. El Asistente de Migración proporciona orientación basada en las mejores prácticas de ingeniería, pero usted debe evaluar sus requisitos específicos y realizar pruebas exhaustivas antes de migrar cargas de trabajo de producción.

Introducción del Asistente de Migración

El Asistente de Migración para Amazon OpenSearch Service está disponible como una AWS solución con documentación completa, plantillas de implementación y código fuente. Para comenzar con el Asistente de Migración:

Para acceder a los recursos del Asistente de Migración:

1. Consulta la documentación completa de la solución en Descripción general de la solución de Migration Assistant for Amazon OpenSearch Service.
2. Comprenda los costos y requisitos leyendo la guía de planificación de la implementación.
3. Despliegue la solución siguiendo las instrucciones de implementación y utilice la guía de uso para realizar su migración.

Para desarrolladores y usuarios avanzados, el código fuente y la documentación adicional de Migration Assistant están disponibles en el GitHub repositorio de OpenSearch migraciones.

Arquitectura de soluciones

Cuando se implementa en AWS, Migration Assistant utiliza varios AWS servicios para ofrecer una solución de migración integral:

- AWS CloudFormation: proporciona plantillas de Infraestructura como Código (IaC) para implementar y configurar el Asistente de Migración.
- OpenSearch Servicio: el servicio de destino para sus cargas de trabajo migradas de búsqueda y análisis
- Amazon Managed Streaming para Apache Kafka: proporciona capacidades de procesamiento de flujo para almacenamiento duradero y reutilización del tráfico HTTP.
- Amazon Elastic Container Service: ejecuta la Consola de Administración de Migraciones y el Reproductor de Tráfico en contenedores seguros y escalables.
- Amazon Elastic File System: proporciona almacenamiento persistente y escalable para los datos de solicitudes y respuestas de los clústeres de origen y destino.
- Amazon Simple Storage Service: almacena instantáneas para tareas de relleno histórico y contenido de Infraestructura como Código.

Para obtener información detallada sobre la arquitectura, consulte Detalles de la arquitectura del Asistente de Migración.

Configuración avanzada de clústeres

Use opciones avanzadas para configurar lo siguiente:

Índices en los cuerpos de las solicitudes

Especifica si se permiten referencias explícitas a índices en el cuerpo de las solicitudes HTTP. Al configurar esta propiedad en `false` se evita que los usuarios eludan el control de acceso para subrecursos. De forma predeterminada, el valor es `true`. Para obtener más información, consulte the section called “Opciones avanzadas y consideraciones de la API”.

Asignación de una caché de datos de campos

Especifica el porcentaje de espacio de montón de Java que se asignará a datos de campos. De forma predeterminada, esta configuración es el 20 % de la pila de JVM.

> **Note**
>
> Muchos clientes consultan índices diarios en rotación. Recomendamos empezar a realizar ensayos de referencia con el valor `indices.fielddata.cache.size` configurado en un 40 % de la pila de JVM para la mayoría de estos casos de uso. Si tiene índices muy grandes es posible que necesite una caché de datos de campos de gran tamaño.

Recuento máximo de cláusulas

Especifica el número máximo de cláusulas permitidas en una consulta de Lucene con operadores booleanos. El valor predeterminado es 1,024. Las consultas que superen el número máximo permitido de cláusulas provocarán un error `TooManyClauses`. Para obtener más información, consulte la documentación de Lucene.

Enriquecimiento semántico automático para Amazon Service OpenSearch

Introducción

Amazon OpenSearch Service utiliza la word-to-word concordancia (búsqueda léxica) para encontrar resultados, de forma similar a otros motores de búsqueda tradicionales. Este enfoque funciona bien para consultas específicas, como códigos de productos o números de modelo, pero tiene dificultades con las búsquedas abstractas, en las que es crucial entender la intención del usuario. Por ejemplo, si buscas «zapatos para la playa», la búsqueda léxica encuentra las palabras individuales «zapatos», «playa», «para» y «el» en los artículos del catálogo, por lo que podrían omitir productos relevantes como «sandalias resistentes al agua» o «calzado de surf» que no contienen los términos de búsqueda exactos.

El enriquecimiento semántico automático resuelve esta limitación al tener en cuenta tanto las coincidencias de palabras clave como el significado contextual de las búsquedas. Esta función comprende la intención de búsqueda y mejora la relevancia de la búsqueda hasta en un 20%. Activa esta función en los campos de texto de tu índice para mejorar los resultados de búsqueda.

> **Note**
>
> El enriquecimiento semántico automático está disponible para los dominios de OpenSearch servicio que ejecutan la versión 2.19 o posterior. Además, los dominios con la OpenSearch versión 2.19 también deben incluir la última actualización de la versión del software de servicio. Actualmente, la función está disponible para los dominios públicos y no se admiten los dominios de VPC.

Detalles del modelo y punto de referencia de rendimiento

Si bien esta función resuelve las complejidades técnicas entre bastidores sin exponer el modelo subyacente, ofrecemos transparencia mediante una breve descripción del modelo y resultados comparativos para ayudarlo a tomar decisiones informadas sobre la adopción de funciones en sus cargas de trabajo críticas.

El enriquecimiento semántico automático utiliza un modelo disperso prediseñado y gestionado por el servicio que funciona de forma eficaz sin necesidad de ajustes personalizados. El modelo analiza los campos que especificas y los expande en vectores dispersos en función de las asociaciones aprendidas a partir de diversos datos de entrenamiento. Los términos ampliados y sus ponderaciones de significancia se almacenan en un formato de índice nativo de Lucene para una recuperación eficiente. Hemos optimizado este proceso mediante el modo de solo documentos, en el que la codificación solo se realiza durante la ingesta de datos. Las consultas de búsqueda simplemente se tokenizan en lugar de procesarse mediante un modelo disperso, lo que hace que la solución sea rentable y eficaz.

Nuestra validación del rendimiento durante el desarrollo de las funciones utilizó el conjunto de datos de recuperación de pasajes de MS MARCO, que incluía pasajes con un promedio de 334 caracteres. Para obtener una puntuación de relevancia, medimos la ganancia acumulada descontada normalizada (NDCG) de los 10 primeros resultados de búsqueda (ndcg @10) según el índice BEIR para contenido en inglés y la media ndcg @10 según MIRACL para contenido multilingüe. Evaluamos la latencia mediante mediciones del percentil 90 (p90) del lado del cliente, y la respuesta de búsqueda p90 tomó valores. Estos puntos de referencia proporcionan indicadores de rendimiento de referencia tanto para la relevancia de la búsqueda como para los tiempos de respuesta. Estas son las principales cifras de referencia:

- Idioma inglés: mejora la relevancia del 20% con respecto a la búsqueda léxica. También redujo la latencia de búsqueda del P90 en un 7,7% con respecto a la búsqueda léxica (BM25 es de 26 ms y el enriquecimiento semántico automático es de 24 ms).
- Multilingüe: la relevancia mejoró un 105% con respecto a la búsqueda léxica, mientras que la latencia de la búsqueda con el P90 aumentó un 38,4% con respecto a la búsqueda léxica (BM25 es de 26 ms y el enriquecimiento semántico automático es de 36 ms).

Dada la naturaleza única de cada carga de trabajo, le recomendamos que evalúe esta función en su entorno de desarrollo utilizando sus propios criterios de evaluación comparativa antes de tomar decisiones de implementación.

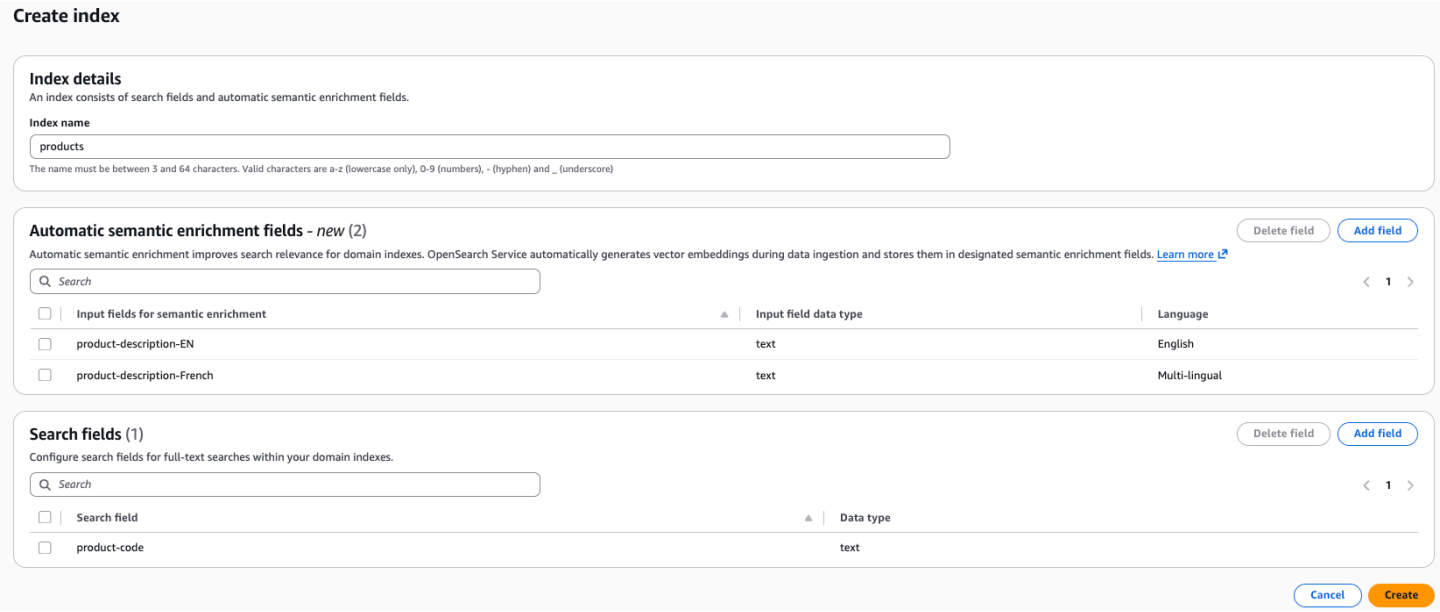
Idiomas compatibles

La función es compatible con el inglés. Además, el modelo también admite árabe, bengalí, chino, finés, francés, hindi, indonesio, japonés, coreano, persa, ruso, español, swahili y telugu.

Configure un índice de enriquecimiento semántico automático para los dominios

Configurar un índice con el enriquecimiento semántico automático activado para sus campos de texto es fácil y puede gestionarlo mediante la consola y las CloudFormation plantillas durante la creación de un nuevo índice. APIs Para habilitarlo en un índice existente, debe volver a crear el índice con el enriquecimiento semántico automático activado para los campos de texto.

Experiencia de consola: la AWS consola le permite crear fácilmente un índice con campos de enriquecimiento semántico automático. Cuando selecciones un dominio, encontrarás el botón de creación de índice en la parte superior de la consola. Al hacer clic en el botón de creación de índice, encontrará opciones para definir campos de enriquecimiento semántico automático. En un índice, puede tener combinaciones de enriquecimiento semántico automático para campos en inglés y multilingüe, así como para campos léxicos.



Experiencia con las API: para crear un índice de enriquecimiento semántico automático mediante la interfaz de línea de comandos (AWS CLI), utilice el AWS comando create-index:

```
aws opensearch create-index \
--domain-name [domain_name] \
--index-name [index_name] \
--index-schema [index_body] \
```

En el siguiente ejemplo de esquema de índice, el campo title_semantic tiene un tipo de campo establecido en texto y el parámetro semantic_enrichment establecido en el estado ENABLED. Si se establece el parámetro semantic_enrichment, se habilita el enriquecimiento semántico automático en el campo title_semantic. Puede usar el campo language_options para especificar el inglés o el multilingüe.

```
    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"
            }
        }
    }
}'
```

Para describir el índice creado, utilice el siguiente comando:

```
aws opensearch get-index \
--domain-name [domain_name] \
--index-name [index_name] \
```

Ingestión y búsqueda de datos

Una vez que haya creado un índice con el enriquecimiento semántico automático activado, la función funcionará automáticamente durante el proceso de ingesta de datos, sin necesidad de realizar ninguna configuración adicional.

Ingesta de datos: al añadir documentos al índice, el sistema automáticamente:

- Analiza los campos de texto que designó para el enriquecimiento semántico
- Genera codificaciones semánticas mediante OpenSearch un modelo disperso gestionado por el servicio
- Almacena estas representaciones enriquecidas junto con los datos originales

Este proceso utiliza los conectores OpenSearch de aprendizaje automático integrados y las canalizaciones de ingesta, que se crean y administran automáticamente entre bastidores.

Búsqueda: los datos de enriquecimiento semántico ya están indexados, por lo que las consultas se ejecutan de manera eficiente sin tener que volver a invocar el modelo de aprendizaje automático.

Esto significa que obtiene una mayor relevancia de búsqueda sin sobrecargar la latencia de búsqueda adicional.

Configuración de permisos para el enriquecimiento semántico automático

Antes de crear un índice con enriquecimiento semántico automático, debe configurar los permisos necesarios. En esta sección se explican los permisos necesarios para las distintas operaciones de indexación y cómo configurarlos tanto para escenarios AWS Identity and Access Management (IAM) como para escenarios de control de acceso detallados.

Permisos de IAM

Los siguientes permisos de IAM son necesarios para las operaciones de enriquecimiento semántico automático. Estos permisos varían en función de la operación de indexación específica que desee realizar.

CreateIndex Permisos de API

Para crear un índice con enriquecimiento semántico automático, necesita los siguientes permisos de IAM:

- `es:CreateIndex`— Cree un índice con capacidades de enriquecimiento semántico.
- `es:ESHttpHead`— Realice solicitudes HEAD para comprobar la existencia del índice.
- `es:ESHttpPut`— Realizar solicitudes PUT para la creación de índices.
- `es:ESHttpPost`— Realizar solicitudes POST para operaciones de indexación.

UpdateIndex Permisos de API

Para actualizar un índice existente con un enriquecimiento semántico automático, necesita los siguientes permisos de IAM:

- `es:UpdateIndex`— Actualizar la configuración y las asignaciones de los índices.
- `es:ESHttpPut`— Realizar solicitudes PUT para actualizaciones de índices.
- `es:ESHttpGet`— Realizar solicitudes GET para recuperar la información del índice.
- `es:ESHttpPost`— Realizar solicitudes POST para operaciones de indexación.

GetIndex Permisos de API

Para recuperar información sobre un índice con enriquecimiento semántico automático, necesita los siguientes permisos de IAM:

- `es:GetIndex`— Recuperar la información y la configuración del índice.
- `es:ESHttpGet`— Realizar solicitudes GET para recuperar los datos del índice.

DeleteIndex Permisos de API

Para eliminar un índice con enriquecimiento semántico automático, necesita los siguientes permisos de IAM:

- `es:DeleteIndex`— Eliminar un índice y sus componentes de enriquecimiento semántico.
- `es:ESHttpDelete`— Realizar solicitudes de ELIMINACIÓN para eliminar el índice.

Ejemplo de política de IAM

El siguiente ejemplo de política de acceso basada en la identidad proporciona los permisos necesarios para que un usuario gestione los índices con un enriquecimiento semántico automático:

```
{
    "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"
        }
```

```
    ]
}
```

Sustituya *aws-region* y por *111122223333* sus valores específicos. *domain-name* Puede restringir aún más el acceso especificando patrones de índice específicos en el ARN del recurso.

Permisos de control de acceso específicos

Si tu dominio de Amazon OpenSearch Service tiene activado un control de acceso detallado, necesitarás permisos adicionales además de los permisos de IAM. Se requieren los siguientes permisos para cada operación de indexación.

CreateIndex Permisos de API

Cuando se habilita el control de acceso detallado, se requieren los siguientes permisos adicionales para crear un índice con enriquecimiento semántico automático:

- `indices:admin/create`— Crear operaciones de indexación.
- `indices:admin/mapping/put`— Crear y actualizar mapeos de índices.
- `cluster:admin/opensearch/ml/create_connector`— Cree conectores de aprendizaje automático para el procesamiento semántico.
- `cluster:admin/opensearch/ml/register_model`— Registre modelos de aprendizaje automático para el enriquecimiento semántico.
- `cluster:admin/ingest/pipeline/put`— Cree canales de ingesta para el procesamiento de datos.
- `cluster:admin/search/pipeline/put`— Crear canales de búsqueda para el procesamiento de consultas.

UpdateIndex Permisos de API

Cuando se habilita el control de acceso detallado, se requieren los siguientes permisos adicionales para actualizar un índice con el enriquecimiento semántico automático:

- `indices:admin/get`— Recuperar información del índice.
- `indices:admin/settings/update`— Actualizar la configuración del índice.
- `indices:admin/mapping/put`— Actualizar las asignaciones de índices.

- `cluster:admin/opensearch/ml/create_connector`— Crear conectores de aprendizaje automático.
- `cluster:admin/opensearch/ml/register_model`— Registrar modelos de aprendizaje automático.
- `cluster:admin/ingest/pipeline/put`— Crea canalizaciones de ingesta.
- `cluster:admin/search/pipeline/put`— Crear canalizaciones de búsqueda.
- `cluster:admin/ingest/pipeline/get`— Recuperar la información de las canalizaciones de ingesta.
- `cluster:admin/search/pipeline/get`— Recuperar la información de la canalización de búsqueda.

GetIndex Permisos de API

Cuando se habilita el control de acceso detallado, se requieren los siguientes permisos adicionales para recuperar información sobre un índice con un enriquecimiento semántico automático:

- `indices:admin/get`— Recuperar información del índice.
- `cluster:admin/ingest/pipeline/get`— Recuperar la información de la canalización de ingesta.
- `cluster:admin/search/pipeline/get`— Recuperar la información de la canalización de búsqueda.

DeleteIndex Permisos de API

Cuando se habilita el control de acceso detallado, se requiere el siguiente permiso adicional para eliminar un índice con el enriquecimiento semántico automático:

- `indices:admin/delete`— Eliminar las operaciones de indexación.

Reescrituras de consultas

El enriquecimiento semántico automático convierte automáticamente las consultas «coincidentes» existentes en consultas de búsqueda semántica sin necesidad de modificar las consultas. Si una consulta de coincidencia forma parte de una consulta compuesta, el sistema recorre la estructura de la consulta, busca las consultas coincidentes y las reemplaza por consultas de dispersión

neuronal. Actualmente, la función solo permite reemplazar las consultas «coincidentes», tanto si se trata de una consulta independiente como si forma parte de una consulta compuesta. No se admite la opción «multi_match». Además, la función admite todas las consultas compuestas para reemplazar sus consultas de coincidencia anidadas. Las consultas compuestas incluyen: bool, boost, constant_score, dis_max, function_score e hybrid.

Limitaciones del enriquecimiento semántico automático

La búsqueda semántica automática es más eficaz cuando se aplica a campos small-to-medium dimensionados que contienen contenido en lenguaje natural, como títulos de películas, descripciones de productos, reseñas y resúmenes. Si bien la búsqueda semántica aumenta la relevancia en la mayoría de los casos de uso, puede que no sea óptima en algunos escenarios. Ten en cuenta las siguientes limitaciones a la hora de decidir si vas a implementar el enriquecimiento semántico automático para tu caso de uso específico.

- Documentos muy largos: el modelo disperso actual procesa solo los primeros 8.192 símbolos de cada documento en inglés. En el caso de los documentos multilingües, se trata de 512 fichas. En el caso de artículos extensos, considere la posibilidad de implementar la fragmentación de documentos para garantizar un procesamiento completo del contenido.
- Cargas de trabajo de análisis de registros: el enriquecimiento semántico aumenta significativamente el tamaño del índice, lo que puede resultar innecesario en el análisis de registros, donde la coincidencia exacta suele ser suficiente. El contexto semántico adicional rara vez mejora la eficacia de la búsqueda de registros lo suficiente como para justificar el aumento de los requisitos de almacenamiento.
- El enriquecimiento semántico automático no es compatible con la función Fuente derivada.
- Limitación: las solicitudes de inferencia de indexación están actualmente limitadas a 200 TPS para los dominios de servicio. OpenSearch Este es un límite flexible; ponte en contacto con AWS Support para obtener límites más altos.

Precios

Amazon OpenSearch Service factura el enriquecimiento semántico automático en función de las unidades de OpenSearch cálculo (OCUs) consumidas durante la generación de vectores dispersos en el momento de la indexación. Solo se le cobrará por el uso real durante la indexación. Puedes monitorizar este consumo mediante la SemanticSearch OCU CloudWatch métrica de Amazon. Para obtener detalles específicos sobre los límites de los tokens del modelo, el volumen de procesamiento por OCU y un ejemplo de cálculo, visita OpenSearch Precios de servicios.

Realizar cambios de configuración en Amazon OpenSearch Service

Amazon OpenSearch Service utiliza un proceso de despliegue azul/verde al actualizar los dominios. Una blue/green implementación crea un entorno inactivo para las actualizaciones de dominio que copia el entorno de producción y redirige a los usuarios al nuevo entorno una vez finalizadas esas actualizaciones. En una blue/green implementación, el entorno azul es el entorno de producción actual. El entorno verde es el entorno de almacenamiento inactivo.

Los datos se migran del entorno azul al entorno verde. Cuando el nuevo entorno está listo, OpenSearch Service cambia los entornos para promover que el entorno ecológico se convierta en el nuevo entorno de producción. La transición se produce sin pérdida de datos. Esta práctica minimiza el tiempo de inactividad y mantiene el entorno original en caso de que la implementación en el nuevo entorno no se produzca correctamente.

Temas

- Cambios que suelen provocar blue/green despliegues
- Cambios que normalmente no provocan blue/green despliegues
- Opciones de despliegue azul/verde
- Determinar si un cambio provocará un blue/green despliegue
- Seguimiento de un cambio de configuración
- Etapas de un cambio de configuración
- Impacto de las blue/green implementaciones en el rendimiento
- Cargos por los cambios de configuración
- Solución de errores de validación

Cambios que suelen provocar blue/green despliegues

Las siguientes operaciones provocan los blue/green despliegues:

- Cambio del tipo de instancia
- Habilitar el control de acceso detallado
- Realización de actualizaciones de software del servicio
- Activación o desactivación de nodos maestros dedicados

- Activación o desactivación de Multi-AZ sin modo de espera
- Cambio del tipo de almacenamiento, el tipo de volumen o el tamaño de volumen
- Selección de diferentes subredes de VPC
- Agregar o eliminar grupos de seguridad de la VPC
- Agregar o eliminar nodos coordinadores dedicados
- Activación o desactivación de la autenticación de Amazon Cognito para paneles OpenSearch
- Elección de otro grupo de usuarios o grupo de identidades de Amazon Cognito
- Modificación de la configuración avanzada
- Actualización a una nueva OpenSearch versión (es posible que los OpenSearch paneles no estén disponibles durante una parte o durante toda la actualización)
- Habilitar el cifrado de los datos en reposo o node-to-node el cifrado
- Habilitar UltraWarm o deshabilitar el almacenamiento en frío
- Desactivación de Auto-Tune y restauración de los cambios asociados
- Asociar un complemento opcional a un dominio y disociar un complemento opcional de un dominio
- Aumento del recuento de nodos maestros dedicados para dominios Multi-AZ con dos nodos maestros dedicados
- Reducción del tamaño del volumen de EBS
- Cambio del tamaño del volumen, las IOPS o el rendimiento de EBS, si el último cambio que hizo está en curso o se produjo hace menos de 6 horas
- Habilitar la publicación de los registros de auditoría en CloudWatch.

En el caso de los dominios Multi-AZ con modo de espera, solo puede realizar una solicitud de cambio a la vez. Si ya hay un cambio en curso, se rechaza la nueva solicitud. Puede comprobar el estado del cambio actual con la API de la `DescribeDomainChangeProgress`.

Cambios que normalmente no provocan blue/green despliegues

En la mayoría de los casos, las siguientes operaciones no provocan blue/green despliegues:

- Modificación de la política de acceso
- Modificación del punto de conexión personalizado
- Cambio de la política de seguridad de la capa de transporte (TLS)
- Cambio de la hora de instantánea automatizada

- Habilitar o deshabilitar Require HTTPS
- Activación de Auto-Tune o desactivación sin revertir sus cambios
- Si su dominio tiene nodos maestros dedicados, cambie el nodo de datos o el recuento de UltraWarm nodos
- Cambio del tipo o recuento de instancias maestras dedicadas (excepto en el caso de los dominios Multi-AZ con dos nodos maestros dedicados), si el dominio tiene nodos maestros dedicados
- Habilitar o deshabilitar la publicación de registros de errores o registros lentos en CloudWatch
- Desactivar la publicación de los registros de auditoría en CloudWatch
- Aumentar el tamaño del volumen, las IOPS o el rendimiento para el tipo gp3 EBS.

> **Note**
>
> Antes del 10 de marzo de 2026, los aumentos de volumen in situ solo se admitían para volúmenes de hasta 3 TiB para gp3. El 10 de marzo de 2026, se eliminó esta limitación, lo que permitió aumentar el volumen in situ más allá de los 3 TiB. Sin embargo, si su clúster tenía un tamaño de volumen superior a 3 TiB antes de eliminar esta limitación, el primer aumento de volumen requerirá 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 implementación.

- Cómo agregar y eliminar etiquetas

> **Note**
>
> Existen algunas excepciones en función de la versión del software de servicio. Si quieres asegurarte de que un cambio no provocará una blue/green implementación, realiza una prueba antes de actualizar tu dominio, si esta opción está disponible. Algunos cambios no ofrecen una opción de ejecución de prueba. Por lo general, le recomendamos hacer cambios en el clúster fuera de las horas de mayor tráfico.

Opciones de despliegue azul/verde

Seleccione una estrategia de implementación para controlar la forma en que su clúster gestiona las implementaciones cuando no hay suficiente capacidad disponible en el momento de la actualización.

1. Full Swap azul/verde: el comportamiento de implementación predeterminado. Requiere la capacidad total de la instancia por adelantado, lo que garantiza la implementación más rápida cuando hay capacidad disponible. El despliegue no se llevará a cabo si no se puede asignar suficiente capacidad.
2. Capacidad optimizada: se recomienda para clústeres con más de 30 nodos de datos. Primero intenta realizar un blue/green intercambio completo y, si la capacidad es insuficiente, procede a la implementación por lotes. Garantiza que las implementaciones se puedan completar incluso cuando la capacidad es limitada. El tiempo de finalización puede aumentar, ya que el despliegue se realizará en lotes.

Console

Para editar el flujo de dominios:

1. Abre la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/home.
2. En el panel de navegación, en Dominios, elija el nombre de dominio para abrir la configuración del clúster.
3. Haga clic en el botón Editar situado a la derecha en la pestaña Configuración del clúster.
4. En las opciones de estrategia de despliegue, elija las configuraciones necesarias para la actualización del dominio:

    1. Full Swap azul/verde: el comportamiento de implementación predeterminado. Requiere la capacidad total de la instancia por adelantado, lo que garantiza la implementación más rápida cuando hay capacidad disponible. El despliegue no se llevará a cabo si no se puede asignar suficiente capacidad.
    2. Capacidad optimizada: se recomienda para clústeres con más de 30 nodos de datos. Primero intenta realizar un blue/green intercambio completo y, si la capacidad es insuficiente, procede a la implementación por lotes. Garantiza que las implementaciones se puedan completar incluso cuando la capacidad es limitada. El tiempo de finalización puede aumentar, ya que el despliegue se realizará en lotes.
5. Seleccione Save changes (Guardar cambios).

Para Create Domain Flow:

1. Abre la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/home.
2. En el panel de navegación izquierdo, seleccione Dominios.
3. Haga clic en el botón Crear dominio.
4. Seleccione todas las configuraciones necesarias para el dominio.
5. En las opciones de estrategia de despliegue, elija las configuraciones necesarias para la actualización del dominio:

    1. Full Swap azul/verde: el comportamiento de implementación predeterminado. Requiere la capacidad total de la instancia por adelantado, lo que garantiza la implementación más rápida cuando hay capacidad disponible. El despliegue no se llevará a cabo si no se puede asignar suficiente capacidad.
    2. Capacidad optimizada: se recomienda para clústeres con más de 30 nodos de datos. Primero intenta realizar un blue/green intercambio completo y, si la capacidad es insuficiente, procede a la implementación por lotes. Garantiza que las implementaciones se puedan completar incluso cuando la capacidad es limitada. El tiempo de finalización puede aumentar, ya que el despliegue se realizará en lotes.
6. Haga clic en el botón Crear en el panel de resumen del dominio situado a la derecha.

API

Puede configurar la estrategia de despliegue mediante la UpdateDomainConfigAPI.

Opción de capacidad optimizada

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/
config
{
   "DeploymentStrategyOptions": {
    "DeploymentStrategy": "CapacityOptimized"
   }
}
```

Opción predeterminada

```
POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/my-domain/
config
```

```
{
   "DeploymentStrategyOptions": {
    "DeploymentStrategy": "Default"
   }
}
```

Determinar si un cambio provocará un blue/green despliegue

Puede probar algunos tipos de cambios de configuración planificados para determinar si provocarán una blue/green implementación, sin tener que comprometerse con esos cambios. Antes de iniciar un cambio de configuración, utilice la consola o una de API para realizar una comprobación de validación con el fin de asegurarse de que el dominio sea apto para una actualización.

Console

Validación de un cambio de configuración

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Dominios.
3. Seleccione el dominio para el que desee realizar un cambio de configuración. Se abrirá la página de detalles del dominio. Seleccione el menú desplegable Acciones y, a continuación, Editar configuración del clúster.
4. Realice cambios en el dominio, como cambiar el tipo de instancia o el número de nodos.
5. En Análisis de prueba, elija Ejecutar. La prueba valida el cambio de configuración para detectar errores y determina si es necesario implementarlo. blue/green
6. Cuando la prueba finaliza, los resultados aparecen en la parte inferior de la página, junto con un ID de prueba. El análisis indica si el cambio de configuración requiere o no una blue/green implementación.

   Cada prueba sobrescribe la anterior. Para retener los detalles de cada prueba, guarde su ID de prueba. Las pruebas están disponibles durante 90 días o hasta que realice una actualización de configuración.
7. Para proceder con la actualización de la configuración, seleccione Guardar cambios. De lo contrario, seleccione Cancelar. Cualquiera de las opciones lo lleva de vuelta a pestaña Configuración del clúster. En esta pestaña, se puede seleccionar Detalles de la ejecución de prueba para ver los detalles de la última ejecución de prueba. Esta página también incluye

una side-by-side comparación entre la configuración anterior a la ejecución en seco y la configuración en prueba.

API

Puede realizar una validación mediante una ejecución de prueba a través de la API de configuración. Para probar los cambios con la API, establezca `DryRun` en `true`, y `DryRunMode` en `Verbose`. El modo detallado ejecuta una comprobación de validación además de determinar si el cambio iniciará una implementación azul/verde. Por ejemplo, esta UpdateDomainConfigsolicitud prueba el tipo de implementación que resulta de habilitar 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 solicitud ejecuta una comprobación de validación y devuelve el tipo de implementación que provocará el cambio, pero no lleva a cabo la actualización:

```
{
   "ClusterConfig": {
     ...
    },
   "DryRunResults": {
      "DeploymentType": "Blue/Green",
      "Message": "This change will require a blue/green deployment."
    }
}
```

Los tipos de implementación posibles son:

- `Blue/Green`: el cambio provocará una implementación azul/verde.

- `DynamicUpdate`: el cambio no provocará una implementación azul/verde.
- `Undetermined`: el dominio sigue en estado de procesamiento, por lo que no se puede determinar el tipo de implementación.
- `None`: sin cambio en la configuración.

Si la validación falla, devuelve una lista de errores de validación.

```
{
   "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."
         }
      ]
   }
}
```

Si el estado es fijo`pending`, puedes usar el ID de prueba en tu UpdateDomainConfig respuesta en DescribeDryRunProgressllamadas posteriores para comprobar el estado de la validación.

```
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."
    }
}
```

Para realizar un análisis de ejecución de prueba sin una comprobación de validación, establezca `DryRunMode` en`Basic` cuando utilice la API de configuración.

Python

El siguiente código de Python usa la UpdateDomainConfigAPI para realizar una comprobación de validación de ejecución en seco y, si la comprobación se realiza correctamente, llama a la misma API sin una ejecución en seco para iniciar la actualización. Si la comprobación no se realiza correctamente, el script imprime el error y se detiene.

```
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)
```

Seguimiento de un cambio de configuración

Puede solicitar un cambio de configuración a la vez o agrupar varios cambios en una sola solicitud. Utilice los campos Estado de procesamiento del dominio y Estado del cambio de configuración en la consola para realizar un seguimiento de los cambios de configuración. Espere a que el estado del dominio se convierta en `Active` antes de solicitar cambios adicionales.

Un dominio puede tener los siguientes estados de procesamiento:

- `Active`: no hay ningún cambio de configuración en curso. Puede enviar una nueva solicitud de cambio de configuración.
- `Creating`: se está creando el dominio.
- `Modifying`: se están haciendo cambios de configuración, como la adición de nuevos nodos de datos, EBS, gp3, el aprovisionamiento de IOPS o la configuración de claves de KMS.
- `Upgrading engine version`: se está actualizando la versión del motor.
- `Updating service software`: se está actualizando el software del servicio.
- `Deleting`: se está eliminando el dominio.
- `Isolated`: el dominio está suspendido.

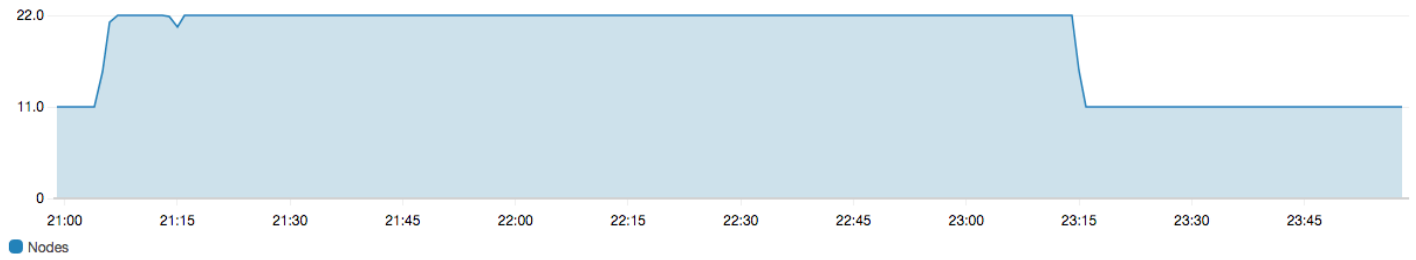
Un dominio puede tener los siguientes estados de cambio de configuración:

- `Pending`: se ha enviado una solicitud de cambio de configuración.
- `Initializing`: el servicio está iniciando un cambio de configuración.
- `Validating`: el servicio está validando los cambios solicitados y los recursos necesarios.
- `Awaiting user inputs`: el servicio espera que continúen los cambios de configuración, como un cambio de tipo de instancia. Puede editar los cambios de configuración.
- `Applying changes`: el servicio está aplicando los cambios de configuración solicitados.
- `Cancelled`: el cambio de configuración está cancelado. Seleccione Cancelar y revertir todos los cambios.
- `Completed`: los cambios de configuración solicitados se han completado correctamente.
- `Validation failed`: los cambios de configuración solicitados no se pudieron completar. No se aplicó ningún cambio de configuración.

> **Note**
>
> Los errores de validación se deben a la presencia de índices rojos en el dominio, a la falta de disponibilidad del tipo de instancia elegido o a la falta de espacio en disco. Para obtener una lista de los errores de validación, consulte the section called “Solución de errores de validación”. Durante un evento de fallo de validación, puede cancelar, reintentar o editar los cambios de configuración.

Cuando se completen los cambios de configuración, el estado del dominio volverá a ser `Active`.

Puedes revisar el estado del clúster y CloudWatch las métricas de Amazon y comprobar que el número de nodos del clúster aumenta temporalmente (a menudo se duplica) mientras se actualiza el dominio. En la siguiente ilustración, puede ver el número de nodos duplicarse de 11 a 22 durante un cambio de configuración y volver a 11 una vez completada la actualización.



Este duplicado provisional puede afectar al rendimiento de los nodos maestros dedicados del clúster, que repentinamente podría tener que administrar más nodos. También puede aumentar las latencias de búsqueda e indexación a medida que OpenSearch Service copia los datos del clúster antiguo al nuevo. Es importante mantener una capacidad suficiente en el clúster para gestionar la sobrecarga asociada a estas blue/green implementaciones.

> **Important**
>
> El usuario no incurrirá en ningún gasto adicional durante los cambios de configuración y el mantenimiento del servicio. Solo se facturará por el número de nodos que solicite para el clúster. Para obtener información detallada, consulte the section called “Cargos por los cambios de configuración”.

Para evitar sobrecargar los nodos maestros dedicados, puedes monitorear el uso con las CloudWatch métricas de Amazon. Para conocer los valores máximos recomendados, consulte the section called “ CloudWatch Alarmas recomendadas”.

Etapas de un cambio de configuración

Tras iniciar un cambio de configuración, OpenSearch Service sigue una serie de pasos para actualizar tu dominio. Puede ver el progreso del cambio de configuración en Estado de cambio de configuración en la consola. Los pasos exactos por los que se realiza una actualización dependen del tipo de cambio que realices. También puede supervisar un cambio de configuración mediante la operación de la API DescribeDomainChangeProgress.

Las siguientes son las posibles etapas que puede realizar una actualización durante un cambio de configuración:

Nombre de la fase	Description (Descripción)
Validación	Validación de que el dominio es apto para una actualización y mostrar problemas de

Nombre de la fase	Description (Descripción)
	validación si es necesario.
Creación de un nuevo entorno	Completar los requisitos previos necesarios y crear los recursos necesarios para iniciar la blue/green implementación.
Provisión de nuevos nodos	Creación de un nuevo conjunto de instancias en el nuevo entorno.
Enrutamiento del tráfico en nodos nuevos	Redireccionamiento del tráfico a los nodos de datos recién creados.
Enrutamiento del tráfico en nodos antiguos	Desactivación del tráfico en los nodos de datos antiguos.

Nombre de la fase	Description (Descripción)
Preparación de los nodos para su eliminación	Preparación para eliminar nodos. Este paso solo ocurre cuando estás reduciendo la escala de su dominio (por ejemplo, de 8 a 6 nodos).
Copia de particiones en nodos nuevos	Trasladando los fragmentos de los nodos antiguos a los nuevos.
Terminación de nodos	Finalización y eliminación de nodos antiguos después de remover los fragmentos.
Eliminación de recursos más antiguos	Eliminación de recursos asociados al entorno anterior (por ejemplo, equilibrador de carga).

Nombre de la fase	Description (Descripción)
Actualización dinámica	Se muestra cuando la actualización no requiere un blue/green despliegue y se puede aplicar de forma dinámica.
Aplicación de cambios específicos relacionados con la instancia maestra	Se muestra cuando se cambia el tipo o el número de la instancia maestra dedicada.
Aplicación de cambios relacionados con el volumen	Se muestra cuando se modifican el tamaño, el tipo, las IOPS y el rendimiento del volumen.

Impacto de las blue/green implementaciones en el rendimiento

Durante el blue/green despliegue, tu clúster OpenSearch de Amazon Service estará disponible para las solicitudes entrantes de búsqueda e indexación. Sin embargo, es posible que observe los siguientes problemas de rendimiento:

- Aumento temporal del uso en los nodos principales, ya que los clústeres tienen que administrar más nodos.
- Aumento de la latencia de búsqueda e indexación a medida que el OpenSearch servicio copia datos de los nodos antiguos a los nuevos.
- Aumento del número de rechazos de solicitudes entrantes a medida que aumenta la carga del clúster durante blue/green las implementaciones.
- Para evitar problemas de latencia y el rechazo de solicitudes, debes ejecutar las blue/green implementaciones cuando el clúster esté en buen estado y haya poco tráfico de red.

Cargos por los cambios de configuración

Si cambias la configuración de un dominio, OpenSearch Service crea un clúster nuevo tal y como se describe en. the section called “Cambios de configuración” Durante la migración de uno a otro, se devengarán los cargos siguientes:

- Si cambia el tipo de instancia, se cobrará por los dos clústeres durante la primera hora. Después de la primera hora, solo se cobrará por el clúster nuevo. Los volúmenes de EBS no se cobran dos veces porque forman parte del clúster, por lo que la facturación corresponde a la facturación de instancias.

  Ejemplo: Supongamos que cambia la configuración y pasa de tres instancias `m3.xlarge` a cuatro instancias `m4.large`. Durante la primera hora, se cobrarán los dos clústeres (3* `m3.xlarge` + 4* `m4.large`). Después de la primera hora, solo se cobrará el clúster nuevo (4 * `m4.large`).
- Si no cambia el tipo de instancia, solo se cobrará el clúster mayor durante la primera hora. Después de la primera hora, solo se cobrará por el clúster nuevo.

  Ejemplo: Supongamos que cambia la configuración y pasa de seis instancias `m3.xlarge` a tres instancias `m3.xlarge`. Durante la primera hora, se cobrará por el clúster mayor (6 * `m3.xlarge`). Después de la primera hora, solo se cobrará el clúster nuevo (3 * `m3.xlarge`).

Solución de errores de validación

Cuando inicias un cambio de configuración o realizas una actualización de una versión OpenSearch o de Elasticsearch, OpenSearch Service primero realiza una serie de comprobaciones de validación para garantizar que tu dominio sea apto para una actualización. Si se produce un error en alguna

de estas comprobaciones, recibirá una notificación en la consola con los problemas específicos que debe corregir antes de actualizar su dominio.

OpenSearch El servicio aísla tu dominio si permanece inutilizable durante más de 60 días. Te enviaremos recordatorios de notificación para resolver estos problemas. Si no solucionas los problemas dentro de este plazo, el OpenSearch Servicio eliminará tu dominio y sus datos.

En la siguiente tabla, se enumeran los posibles problemas de dominio que puedan surgir en OpenSearch Service y los pasos para resolverlos.

Problema	Código de error	Pasos para la solución de problemas
No se encontró el grupo de seguridad	`SecurityGroupNotFound`	El grupo de seguridad asociado a su dominio de OpenSearch servicio no existe. Para solucionar este problema, cree un grupo de seguridad con el nombre especificado.
No se encontró la subred	`SubnetNotFound`	La subred asociada a su dominio OpenSearch de servicio no existe. Para solucionar este problema, cree una subred en la VPC.
No se configuró el rol vinculado al servicio	`SLRNotConfigured`	La función vinculada al servicio para el OpenSearch Servicio no está configurada. El rol vinculado al servicio está predefinido por el OpenSearch Servicio e incluye todos los permisos que el servicio necesita para llamar a otros AWS servicios en su nombre. Si el rol no existe, es posible que tenga que crearlo manualmente.
No hay suficientes direcciones IP	`InsufficientFreeIPsForSubnets`	Una o más de las subredes de VPC no tienen suficientes direcciones IP para actualizar el dominio. Para calcular el número de direcciones IP que necesita, consulte the section called “Reserva de direcciones IP en una subred de una VPC”.
No existe el grupo de usuarios de Cognito	`CognitoUserPoolNotFound`	OpenSearch El servicio no encuentra el grupo de usuarios de Amazon Cognito. Verifique que creó uno y que tiene el ID correcto. Para encontrar el ID, puede utilizar la consola de Amazon Cognito o el siguiente comando de la AWS CLI :

Problema	Código de error	Pasos para la solución de problemas
		`aws cognito-idp list-user-pools --max-results 60 --region us-east-1`
El grupo de identidades de Cognito no existe	`CognitoIdentityPoolNotFound`	OpenSearch El servicio no encuentra el grupo de identidades de Cognito. Verifique que creó uno y que tiene el ID correcto. Para encontrar el ID, puede utilizar la consola de Amazon Cognito o el siguiente comando de la AWS CLI : `aws cognito-identity list-identity-pools --max-results 60 --region us-east-1`
No se encontró el dominio de Cognito para el grupo de usuarios	`CognitoDomainNotFound`	El grupo de usuarios no tiene un nombre de dominio. Puede configurar uno mediante la consola de Amazon Cognito o el siguiente comando: AWS CLI `aws cognito-idp create-user-pool-domain --domain my-domain --user-pool-id id`
No se configuró el rol de Cognito	`CognitoRoleNotConfigured`	La función de IAM que concede al OpenSearch Servicio el permiso para configurar los grupos de usuarios e identidades de Amazon Cognito y utilizarlos para la autenticación no está configurada. Configure el rol con un conjunto de permisos y una relación de confianza adecuados. Puede usar la consola, que crea el CognitoAccessForAmazonOpenSearchrol predeterminado para usted, o puede configurar un rol manualmente mediante el SDK AWS CLI o el AWS mismo.
No se puede describir el grupo de usuarios	`UserPoolNotDescribable`	El rol de Amazon Cognito especificado no tiene permiso para describir el grupo de usuarios asociado a su dominio. Asegúrese de que la política de permisos de rol permita la acción `cognito-identity:DescribeUserPool` . Consulte the section called “Acerca del CognitoAccessForAmazonOpenSearch rol” para ver la política de permisos completa.

Problema	Código de error	Pasos para la solución de problemas
No se puede describir el grupo de identidades	`IdentityPoolNotDescribable`	El rol de Amazon Cognito especificado no tiene permiso para describir el grupo de identidades asociado a su dominio. Asegúrese de que la política de permisos de rol permita la acción `cognito-identity:DescribeIdentityPool`. Consulte the section called “Acerca del CognitoAccessForAmazonOpenSearch rol” para ver la política de permisos completa.
No se pueden describir los grupos de usuarios y de identidades	`CognitoPoolsNotDescribable`	El rol de Amazon Cognito especificado no tiene permiso para describir los grupos de usuarios e identidades asociados a su dominio. Asegúrese de que la política de permisos de rol permita las acciones `cognito-identity:DescribeIdentityPool` y `cognito-identity:DescribeUserPool`. Consulte the section called “Acerca del CognitoAccessForAmazonOpenSearch rol” para ver la política de permisos completa.
La clave de KMS no está habilitada	`KMSKeyNotEnabled`	La clave AWS Key Management Service (AWS KMS) utilizada para cifrar tu dominio está deshabilitada. Vuelva a habilitar la clave inmediatamente.
El estado del certificado personalizado no es ISSUED (EMITIDO)	`InvalidCertificate`	Si tu dominio usa un punto final personalizado, lo proteges generando un certificado SSL en AWS Certificate Manager (ACM) o importando uno propio. El estado del certificado debe ser Emitido. Si aparece este error, compruebe el estado del certificado en la consola de ACM. Si el estado aparece como Expired (Vencido), Failed (Error), Inactive (Inactivo) o Pending validation (Pendiente de validación), consulte la documentación sobre solución de problemas de ACM para solucionar el problema.

Problema	Código de error	Pasos para la solución de problemas
No hay capacidad suficiente para lanzar el tipo de instancia elegido	`InsufficientInstanceCapacity`	La capacidad del tipo de instancia solicitada no está disponible. Por ejemplo, es posible que hayas solicitado cinco `i3.16xlarge.search` nodos, pero el OpenSearch Servicio no tiene suficientes `i3.16xlarge.search` hosts disponibles, por lo que no se puede tramitar la solicitud. Comprueba los tipos de instancia compatibles en OpenSearch Service y elige un tipo de instancia diferente.
Índices rojos en clúster	`RedCluster`	Uno o más índices del clúster tienen un estado rojo, lo que lleva a un estado rojo general del clúster. Para solucionar este problema, consulte the section called “Estado rojo del clúster”.
Interruptor de circuitos de memoria, demasiadas solicitudes	`TooManyRequests`	Hay demasiadas solicitudes de búsqueda y escritura en tu dominio, por lo que OpenSearch Service no puede actualizar su configuración. Puede reducir la cantidad de solicitudes, escalar las instancias verticalmente hasta 64 GiB de RAM o escalar horizontalmente mediante la adición de instancias.
La nueva configuración no puede almacenar datos (poco espacio en disco)	`InsufficientStorageCapacity`	El tamaño de almacenamiento configurado no puede contener todos los datos de su dominio. Para solucionar este problema, elija un volumen mayor, elimine los índices que no se utilizan o aumente la cantidad de nodos del clúster para liberar espacio en el disco de inmediato.

Problema	Código de error	Pasos para la solución de problemas

Particiones ancladas a nodos específicos	`ShardMovementBlocked`	Uno o más índices de su dominio están adjuntos a nodos específicos y no se pueden reasignar. Lo más probable es que esto suceda porque se configuró el filtrado de asignación de particiones, que le permite especificar qué nodos tienen permiso para alojar las particiones de un índice en particular.
		
		Para solucionar este problema, elimine los filtros de asignación de particiones de todos los índices afectados:
		
		`PUT my-index/_settings`
		`{`
		`  "settings": {`
		`  "index.routing.allocation.require._name": null`
		`  }`
		`}`

La nueva configuración no puede contener todas las particiones (recuento de particiones)	`TooManyShards`	El número de fragmentos de tu dominio es demasiado alto, lo que impide que OpenSearch Service los traslade a la nueva configuración. Para solucionar este problema, escale su dominio horizontalmente. Para ello, agregue nodos del mismo tipo de configuración que los nodos actuales del clúster. Tenga en cuenta que el tamaño máximo de volumen de EBS depende del tipo de instancia del nodo. Para evitar este problema en el futuro, consulte the section called “Selección del número de particiones” y defina una estrategia de partición que sea adecuada para su caso de uso.
La subred asociada a tu dominio no admite direcciones IPv4	`ResultCodeIPv4BlockNotExists`	Para resolver este problema, cree una subred o actualice la subred existente en la VPC según el tipo de dirección IP configurada del dominio. Si tu dominio usa un tipo de dirección IPv4exclusivo, usa una subred IPv4 solo. Si su dominio usa el modo de doble pila, use una subred de doble pila.

Problema	Código de error	Pasos para la solución de problemas
La subred asociada a tu dominio no admite direcciones IPv6	`ResultCodeIPv6BlockNotExists`	Para resolver este problema, cree una subred o actualice la subred existente en la VPC según el tipo de dirección IP configurada del dominio. Si tu dominio usa un tipo de dirección IPv4exclusivo, usa una subred IPv4 solo. Si su dominio usa el modo de doble pila, use una subred de doble pila.

Actualizaciones de software de servicio en Amazon OpenSearch Service

> **Note**
>
> Para obtener una explicación de los cambios y adiciones realizados en cada actualización del software de servicio principal (sin parches), consulte las notas de la versión.

Amazon OpenSearch Service publica periódicamente actualizaciones de software de servicio que añaden funciones o mejoran tus dominios. El panel Notificaciones en la consola es la forma más sencilla para ver si hay disponible una actualización o para verificar el estado de una actualización. Cada notificación incluye detalles sobre la actualización del software del servicio. Todas las actualizaciones de software utilizan implementaciones azul/verde para minimizar el tiempo de inactividad.

Las actualizaciones del software del servicio son diferentes de las actualizaciones de las versiones del OpenSearch servicio. Para obtener información sobre la actualización a una versión posterior del OpenSearch Servicio, consultethe section called “Actualización de dominios”.

OpenSearch El servicio requiere que aplique las actualizaciones de software de servicio requeridas en un plazo de 30 días a partir de su disponibilidad. Estas actualizaciones son fundamentales para mantener el cumplimiento de las normas de seguridad.

Si no aplicas las actualizaciones obligatorias dentro del período de 30 días, recibirás notificaciones recordatorias cada 15 días durante 30 días. Tras este periodo sin conformidad, tu dominio quedará aislado con los siguientes efectos:

- Se eliminarán todos los accesos de red a tu dominio
- El estado del dominio cambia a aislado
- El dominio permanece inutilizable hasta que aplique las actualizaciones necesarias

Durante el aislamiento, seguirás recibiendo notificaciones recordatorias cada 15 días durante 60 días. Si no aplicas las actualizaciones necesarias dentro de este período, tu dominio de OpenSearch servicio y todos los datos asociados se eliminarán permanentemente. Para obtener más información, consulta Solución de errores de validación.

Actualizaciones opcionales frente a actualizaciones obligatorias

OpenSearch El servicio tiene dos categorías amplias de actualizaciones de software de servicio:

Actualizaciones opcionales

Las actualizaciones del software de servicio opcionales generalmente incluyen mejoras y soporte para nuevas características o funcionalidades. Las actualizaciones opcionales no se aplican a sus dominios y no hay un plazo fijo para instalarlas. La disponibilidad de la actualización se comunica por correo electrónico y una notificación en la consola. Puede optar por aplicar la actualización inmediatamente o reprogramarla para una fecha y hora más adecuadas. También puede programarla durante la ventana de menor actividad del dominio. La mayoría de las actualizaciones de software son opcionales.

Independientemente de si programa o no una actualización, si realiza un cambio en el dominio que provoque una implementación azul/verde, OpenSearch Service actualizará automáticamente el software del servicio por usted.

Puede configurar su dominio para que aplique automáticamente las actualizaciones opcionales durante la ventana de menor actividad. Cuando esta opción está activada, el OpenSearch Servicio espera al menos 13 días desde que esté disponible una actualización opcional y, a continuación, programa la actualización después de siete días. Recibirá una notificación en la consola cuando la actualización esté programada y podrá elegir reprogramarla para una fecha posterior.

Para activar las actualizaciones de software automáticas, seleccione Habilitar la actualización automática de software al crear o actualizar su dominio. Para configurar el mismo ajuste mediante el AWS CLI, `--software-update-options true` establézcalo al crear o actualizar el dominio.

Actualizaciones requeridas

Las actualizaciones de software de servicio requeridas suelen incluir correcciones de seguridad críticas u otras actualizaciones obligatorias para garantizar la integridad y la funcionalidad continuas de su dominio. Algunos ejemplos de actualizaciones obligatorias son las vulnerabilidades y exposiciones comunes de Log4j () y la aplicación de la versión 2 (CVEs) del Servicio de Metadatos de Instancia. IMDSv2 El número de actualizaciones obligatorias en un año suele ser inferior a tres.

OpenSearch El servicio programa automáticamente estas actualizaciones y le notifica siete días antes de la actualización programada por correo electrónico y una notificación de consola. Puede optar por aplicar la actualización inmediatamente o reprogramarla para una fecha y hora más adecuadas dentro del plazo permitido. También puede programarla durante la próxima ventana de menor actividad del dominio. Si no realizas ninguna acción en relación con una actualización obligatoria y no realizas ningún cambio en el dominio que provoque una blue/green implementación, el OpenSearch Servicio puede iniciar la actualización en cualquier momento después de la fecha límite especificada (normalmente 14 días a partir de la disponibilidad), dentro del período de menor actividad del dominio.

Independientemente de cuándo esté programada la actualización, si realizas un cambio en el dominio que provoque un despliegue azul o verde, OpenSearch Service actualizará tu dominio automáticamente.

Actualizaciones de revisión

Las versiones del software de servicio que terminan en «-P» y en un número, como R20211203-, son versiones de parches. *P4* Es probable que las revisiones incluyan mejoras de rendimiento, correcciones de errores menores y correcciones de seguridad o mejoras de posición. Las versiones de revisión no incluyen nuevas características ni cambios extraordinarios y, por lo general, no tienen un impacto directo o notable en los usuarios. La notificación del software de servicio le indica si el lanzamiento de una revisión es opcional u obligatoria.

Consideraciones

Tenga en cuenta lo siguiente cuando tenga que tomar una decisión sobre la actualización del dominio:

- La actualización manual del dominio permite aprovechar las nuevas características de forma más rápida. Si selecciona Actualizar, OpenSearch Service coloca la solicitud en una cola e inicia la actualización cuando tiene tiempo.

- Al iniciar una actualización del software del servicio, el OpenSearch Servicio envía una notificación cuando se inicia la actualización y cuando se completa.
- Las actualizaciones de software utilizan blue/green las implementaciones para minimizar el tiempo de inactividad. Las actualizaciones pueden sobrecargar temporalmente los nodos maestros dedicados de un clúster, así que asegúrese de mantener la capacidad suficiente para manejar la sobrecarga asociada.
- Normalmente, las actualizaciones se completan en cuestión de minutos, pero también pueden tardar varias horas o incluso días si el sistema experimenta una carga pesada. Considere actualizar el dominio durante la ventana de menor actividad configurada para evitar largos periodos de actualización.

Iniciar una actualización del software de servicio

Puede solicitar una actualización del software del servicio a través de la consola de OpenSearch servicio AWS CLI, la o una de las SDKs.

Consola

Cómo solicitar una actualización del software de servicio

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Seleccione el nombre de dominio para abrir su configuración.
3. Seleccione Acciones, Actualizar y seleccione una de las siguientes opciones:

    - Aplicar la actualización ahora: programa de inmediato la acción para que se lleve a cabo en la hora actual si hay capacidad disponible. Si no hay capacidad disponible, ofrecemos otras franjas horarias disponibles para elegir.
    - Programarla en una ventana de menor actividad: solo está disponible si la ventana de menor actividad está activada para el dominio. Programa la actualización para que se lleve a cabo durante la ventana de menor actividad configurada en el dominio. No hay garantía de que la actualización se produzca durante la siguiente ventana inmediata. Dependiendo de la capacidad, podría producirse en los días siguientes. Para obtener más información, consulte the section called “Intervalos de menor demanda”.
    - Programar para una fecha y hora específicas: programa la actualización para que se lleve a cabo en una fecha y hora específicas. Si la hora que especifica no está disponible por motivos de capacidad, puede seleccionar una franja horaria diferente.

Si programa la actualización para una fecha posterior (dentro o fuera de la ventana de menor actividad del dominio), puede reprogramarla en cualquier momento. Para obtener instrucciones, consulte the section called “Reprogramación de acciones”.

4. Seleccione Confirmar.

AWS CLI

Envía una start-service-software-update AWS CLI solicitud para iniciar una actualización del software del servicio. En este ejemplo, se añade la actualización a la cola de forma inmediata:

```
aws opensearch start-service-software-update \
  --domain-name my-domain \
  --schedule-at "NOW"
```

Respuesta:

```
{
    "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

Después de solicitar una actualización, dispone de una pequeña ventana de tiempo para cancelarla. La duración de este `PENDING_UPDATE` estado puede variar considerablemente y depende de usted Región de AWS y del número de actualizaciones simultáneas que realice el OpenSearch Servicio. Para cancelar una actualización, usa la consola o el `cancel-service-software-update` AWS CLI comando.

Si la solicitud falla con una `BaseException`, eso significa que la hora que especificó no está disponible por motivos de capacidad y debe especificar una hora diferente. OpenSearch El servicio proporciona sugerencias alternativas de ranuras disponibles en la respuesta.

AWS SDKs

Este ejemplo de script de Python utiliza los métodos describe_domain y start_service_software_update de AWS SDK para Python (Boto3) para comprobar si un dominio es apto para una actualización de software de servicio y, de ser así, inicia la actualización. Se debe proporcionar un valor para `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)
```

Programación de actualizaciones del software durante las ventanas de menor actividad

Cada dominio OpenSearch de servicio creado después del 16 de febrero de 2023 tiene un período diario de 10 horas entre las 22:00 y las 8:00 a. m., hora local, que consideramos el período de menor actividad. OpenSearch El servicio usa esta ventana para programar las actualizaciones del software del servicio para el dominio. Las actualizaciones fuera de las horas pico ayudan a minimizar la carga de los nodos maestros dedicados de un clúster durante los períodos de mayor tráfico. OpenSearch El servicio no puede iniciar las actualizaciones fuera de este período de 10 horas sin su consentimiento.

- En el caso de las actualizaciones opcionales, el OpenSearch Servicio te notifica la disponibilidad de la actualización y te pide que la programes durante las próximas horas de menor actividad.
- En el caso de las actualizaciones necesarias, el OpenSearch Servicio programa automáticamente la actualización durante las próximas horas de menor actividad y te lo notifica con tres días de antelación. Puede reprogramar la actualización (dentro o fuera de la ventana de menor actividad), pero solo dentro del plazo requerido para que se complete la actualización.

Para cada dominio, puede optar por anular la hora de inicio predeterminada a las 22:00 por una hora personalizada. Para obtener instrucciones, consulte the section called “Configuración de un intervalo de menor demanda personalizado”.

Consola

Cómo programar una actualización durante una próxima ventana de menor actividad

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Seleccione el nombre de dominio para abrir su configuración.
3. Seleccione Acciones, Actualizar.
4. Seleccione Programarla en la ventana de menor actividad.
5. Seleccione Confirmar.

Puede ver la acción programada en la pestaña de la ventana de menor actividad y reprogramarla en cualquier momento. Consulte the section called “Visualización de acciones programadas”.

CLI

Para programar una actualización durante una próxima temporada de menor actividad mediante el AWS CLI, envíe una StartServiceSoftwareUpdatesolicitud y especifique `OFF_PEAK_WINDOW` el `--schedule-at` parámetro:

```
aws opensearch start-service-software-update \
  --domain-name my-domain \
  --schedule-at "OFF_PEAK_WINDOW"
```

Supervisar actualizaciones de software de servicio

OpenSearch El servicio envía una notificación cuando una actualización del software del servicio está disponible, es necesaria, se ha iniciado, se ha completado o se ha producido un error. Puede

ver estas notificaciones en el panel de notificaciones de la consola de OpenSearch servicio. La gravedad de la notificación es `Informational` si la actualización es opcional y `High` si es necesaria.

OpenSearch El servicio también envía eventos de software de servicio a Amazon EventBridge. Puede utilizarlas EventBridge para configurar reglas que envíen un correo electrónico o realicen una acción específica cuando se recibe un evento. Para ver un ejemplo de recorrido, consultethe section called “Tutorial: Sending SNS alerts for available updates”.

Para ver el formato de cada evento de software de servicio enviado a Amazon EventBridge, consultathe section called “Eventos de actualización del software de servicio”.

Cuando los dominios no son aptos para una actualización

El dominio no es apto para una actualización del software de servicio si está en cualquiera de los estados que se muestran a continuación:

Estado	Description (Descripción)
Dominio en procesamiento	El dominio está en medio de un cambio de configuración. Compruebe si es posible optar a la actualización una vez que se complete la operación.
Estado rojo del clúster	Uno o varios índices en el clúster aparecen en rojo. Para ver los pasos de solución de problemas, consulte the section called “Estado rojo del clúster”.
Tasa de errores alta	El OpenSearch clúster devuelve una gran cantidad de errores de 5 xx al intentar procesar las solicitudes. Este problema suele ser el resultado de demasiadas solicitudes de lectura o escritura simultáneas. Considere la posibilidad de reducir el tráfico hacia el clúster o el escalado del dominio.
Cerebro dividido	El cerebro dividido significa que el OpenSearch clúster tiene más de un nodo maestro y se ha dividido en dos clústeres que nunca volverán a unirse por sí solos. Puede evitar un cerebro dividido con el número recomendado de nodos maestros dedicados. Para obtener ayuda a fin de recuperar un cerebro dividido, póngase en contacto con Soporte.

Estado	Description (Descripción)
Problema de integración de Amazon Cognito	Tu dominio usa la autenticación para los OpenSearch paneles y el OpenSearch Servicio no puede encontrar uno o más recursos de Amazon Cognito. Este problema suele producirse si no se encuentra el grupo de usuarios de Amazon Cognito. Para corregir el problema, vuelva a crear el recurso que falta y configure el dominio del OpenSearch servicio para usarlo.
Otro problema de servicio	Los problemas relacionados con el propio OpenSearch Servicio pueden hacer que tu dominio aparezca como no apto para una actualización. Si en su dominio no se aplican ninguna de las condiciones anteriores y el problema persiste durante más de un día, póngase en contacto con Soporte.

Definición de ventanas fuera de las horas pico para Amazon Service OpenSearch

Cuando creas un dominio de Amazon OpenSearch Service, defines un período diario de 10 horas que se considera horas de menor actividad. OpenSearch El servicio utiliza esta ventana para programar las actualizaciones del software del servicio y las optimizaciones de Auto-Tune que requieren una implementación azul/verde durante tiempos de tráfico comparativamente más bajos, siempre que sea posible. Blue/green se refiere al proceso de crear un nuevo entorno para las actualizaciones del dominio y redirigir a los usuarios al nuevo entorno una vez finalizadas esas actualizaciones.

Si bien blue/green las implementaciones no son disruptivas, para minimizar cualquier posible impacto en el rendimiento mientras se consumen los recursos de una blue/green implementación, le recomendamos que programe estas implementaciones durante el período de menor actividad configurado por el dominio. Las actualizaciones, como las sustituciones de nodos o las que deben implementarse en el dominio de forma inmediata, no utilizan el intervalo de menor demanda.

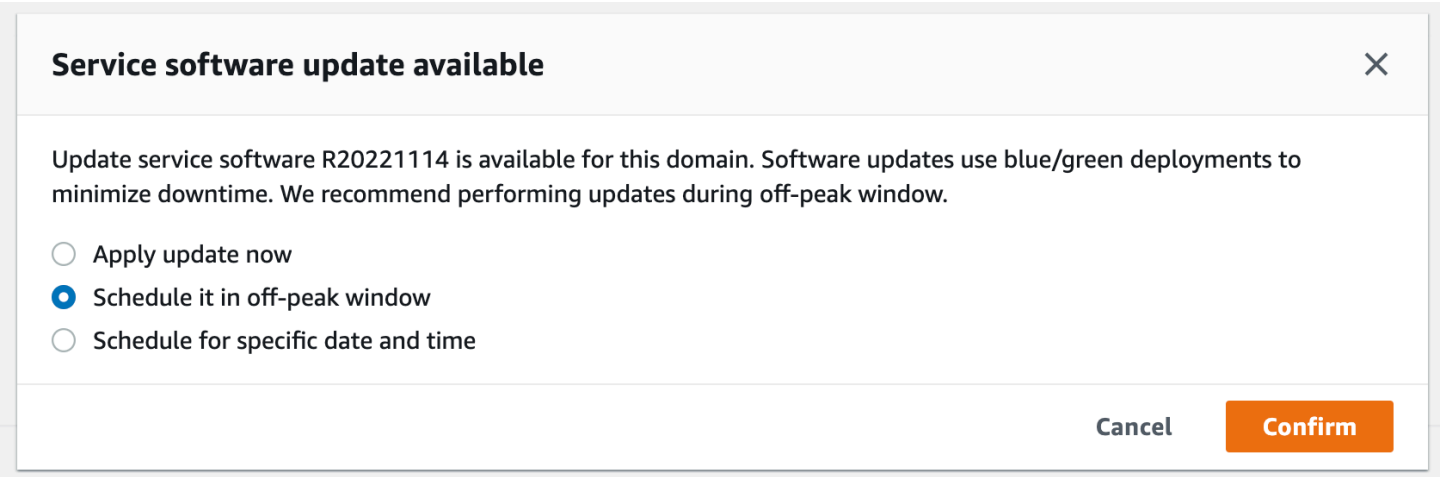
Puede modificar la hora de inicio del intervalo de menor demanda, pero no puede modificar la longitud del intervalo.

> **Note**
>
> Los intervalos de menor demanda se introdujeron el 16 de febrero de 2023. Todos los dominios creados antes de esta fecha tienen el intervalo de menor demanda desactivado de forma predeterminada. Debe habilitar y configurar manualmente el intervalo de menor demanda para estos dominios. Todos los dominios creados después de esta fecha tendrán habilitado el intervalo de menor demanda de forma predeterminada. No puede deshabilitar el intervalo de menor demanda de un dominio una vez que esté habilitado.

Actualizaciones de software de servicio de menor demanda

OpenSearch El servicio tiene dos categorías amplias de actualizaciones de software de servicio: opcionales y obligatorias. Ambos tipos requieren blue/green implementaciones. Las actualizaciones opcionales no se aplican en los dominios, mientras que las actualizaciones obligatorias se instalan automáticamente si no realiza ninguna acción antes de la fecha límite especificada (normalmente dos semanas después de la disponibilidad). Para más información, consulte the section called "Actualizaciones opcionales frente a actualizaciones obligatorias".

Cuando inicia una actualización opcional, tiene la opción de aplicarla inmediatamente, programarla para un intervalo de menor demanda posterior o especificar una fecha y hora personalizadas para aplicarla.



Para las actualizaciones necesarias, el OpenSearch servicio programa automáticamente una fecha y una hora durante las horas de menor actividad para realizar la actualización. Recibe una notificación tres días antes de la actualización programada, y puede optar por reprogramarla para una fecha

y hora posteriores dentro del intervalo de implementación obligatorio. Para ver las instrucciones, consulte the section called “Reprogramación de acciones”.

Optimizaciones de ajuste automático durante horas de menor demanda

Anteriormente, Auto-Tune utilizaba ventanas de mantenimiento para programar los cambios que requerían una blue/green implementación. Los dominios que ya tenían habilitados ajuste automático y los intervalos de mantenimiento antes de la introducción de los intervalos de menor demanda seguirán utilizando los intervalos de mantenimiento para estas actualizaciones, a menos que los migre para utilizar los intervalos de menor demanda.

Le recomendamos que migre sus dominios para utilizar el intervalo de menor demanda, ya que se utiliza para programar otras actividades en el dominio, como las actualizaciones del software de servicio. Para ver las instrucciones, consulte the section called “Migración desde los intervalos de mantenimiento de ajuste automático”. No puede volver a usar los intervalos de mantenimiento después de migrar su dominio a los intervalos de menor demanda.

Todos los dominios creados después del 16 de febrero de 2023 utilizarán el período de menor actividad, en lugar de los períodos de mantenimiento tradicionales, para programar blue/green las implementaciones. No puede deshabilitar el intervalo de menor demanda de un dominio. Para ver una lista de las optimizaciones de ajuste automático que requieren implementaciones azul/verde, consulte the section called “Tipos de cambios”.

Habilitación del intervalo de menor demanda

Todos los dominios creados antes del 16 de febrero de 2023 (cuando se introdujeron los intervalos de menor demanda) tienen esa característica desactivada de forma predeterminada. Debe habilitarlo manualmente para estos dominios. No puede deshabilitar el intervalo de menor demanda después de activarlo.

Consola

Cómo habilitar el intervalo de menor demanda en un dominio

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Seleccione el nombre del dominio para abrir la configuración.
3. Acceda a la pestaña de intervalo de menor demanda y seleccione Editar.

4. Especifique una hora de inicio personalizada de acuerdo con el tiempo universal coordinado (UTC). Por ejemplo, para configurar una hora de inicio a las 23:30 en la región Oeste de EE. UU. (Oregón), especifique 07:30.
5. Seleccione Guardar cambios.

CLI

Para modificar la ventana de menor actividad mediante el AWS CLI, envíe una UpdateDomainConfigsolicitud:

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --off-peak-window-options 'Enabled=true,
 OffPeakWindow={WindowStartTime={Hours=02,Minutes=00}}'
```

Si no especifica una hora de inicio personalizada para el intervalo, el valor predeterminado es 00:00 UTC.

Configuración de un intervalo de menor demanda personalizado

Debe especificar un intervalo de menor demanda personalizado para su dominio usando el tiempo universal coordinado (UTC). Por ejemplo, si desea que el intervalo de menor demanda comience a las 23:00 para un dominio en la región Este de EE. UU. (Norte de Virginia), debe especificar a las 04:00 UTC.

Consola

Cómo modificar el intervalo de horas de menor demanda de un dominio

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Seleccione el nombre del dominio para abrir la configuración.
3. Acceda a la pestaña del intervalo de menor demanda. Puede ver el intervalo de menor demanda configurado y una lista de las próximas acciones programadas para el dominio.
4. Seleccione Editar y especifique una nueva hora de inicio en UTC. Por ejemplo, para configurar la hora de inicio a las 21:00 en la región Este de EE. UU. (Norte de Virginia), especifique 02:00 UCT.
5. Seleccione Guardar cambios.

CLI

Para configurar una ventana personalizada fuera de las horas pico mediante el AWS CLI, envíe una UpdateDomainConfigsolicitud y especifique la hora y los minutos en formato horario de 24 horas.

Por ejemplo, la siguiente solicitud cambia la hora de inicio del intervalo a las 2:00 UTC:

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --off-peak-window-options 'OffPeakWindow={WindowStartTime={Hours=02,Minutes=00}}'
```

Si no especifica la hora de inicio del intervalo, el valor predeterminado es las 22:00, hora local, para la Región de AWS en la que se crea el dominio.

Visualización de acciones programadas

Puede ver todas las acciones que están actualmente programadas, en curso o pendientes para cada uno de sus dominios. Las acciones pueden tener una gravedad de `HIGH`, `MEDIUM` y `LOW`.

Las acciones pueden tener los siguientes estados:

- `Pending update`: la acción está en la cola para ser procesada.
- `In progress`: la acción está actualmente en curso.
- `Failed`: la acción no pudo completarse.
- `Completed`: la acción ha finalizado correctamente.
- `Not eligible`: solo para actualizaciones de software de servicio. La actualización no puede continuar porque el clúster está en mal estado.
- `Eligible`: solo para actualizaciones de software de servicio. El dominio es elegible para una actualización.

Consola

La consola OpenSearch de servicio muestra todas las acciones programadas en la configuración del dominio, junto con la gravedad y el estado actual de cada acción.

Cómo visualizar las acciones programadas de un dominio

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.

2. Seleccione el nombre del dominio para abrir la configuración.
3. Acceda a la pestaña del intervalo de menor demanda.
4. En Acciones programadas, consulte todas las acciones que están actualmente programadas, en curso o pendientes en el dominio.

CLI

Para ver las acciones programadas mediante el AWS CLI, envía una ListScheduledActionssolicitud:

```
aws opensearch list-scheduled-actions \
  --domain-name my-domain
```

Respuesta:

```
{
    "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",
        }
    ]
}
```

Reprogramación de acciones

OpenSearch El servicio le notifica las actualizaciones programadas del software del servicio y las optimizaciones de Auto-Tune. Puede optar por aplicar el cambio de inmediato o reprogramarlo para una fecha y hora posteriores.

> **Note**
>
> OpenSearch El servicio puede programar la acción una hora después de la hora que usted seleccione. Por ejemplo, si decide aplicar una actualización a las 17:00 horas, puede aplicarla entre las 17:00 y las 18:00.

Consola

Cómo reprogramar una acción

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Seleccione el nombre del dominio para abrir la configuración.
3. Acceda a la pestaña del intervalo de menor demanda.
4. En Acciones programadas, seleccione la acción y elija Reprogramar.
5. Seleccione una de las siguientes opciones:

   - Aplicar la actualización ahora: programa de inmediato la acción para que se lleve a cabo en la hora actual si hay capacidad disponible. Si no hay capacidad disponible, ofrecemos otras franjas horarias disponibles para elegir.
   - Programarla en un intervalo de menor demanda: marca la acción para que se retome durante un próximo intervalo de menor demanda. No hay garantía de que el cambio se implemente durante el intervalo inmediatamente siguiente. Dependiendo de la capacidad, podría producirse en los días siguientes.
   - Reprogramar esta actualización: permite especificar una fecha y hora personalizadas para aplicar el cambio. Si la hora que especifica no está disponible por motivos de capacidad, puede seleccionar una franja horaria diferente.
   - Cancelar actualización programada: cancela la actualización. Esta opción solo está disponible para las actualizaciones de software de servicio opcionales. No está disponible para las acciones de ajuste automático ni para las actualizaciones de software obligatorias.
6. Seleccione Guardar cambios.

CLI

Para reprogramar una acción mediante el AWS CLI, envía una UpdateScheduledActionsolicitud. Para recuperar el ID de la acción, envíe una solicitud `ListScheduledActions`.

La siguiente solicitud reprograma una actualización del software de servicio para una fecha y hora específicas:

```
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
```

Respuesta:

```
{
   "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"
   }
}
```

Si la solicitud falla con un`SlotNotAvailableException`, significa que la hora que especificaste no está disponible por motivos de capacidad y debes especificar una hora diferente. OpenSearch El servicio proporciona sugerencias alternativas de franjas horarias disponibles en la respuesta.

Migración desde los intervalos de mantenimiento de ajuste automático

Si un dominio se creó antes del 16 de febrero de 2023, podría utilizar los períodos de mantenimiento para programar las optimizaciones de Auto-Tune que requieran una blue/green implementación. Puede migrar los dominios de ajuste automático existentes para utilizar en su lugar el intervalo de menor demanda.

> **Note**
>
> No puede volver a usar los intervalos de mantenimiento después de migrar su dominio para usar los intervalos de menor demanda.

Consola

Cómo migrar un dominio para utilizar el intervalo de menor demanda

1. En la consola OpenSearch de Amazon Service, selecciona el nombre del dominio para abrir su configuración.
2. Vaya a la pestaña Ajuste automático y seleccione Editar.
3. Seleccione Migrar al intervalo de menor demanda.
4. En Hora de inicio (UTC), indique una hora de inicio diaria para el intervalo de menor demanda en el tiempo universal coordinado (UTC).
5. Seleccione Guardar cambios.

CLI

Para migrar de un período de mantenimiento de Auto-Tune a un período de menor actividad mediante el AWS CLI, envíe una UpdateDomainConfigsolicitud:

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --auto-tune-options
 DesiredState=ENABLED,UseOffPeakWindow=true,MaintenanceSchedules=[]
```

El intervalo de menor demanda debe estar activado para poder migrar un dominio del intervalo de mantenimiento de ajuste automático al intervalo de menor demanda. Puede activar el intervalo de menor demanda en una solicitud independiente o en la misma solicitud. Para obtener instrucciones, consulte the section called “Habilitación del intervalo de menor demanda”.

Notificaciones en Amazon OpenSearch Service

Las notificaciones de Amazon OpenSearch Service contienen información importante sobre el rendimiento y el estado de tus dominios. OpenSearch El servicio le notifica sobre las actualizaciones del software del servicio, las mejoras de Auto-Tune, los problemas de estado del clúster y los

errores de dominio. Las notificaciones están disponibles para todas las versiones de Elasticsearch OpenSearch OSS.

Puede ver las notificaciones en el panel de notificaciones de la consola de OpenSearch servicio. Todas las notificaciones OpenSearch del Servicio también aparecen en Amazon EventBridge. Para obtener una lista completa de las notificaciones y los eventos de muestra, consulte the section called "Supervisión de eventos".

Introducción a las notificaciones

Las notificaciones se habilitan automáticamente al crear un dominio. Ve al panel de notificaciones de la consola de OpenSearch servicio para supervisar y confirmar las notificaciones. Cada notificación incluye información, como la hora en que se publicó, el dominio con el que se relaciona, un nivel de severidad y estado y una breve explicación. Puede ver notificaciones históricas de hasta 90 días en la consola.

Después de acceder al panel Notifications (Notificaciones) o de confirmar la recepción de una notificación, es posible que reciba un mensaje de error en el que se indique que no tiene permisos para realizar `es:ListNotificationsV2` o `es:UpdateNotificationStatus`. Para resolver este problema, otorgue a su usuario o rol los siguientes permisos en IAM:

JSON

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

La consola de IAM genera el error "IAM does not recognize one or more actions" ("IAM no reconoce una o varias acciones") que se puede ignorar sin riesgo. También puede restringir la acción `es:UpdateNotificationStatus` en determinados dominios. Para obtener más información, consulte the section called "Referencia de los elementos de las políticas".

Severidad de las notificaciones

Las notificaciones del OpenSearch Servicio pueden ser informativas, relacionadas con cualquier acción que ya hayas realizado o con las operaciones de tu dominio, o procesables, ya que requieren que tomes medidas específicas, como la aplicación de un parche de seguridad obligatorio. Cada notificación tiene una severidad asociada, que puede ser `Informational`, `Low`, `Medium`, `High` o `Critical`. En la siguiente tabla se resumen las severidades:

Gravedad	Description (Descripción)	Ejemplos
`Informational`	Información relacionada con la operación del dominio.	• Actualización del software de servicio disponible • Ajuste automático iniciado
`Low`	Una acción recomendada, pero que no tiene ningún impacto negativo en la disponibilidad o el rendimiento del dominio si no se la realiza.	• Ajuste automático cancelado • Advertencia de alto recuento de fragmentos
`Medium`	Puede haber una consecuencia si no se realiza la acción recomendada, pero cuenta con una ventana de tiempo extendida para la acción que se va a realizar.	• Error al actualizar el software del servicio • Límite de recuento de fragmentos superado
`High`	Es necesario realizar acciones urgentes para evitar efectos adversos.	• Se requiere actualización del software de servicio • Clave KMS inaccesible
`Critical`	Se requiere una acción inmediata para evitar efectos adversos, o para recuperarse de ellos.	Ninguno disponible

Ejemplo de evento EventBridge

El siguiente ejemplo muestra un evento OpenSearch de notificación de servicio enviado a Amazon EventBridge. La notificación tiene una severidad `Informational` porque la actualización es opcional:

```
{
  "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."
  }
}
```

Configuración de un dominio Multi-AZ en Amazon Service OpenSearch

Para evitar la pérdida de datos y minimizar el tiempo de inactividad del clúster de Amazon OpenSearch Service en caso de una interrupción del servicio, puede distribuir los nodos en dos o tres zonas de disponibilidad de la misma región, una configuración conocida como Multi-AZ. Las zonas de disponibilidad son ubicaciones aisladas dentro de cada AWS región.

Para los dominios que ejecutan cargas de trabajo de producción, recomendamos la opción de implementación de Multi-AZ con modo de espera, que crea la siguiente configuración:

- El dominio implementado en tres zonas.
- Tipo de instancia de la generación actual para los nodos maestros dedicados y los nodos de datos.
- Tres nodos maestros dedicados y tres (o un múltiplo de tres) nodos de datos.
- Al menos dos réplicas para cada índice de su dominio o un múltiplo de tres copias de datos (incluidos los nodos principales y las réplicas).

El resto de esta sección contiene explicaciones y contextos sobre estas configuraciones.

Multi-AZ con modo de espera

Multi-AZ with Standby es una opción de implementación para los dominios de Amazon OpenSearch Service que ofrece una disponibilidad del 99,99%, un rendimiento uniforme para las cargas de trabajo de producción y una configuración y administración de dominios simplificadas. Cuando utilice Multi-AZ con modo de espera, los dominios son resistentes a los fallos de infraestructura, sin que ello afecte al rendimiento ni a la disponibilidad. Esta opción de implementación cumple con este estándar al exigir una serie de prácticas recomendadas, como un recuento específico de nodos de datos, un recuento de nodos maestros, un tipo de instancia, un recuento de réplicas, la configuración de las actualizaciones de software y la activación del ajuste automático.

> **Note**
>
> La opción Multi-AZ con modo de espera está disponible para OpenSearch la versión 1.3 y versiones posteriores, y requiere regiones con al menos tres zonas de disponibilidad.

Al utilizar Multi-AZ con modo de espera, el OpenSearch servicio crea un dominio en tres zonas de disponibilidad, cada una de las cuales contiene una copia completa de los datos y los datos se distribuyen equitativamente en cada una de las zonas. Su dominio reserva los nodos de una de estas zonas como en espera, lo que significa que no atienden solicitudes de búsqueda. Cuando el OpenSearch servicio detecta un fallo en la infraestructura subyacente, activa automáticamente los nodos en espera en menos de un minuto. El dominio sigue atendiendo las solicitudes de indexación y búsqueda, y cualquier impacto se limita al tiempo que se tarda en realizar la conmutación por error. No hay redistribución de los datos o los recursos, por lo que el rendimiento del clúster no se ve afectado y no hay riesgo de que se degrade la disponibilidad. Multi-AZ con modo de espera está disponible sin costo adicional.

Tiene dos opciones para crear un dominio con el modo de espera en Consola de administración de AWS. En primer lugar, puede crear un dominio con el método de creación fácil y OpenSearch Service utilizará automáticamente una configuración predeterminada, que incluye lo siguiente:

- Tres zonas de disponibilidad, una de las cuales actúa como zona de espera
- Tres nodos maestros y de datos dedicados
- El ajuste automático está habilitado en el dominio
- GP3 almacenamiento para los nodos de datos

También puede elegir el método de Creación estándar y seleccionar Dominio con modo de espera como opción de implementación. Esto le permite personalizar su dominio sin dejar de utilizar las características clave del modo de espera, como tres zonas y tres nodos maestros. Recomendamos elegir un recuento de nodos de datos que sea múltiplo de tres (el número de zonas de disponibilidad).

Una vez que haya creado su dominio, puede ir a las páginas de detalles del dominio y, en la pestaña Configuración del clúster, confirmar que en las zonas de disponibilidad aparezca 3-AZ con modo de espera.

Si tiene problemas para migrar un dominio existente a Multi-AZ con modo de espera, consulte Error migrating to Multi-AZ with Standby en la guía de solución de problemas.

Limitaciones

Al configurar un dominio con Multi-AZ con modo de espera, tenga en cuenta las siguientes limitaciones:

- El número total de particiones de un nodo no puede superar los 1000, el número total de particiones de un clúster no puede superar los 75 000 y el tamaño de una sola partición no puede superar los 65 GB.
- Multi-AZ con modo de espera solo funciona con los tipos de instancia `m5` `c5` `r5r6g,c6g,m6g,r6gd,,r7g,c7g,m7g,r7gd`, y de `i3` instancia. Para obtener más información sobre los tipos de instancias admitidos, consulte Tipos de instancias admitidos.
- Solo puedes usar una SSD aprovisionada, una IOPs SSD de uso general (GP3) o un almacenamiento respaldado por instancias con capacidad de espera.
- Si lo habilita UltraWarmen un dominio Multi-AZ con modo de espera, la cantidad de nodos en espera debe ser un múltiplo de la cantidad de zonas de disponibilidad que se estén utilizando.

Multi-AZ sin modo de espera

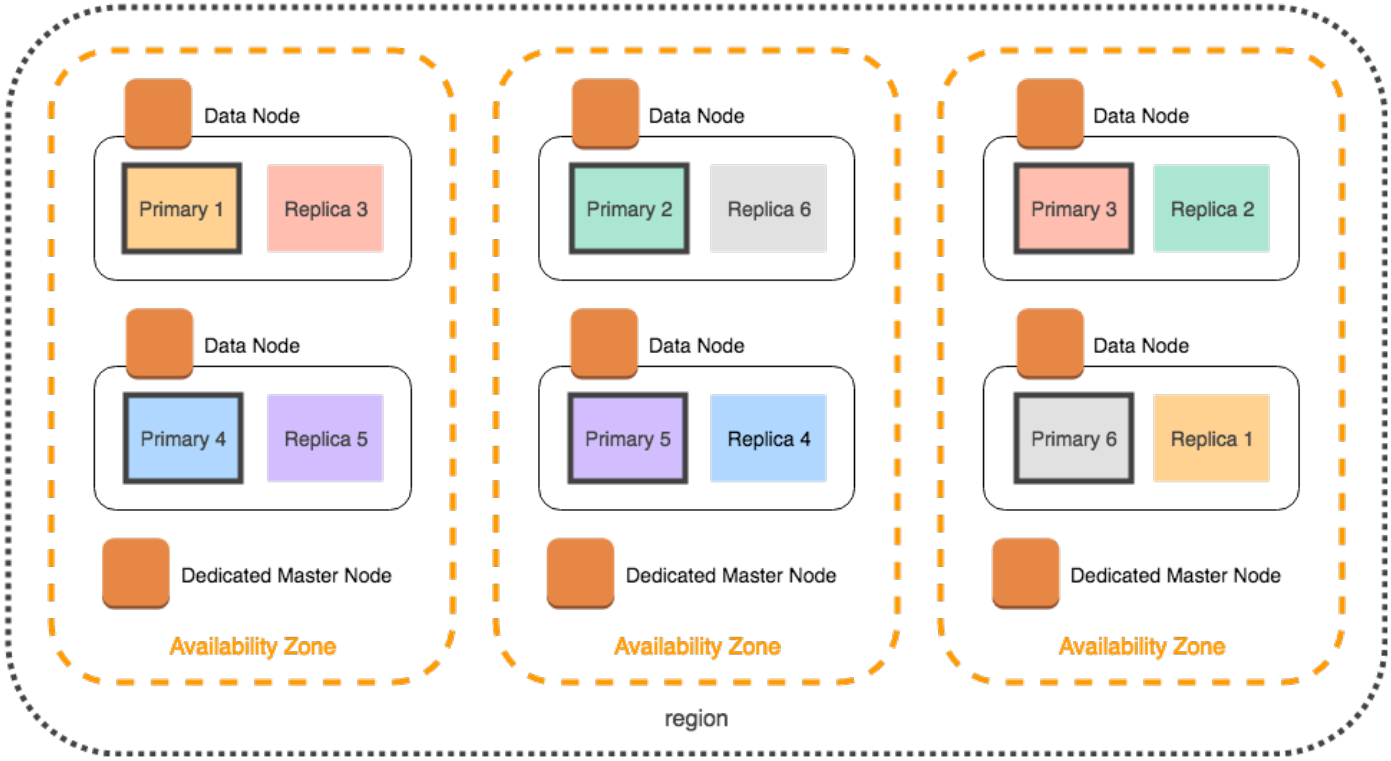
OpenSearch El servicio sigue siendo compatible con Multi-AZ sin modo de espera, lo que ofrece una disponibilidad del 99,9%. Los nodos se distribuyen en las zonas de disponibilidad y la disponibilidad depende del número de zonas de disponibilidad y de las copias de los datos. Mientras que en el modo de espera hay que configurar el dominio según las prácticas recomendadas, sin el modo de espera puede elegir su propio número de zonas de disponibilidad, nodos y réplicas. No recomendamos esta opción a menos que tenga flujos de trabajo existentes que se verían interrumpidos al crear dominios con el modo de espera.

Si elige esta opción, le recomendamos que seleccione tres zonas de disponibilidad para seguir siendo resistente a los errores de nodos, discos y zonas de disponibilidad única. Cuando se produce un error, el clúster redistribuye los datos entre los recursos restantes para mantener la disponibilidad y la redundancia. Este movimiento de datos aumenta el uso de recursos en el clúster y puede afectar al rendimiento. Si el clúster no tiene el tamaño adecuado, puede experimentar una disminución de la disponibilidad, lo que en gran medida va en contra del propósito de las zonas de disponibilidad múltiples.

La única forma de configurar un dominio sin modo de espera Consola de administración de AWS es elegir el método de creación estándar y seleccionar Dominio sin modo de espera como opción de despliegue.

Distribución de particiones

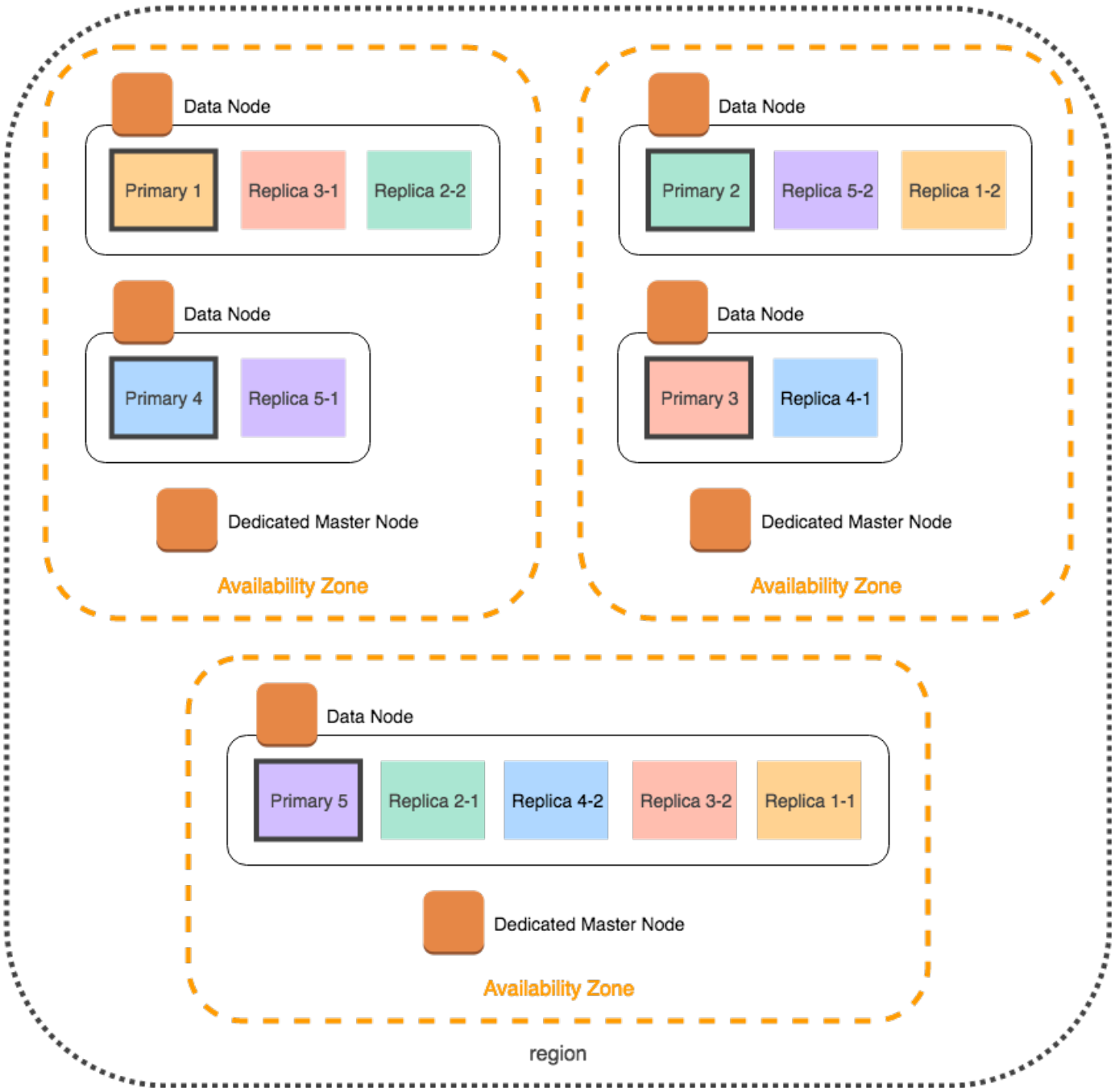
Si habilita Multi-AZ sin modo de espera, debe haber al menos una réplica por cada índice del clúster. Sin réplicas, OpenSearch Service no puede distribuir copias de sus datos a otras zonas de disponibilidad. Afortunadamente, la configuración predeterminada para cualquier índice es un recuento de réplicas de 1. Como se muestra en el siguiente diagrama, OpenSearch Service hace todo lo posible por distribuir los fragmentos principales y sus correspondientes fragmentos de réplica en distintas zonas.



Además de distribuir los fragmentos por zona de disponibilidad, OpenSearch Service los distribuye por nodo. Sin embargo, ciertas configuraciones de dominio pueden generar un número desequilibrado de particiones. Imagine el siguiente dominio:

- 5 nodos de datos
- 5 particiones principales
- 2 réplicas
- 3 zonas de disponibilidad

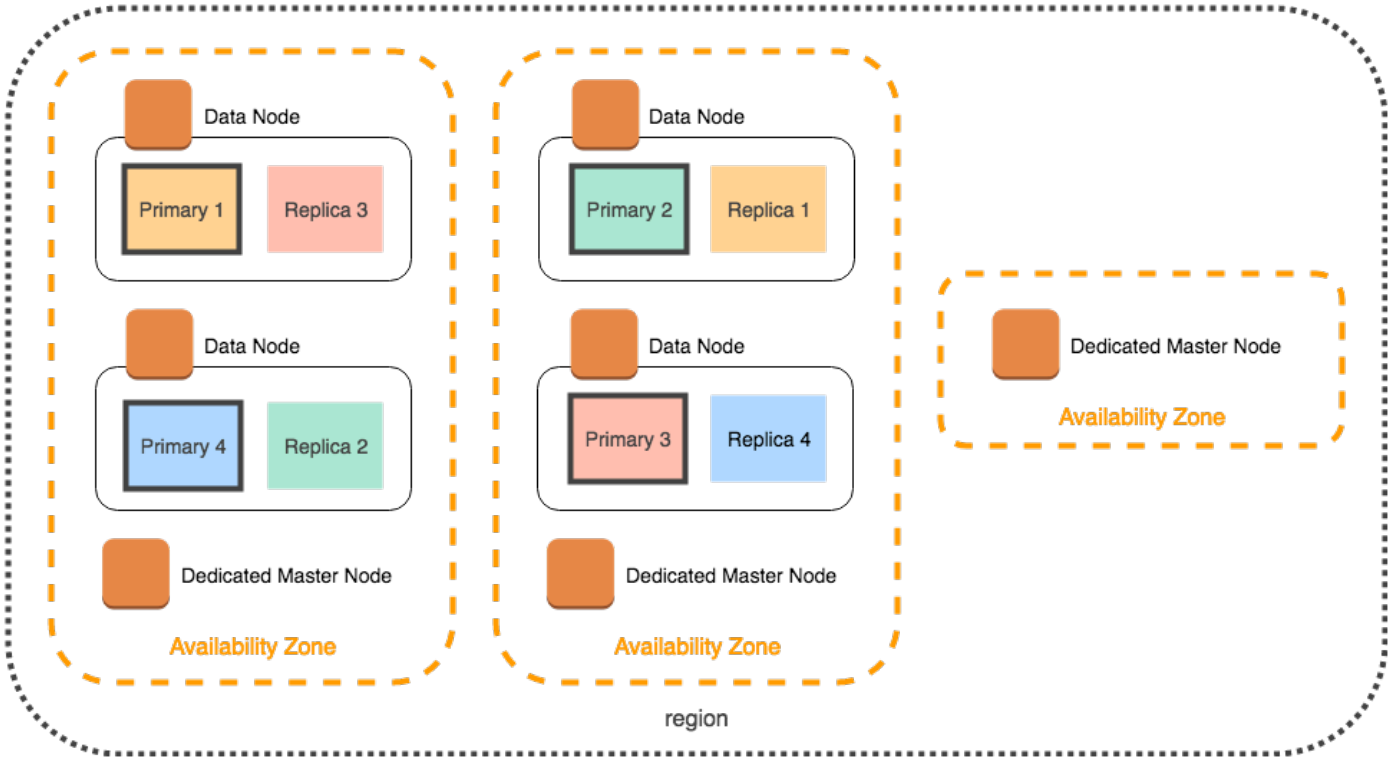
En esta situación, el OpenSearch servicio tiene que sobrecargar un nodo para poder distribuir los fragmentos principales y de réplica entre las zonas, como se muestra en el siguiente diagrama.



Para evitar este tipo de situaciones, que pueden ejercer demasiada presión sobre determinados nodos y perjudicar el rendimiento, le recomendamos que elija multi-AZ con modo de espera o que elija un número de instancias que sea múltiplo de tres cuando planee tener dos o más réplicas por cada índice.

Distribución de nodos maestros dedicados

Incluso si selecciona dos zonas de disponibilidad al configurar su dominio, OpenSearch Service distribuye automáticamente los nodos principales dedicados en tres zonas de disponibilidad. Esta distribución ayuda a evitar el tiempo de inactividad del clúster si una zona experimenta una interrupción del servicio. Si está utilizando los tres nodos principales dedicados recomendados y una zona de disponibilidad deja de funcionar, el clúster seguirá teniendo un quorum (2) de nodos principales dedicados y podrá seleccionar un nuevo nodo principal. En el siguiente diagrama, se muestra esta configuración.



Si selecciona un tipo de instancia de una generación anterior que no esté disponible en tres zonas de disponibilidad, pueden producirse las siguientes situaciones:

- Si ha elegido tres zonas de disponibilidad para el dominio, el OpenSearch servicio generará un error. Seleccione un tipo de instancia distinto y vuelva a intentarlo.
- Si elige dos zonas de disponibilidad para el dominio, OpenSearch Service distribuye los nodos principales dedicados en dos zonas.

Interrupción de las zonas de disponibilidad

No es frecuente que el servicio se interrumpa en las zonas de disponibilidad, pero puede ocurrir. En la siguiente tabla, se muestran diferentes configuraciones de Multi-AZ y los comportamientos durante una interrupción. La última fila de la tabla se aplica a Multi-AZ con modo de espera, mientras que todas las demás filas tienen configuraciones que solo se aplican a Multi-AZ sin modo de espera.

Número de zonas de disponibilidad de una región	Número de zonas de disponibilidad seleccionado	Número de nodos maestros dedicados	Comportamiento si se interrumpe una zona de disponibilidad
2 o más	2	0	Tiempo de inactividad. El clúster pierde la mitad de los nodos de datos y debe sustituir al menos uno en la zona de disponibilidad que queda antes de poder elegir un nodo principal.
2	2	3	Posibilidad de que se produzca un tiempo de inactividad del 50/50. OpenSearch El servicio distribuye dos nodos maestros dedicados en una zona de disponibilidad y uno en la otra: • Si la zona de disponibilidad con un nodo principal dedicado se interrumpe, los dos nodos principales dedicados de la otra zona de disponibilidad pueden elegir un nodo principal. • Si la zona de disponibilidad con dos nodos maestros dedicados experimenta una interrupción, el clúster no estará disponible hasta que la zona de disponibilidad restante se recupere.
3 o más	2	3	Sin tiempo de inactividad. OpenSearch El servicio distribuye automáticamente los nodos maestros dedicados en tres zonas de disponibilidad, de modo que los dos nodos maestros

Número de zonas de disponibilidad de una región	Número de zonas de disponibilidad seleccionado	Número de nodos maestros dedicados	Comportamiento si se interrumpe una zona de disponibilidad
			dedicados restantes pueden elegir un nodo maestro.
3 o más	3	0	No hay tiempo de inactividad. Aproximad amente dos tercios de los nodos de datos seguirán estando disponibles para elegir un nodo principal.
3 o más	3	3	No hay tiempo de inactividad. Los otros dos nodos principales dedicados pueden elegir un nodo principal.

En todas las configuraciones, independientemente de la causa, las fallas de los nodos pueden provocar que los nodos de datos restantes del clúster experimenten un período de mayor carga, mientras que OpenSearch Service configura automáticamente nuevos nodos para reemplazar los que ahora faltan.

Por ejemplo, en caso de que se interrumpa una zona de disponibilidad de una configuración de tres zonas, solo dos tercios de los nodos de datos tendrán que procesar tantas solicitudes como lleguen al clúster. A medida que se procesen estas solicitudes, los nodos restantes irán replicando particiones en los nuevos nodos tan pronto como se conecten, lo que podría afectar aún más al rendimiento. Si la disponibilidad es esencial para la carga de trabajo, considere la posibilidad de agregar más recursos al clúster para mitigar este problema.

> **Note**
>
> OpenSearch El servicio administra los dominios Multi-AZ de forma transparente, por lo que no es posible simular manualmente las interrupciones en las zonas de disponibilidad.

Lanzamiento de tus dominios OpenSearch de Amazon Service dentro de una VPC

Puede lanzar AWS recursos, como dominios de Amazon OpenSearch Service, en una nube privada virtual (VPC). Una VPC es una red virtual dedicada a usted. Cuenta de AWS Está aislada de forma lógica de otras redes virtuales de la AWS nube. La ubicación de un dominio de OpenSearch servicio dentro de una VPC permite una comunicación segura entre el OpenSearch Servicio y otros servicios de la VPC sin necesidad de una puerta de enlace a Internet, un dispositivo NAT o una conexión VPN. Todo el tráfico permanece seguro en la AWS nube.

> **Note**
>
> Si coloca su dominio OpenSearch de servicio en una VPC, su equipo debe poder conectarse a la VPC. Esta conexión a menudo adopta la forma de VPN, transit gateway, red administrada o servidor proxy. No puede acceder directamente a sus dominios desde fuera de la VPC.

VPC frente a dominios públicos

Las siguientes son algunas de las formas en que los dominios de VPC difieren de los dominios públicos. Cada diferencia se describe más adelante con más detalle.

- Debido a su aislamiento lógico, los dominios que residen dentro de una VPC tienen una capa adicional de seguridad en comparación con los dominios que utilizan puntos de conexión públicos.
- Aunque se puede acceder a los dominios públicos desde cualquier dispositivo conectado a Internet, los dominios de VPC requieren algún tipo de VPN o proxy.
- En comparación con los dominios públicos, los dominios VPC muestran menos información en la consola de . En concreto, la pestaña Cluster health (Estado del clúster) no incluye información de particiones y la pestaña Indices (Índices) no aparece.
- Los puntos de conexión del dominio adoptan formas diferentes (`https://search-domain-name` frente a `https://vpc-domain-name`).
- No puede aplicar estas políticas de acceso basado en IP a los dominios que residan dentro de una VPC porque los grupos de seguridad ya aplican políticas de acceso basadas en IP.

Limitaciones

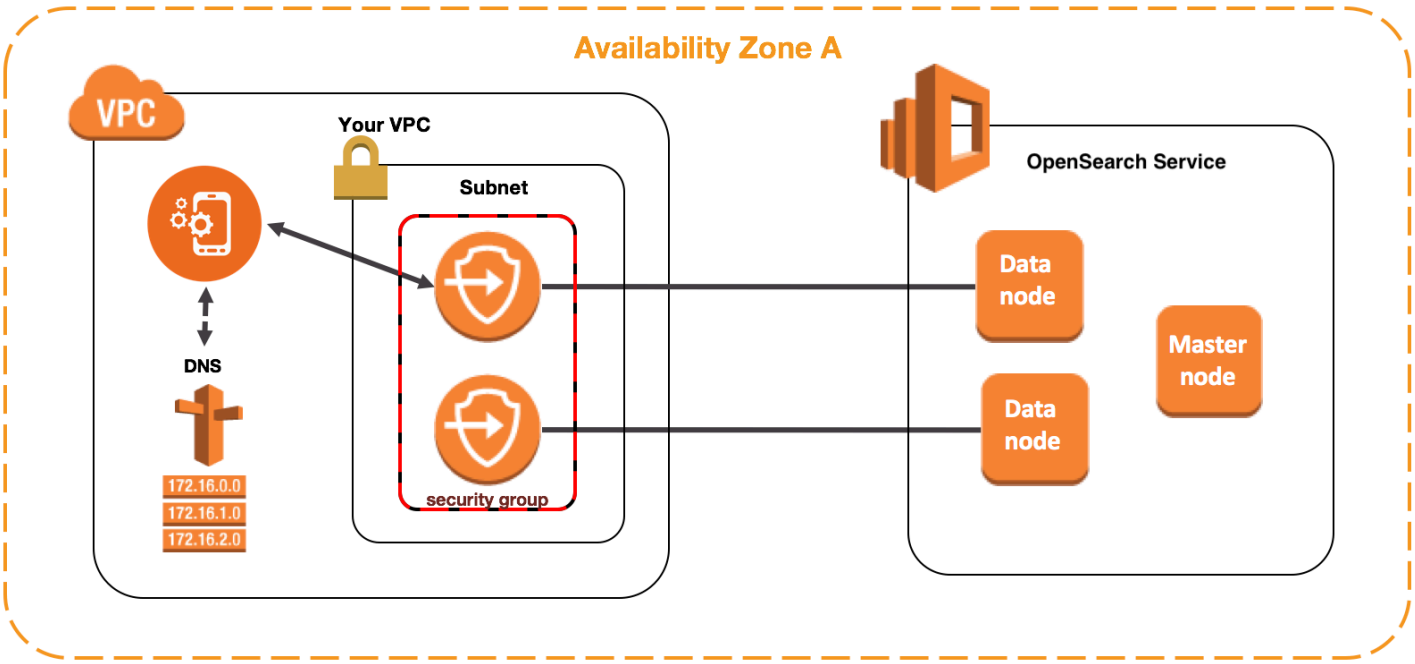
La operación de un dominio de OpenSearch servicio dentro de una VPC tiene las siguientes limitaciones:

- Si se lanza un dominio en una VPC, no se puede cambiar posteriormente para que utilice un punto de conexión público. Lo contrario también es cierto: si se crea un dominio con un punto de conexión público, no es posible situarlo posteriormente en una VPC. En lugar de ello, se debe crear un dominio nuevo y migrar los datos.
- Es posible lanzar el dominio dentro de una VPC o utilizar un punto de conexión público, pero no es posible hacer las dos cosas. Se debe elegir una de ellas al crear un dominio.
- No puede lanzar el dominio en una VPC que utilice tenencia dedicada. Debe utilizar una VPC con la tenencia establecida en Default (Predeterminada).
- Después de situar un dominio dentro de una VPC, no se puede mover a otra VPC, pero puede cambiar la configuración de subredes y grupos de seguridad.
- Para acceder a la instalación predeterminada de los OpenSearch paneles de control de un dominio que reside en una VPC, los usuarios deben tener acceso a la VPC. Este proceso depende de la configuración de la red, pero probablemente implica conectarse a una VPN o a una red gestionada, o utilizar un servidor proxy o transit gateway. Para obtener más información, consulte the section called “Acerca de las políticas de acceso en los dominios de VPC”, la Guía del usuario de Amazon VPC y the section called “Control del acceso a Dashboards”.

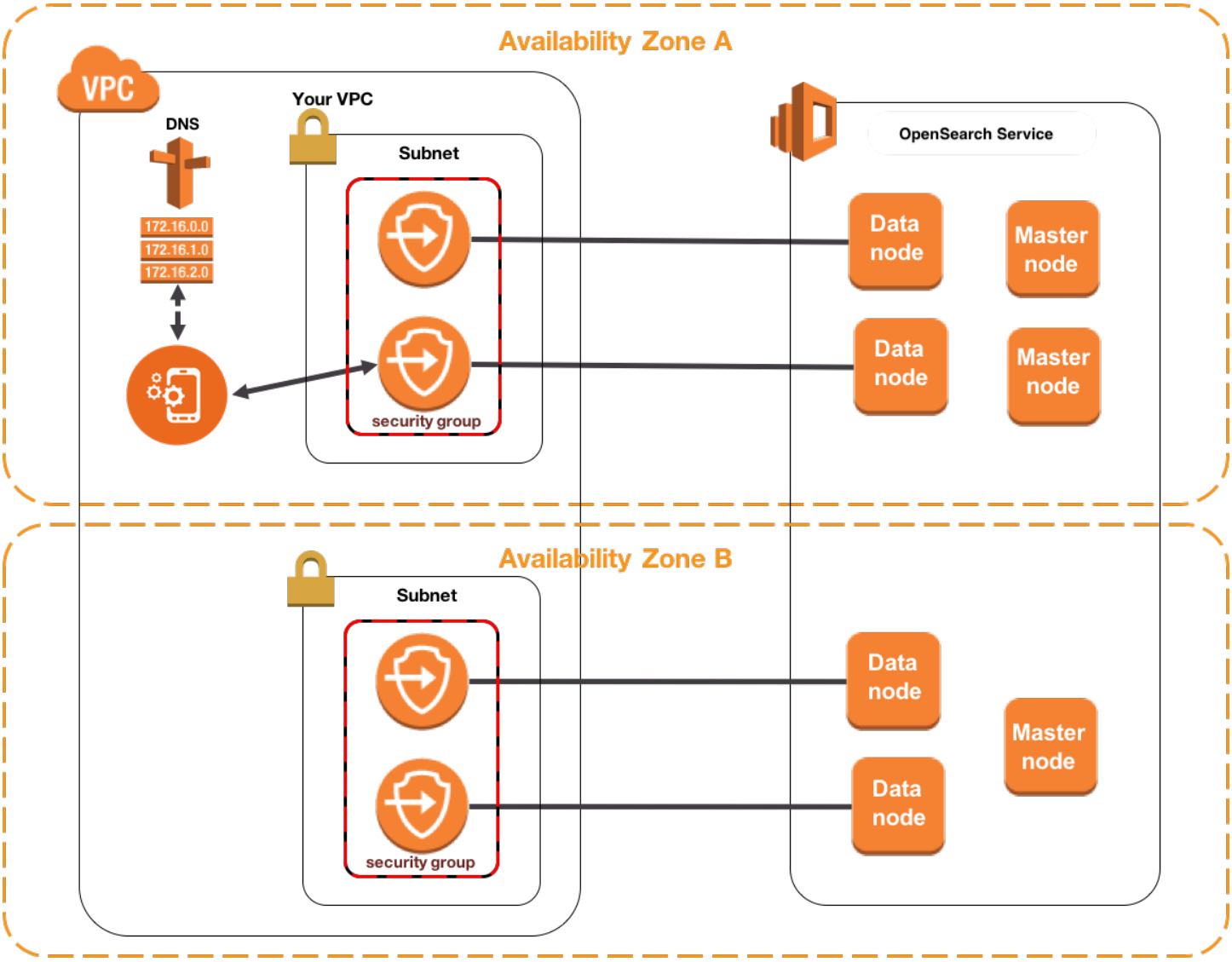
Arquitectura

Para brindar soporte VPCs, OpenSearch Service coloca un punto final en una, dos o tres subredes de su VPC. Si habilita varias zonas de disponibilidad para su dominio, cada subred debe estar en una zona de disponibilidad diferente de la misma región. Si solo usa una zona de disponibilidad, el OpenSearch servicio coloca un punto final en una sola subred.

En la siguiente ilustración, se muestra la arquitectura de una VPC para una zona de disponibilidad.



En la siguiente ilustración, se muestra la arquitectura de una VPC para dos zonas de disponibilidad.



OpenSearch El servicio también coloca una interface de red elástica (ENI) en la VPC para cada uno de los nodos de datos. OpenSearch El servicio asigna a cada ENI una dirección IP privada del rango de IPv4 direcciones de su subred. El servicio también asigna un nombre de host de DNS pública (que es el punto de conexión del dominio) para las direcciones IP. Es necesario utilizar un servicio DNS público para resolver el punto de conexión (que es un nombre de host DNS) a las direcciones IP adecuadas de los nodos de datos:

- Si su VPC utiliza el servidor DNS proporcionado por Amazon configurando la `enableDnsSupport` opción en `true` (el valor predeterminado), la resolución del punto final del OpenSearch servicio se realizará correctamente.

- Si la VPC usa un servidor DNS privado y el servidor puede acceder a los servidores DNS públicos autorizados para resolver los nombres de host DNS, la resolución del punto final del OpenSearch servicio también se realizará correctamente.

Dado que las direcciones IP podrían cambiar, conviene resolver periódicamente el punto de conexión del dominio para no dejar de tener acceso a los nodos de datos correctos. Recomendamos establecer el intervalo de resolución de DNS en un minuto. Si se utiliza un cliente, también conviene asegurarse de que se borra la caché de DNS del cliente.

Migración del acceso público al acceso mediante VPC

Cuando se crea un dominio, se debe especificar si debe tener un punto de conexión público o residir dentro de una VPC. Una vez creado, no se puede cambiar de un tipo a otro. Lo que hay que hacer es crear un dominio nuevo y migrar los datos o reindexarlo manualmente. Las instantáneas son un método cómodo para migrar los datos. Para obtener información sobre cómo tomar instantáneas y restaurarlas, consulte the section called “Crear instantáneas de índice”.

Acerca de las políticas de acceso en los dominios de VPC

Colocar su dominio de OpenSearch servicio dentro de una VPC proporciona una capa de seguridad sólida e inherente. Cuando crea un dominio con el acceso público, el punto de conexión adopta la siguiente forma:

```
https://search-domain-name-identifier.region.es.amazonaws.com
```

Como sugiere la etiqueta “público”, esto es punto de conexión es accesible desde cualquier dispositivo conectado a Internet, aunque puede (y debería) controlar el acceso al mismo. Si accede al punto de conexión en un navegador web, es posible que reciba un mensaje `Not Authorized`, pero la solicitud llega al dominio.

Cuando crea un dominio con acceso VPC, el punto de enlace parece similar a un punto de conexión público:

```
https://vpc-domain-name-identifier.region.es.amazonaws.com
```

Si intenta acceder al punto de conexión en un navegador web, sin embargo, podría encontrar que la solicitud agota el tiempo de espera. Para realizar solicitudes `GET` incluso básicas, su equipo debe poder conectarse a la VPC. Esta conexión a menudo adopta la forma de VPN, transit gateway, red administrada o servidor proxy. Para obtener más información sobre las distintas formas puede

adoptar, consulte Ejemplos de VPC en la Guía del usuario del Amazon VPC. Para un ejemplo de desarrollo detallado, consulte the section called “Probar los dominios de VPC”.

Además de este requisito de conectividad, le VPCs permite administrar el acceso al dominio a través de grupos de seguridad. Para muchos casos de uso, esta combinación de características de seguridad es suficiente, y es posible que se sienta cómodo aplicando una política de acceso abierto al dominio.

Operar con una política de acceso abierto no significa que cualquier usuario de Internet pueda acceder al dominio del OpenSearch Servicio. Más bien, significa que si una solicitud llega al dominio del OpenSearch servicio y los grupos de seguridad asociados lo permiten, el dominio la acepta. La única excepción es si se utiliza un control de acceso detallado o una política de acceso que especifique roles de IAM. En estas situaciones, para que el dominio acepte una solicitud, los grupos de seguridad deben permitirla y debe estar firmada con credenciales válidas.

> **Note**
>
> Como los grupos de seguridad ya aplican políticas de acceso basadas en IP, no puede aplicar políticas de acceso basadas en IP a los dominios de OpenSearch servicio que residen en una VPC. Si utiliza acceso público, las políticas basadas en IP siguen estando disponibles.

Antes de comenzar: Requisitos previos para el acceso mediante VPC

Antes de poder habilitar una conexión entre una VPC y su nuevo dominio de OpenSearch servicio, debe hacer lo siguiente:

- Creación de una VPC

  Para crear su VPC, puede utilizar la consola de Amazon VPC, la AWS CLI o una de las. AWS SDKs Para obtener más información, consulte Trabajar con VPCs en la Guía del usuario de Amazon VPC. Si ya tiene una VPC, puede omitir este paso.
- Reservar direcciones IP

  OpenSearch El servicio permite la conexión de una VPC a un dominio al colocar las interfaces de red en una subred de la VPC. Cada interfaz de red tiene asociada una dirección IP. Debe reservar un número suficiente de direcciones IP en la subred para las interfaces de red. Para obtener más información, consulte Reserva de direcciones IP en una subred de una VPC.

Probar los dominios de VPC

La seguridad mejorada de una VPC puede hacer que la conexión a su dominio y la ejecución de pruebas básicas se convierta en un reto. Si ya tienes un dominio de VPC de OpenSearch servicio y prefieres no crear un servidor VPN, prueba el siguiente proceso:

1. Para la política de acceso del dominio, elija Only use fine-grained access control (Utilizar únicamente control de acceso detallado). Siempre puede actualizar esta configuración después de finalizar las pruebas.
2. Cree una instancia Amazon EC2 de Amazon Linux en la misma VPC, subred y grupo de seguridad que su dominio de servicio. OpenSearch

   Dado que esta instancia es para fines de prueba con muy poco trabajo, elija un tipo de instancia económica como `t2.micro`. Asigne a la instancia una dirección IP pública y cree un nuevo par de claves o elija uno existente. Si crea una nueva clave, descárguela en su directorio `~/.ssh`.

   Para obtener más información acerca de la creación de instancias, consulte Introducción a las instancias de Amazon EC2 Linux.
3. Agregue un gateway de Internet a su VPC.
4. En la tabla de ruta de la VPC, agregue una nueva ruta. En Destination (Destino), especifique un bloque de CIDR que contenga la dirección IP pública de su equipo. En Target (Destino), especifique el gateway de Internet que acaba de crear.

   Por ejemplo, puede especificar solo `123.123.123.123/32` para su equipo o `123.123.123.0/24` para varios equipos.
5. Para el grupo de seguridad, especifique dos reglas de entrada:

Tipo	Protocolo	Rango de puertos	origen
SSH (22)	TCP (6)	22	*your-cidr-block*
HTTPS (443)	TCP (6)	443	*your-security-group-id*

   La primera regla le permite usar SSH en la instancia EC2. La segunda permite que la instancia EC2 se comunique con el dominio del OpenSearch servicio a través de HTTPS.
6. En el terminal ejecute el comando siguiente:

```
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
```

Este comando crea un túnel SSH que reenvía las solicitudes a https://localhost:9200 a su dominio de OpenSearch servicio a través de la instancia EC2. Al especificar el puerto 9200 en el comando, se simula una OpenSearch instalación local, pero utilice el puerto que desee. OpenSearch El servicio solo acepta conexiones a través del puerto 80 (HTTP) o 443 (HTTPS).

El comando no proporciona comentarios y se ejecuta de forma indefinida. Para detenerlo, presione `Ctrl + C`.

7. Navegue hasta https://localhost:9200/_dashboards/ en su navegador web. Es posible que tenga que reconocer una excepción de seguridad.

   Si lo desea, puede enviar solicitudes a https://localhost:9200 mediante curl, Postman o su lenguaje de programación favorito.

> **Tip**
>
> Si detecta errores de curl debido a que los certificados no coinciden, pruebe con la marca `--insecure`.

Reserva de direcciones IP en una subred de una VPC

OpenSearch El servicio conecta un dominio a una VPC mediante la colocación de las interfaces de red en una subred de la VPC (o en varias subredes de la VPC si se habilitan varias zonas de disponibilidad). Cada interfaz de red tiene asociada una dirección IP. Antes de crear el dominio de OpenSearch servicio, debe tener un número suficiente de direcciones IP disponibles en cada subred para dar cabida a las interfaces de red.

Esta es la fórmula básica: la cantidad de direcciones IP que el OpenSearch Servicio reserva en cada subred es tres veces la cantidad de nodos de datos, dividida por la cantidad de zonas de disponibilidad.

Ejemplos

- Si un dominio tiene diez nodos de datos en tres zonas de disponibilidad, el recuento de direcciones IP por cada subred será 9 * 3 / 3 = 9.

- Si un dominio tiene ocho nodos de datos en dos zonas de disponibilidad, el recuento de direcciones IP por cada subred será 8 * 3 / 2 = 12.
- Si un dominio tiene seis nodos de datos en una zona de disponibilidad, el recuento de direcciones IP por cada subred será 6 * 3 / 1 = 18.

Al crear el dominio, el OpenSearch Servicio reserva las direcciones IP, utiliza algunas para el dominio y reserva el resto para despliegues azules o verdes. Puede ver las interfaces de red y sus direcciones IP asociadas en la sección Network Interfaces (Interfaces de red) de la consola de Amazon EC2. La columna Descripción muestra el dominio del OpenSearch servicio al que está asociada la interfaz de red.

> **Tip**
>
> Se recomienda crear subredes dedicadas para las direcciones IP reservadas del OpenSearch Servicio. Mediante el uso de subredes dedicadas, evitará solapamientos con otras aplicaciones y servicios, y se asegurará de que podrá reservar direcciones IP adicionales si necesita escalar el clúster en el futuro. Para obtener más información, consulte Creación de una subred en la VPC.

También puedes considerar la posibilidad de aprovisionar nodos coordinadores dedicados para reducir la cantidad de reservas de direcciones IP privadas necesarias para tu dominio de VPC. OpenSearchconecta una interface de red elástica (ENI) a sus nodos coordinadores dedicados en lugar de a sus nodos de datos. Los nodos coordinadores dedicados suelen representar alrededor del 10 % del total de nodos de datos. Como resultado, se reservará un número menor de direcciones IP privadas para los dominios de VPC.

Roles vinculados a servicios para el acceso mediante VPC

Un rol vinculado a un servicio es un tipo único de rol de IAM; que delega permisos en un servicio para que pueda crear y administrar recursos en nombre del usuario. OpenSearch El servicio requiere un rol vinculado al servicio para acceder a la VPC, crear el punto final del dominio y colocar las interfaces de red en una subred de la VPC.

OpenSearch El servicio crea automáticamente el rol cuando usas la consola de OpenSearch servicio para crear un dominio dentro de una VPC. Para que esta creación automática se realice correctamente, es necesario disponer de permisos para la acción

`iam:CreateServiceLinkedRole`. Para obtener más información, consulte Permisos de roles vinculados a servicios en la Guía del usuario de IAM.

Una vez que OpenSearch Service haya creado el rol, podrá verlo (`AWSServiceRoleForAmazonOpenSearchService`) mediante la consola de IAM.

Para obtener más información acerca de los permisos de este rol y como eliminarlo, consulte the section called “Cómo utilizar roles vinculados a servicios”.

Creación de instantáneas de índices en Amazon Service OpenSearch

Las instantáneas de Amazon OpenSearch Service son copias de seguridad de los índices y el estado de un clúster. Estado incluye configuraciones de clúster, información de nodos, configuración de índices y asignación de partición.

OpenSearch Las instantáneas de servicio se presentan en las siguientes formas:

- Las instantáneas automatizadas son solo para la recuperación de clústeres. Puede utilizarlas para restaurar el dominio en caso de que se produzca un estado rojo del clúster o se pierdan datos. Para obtener más información, consulte Restauración de instantáneas a continuación. OpenSearch El servicio almacena las instantáneas automatizadas en un bucket de Amazon S3 preconfigurado sin coste adicional.
- Las instantáneas manuales son para la recuperación de clústeres o para el traslado de datos de un clúster a otro. Debe iniciar instantáneas manuales. Estas instantáneas se almacenan en su propio bucket de Amazon S3 y se aplican cargos estándar de S3. Si tiene una instantánea de un OpenSearch clúster autogestionado, puede utilizarla para migrar a un OpenSearch dominio de servicio. Para obtener más información, consulta Migración a Amazon OpenSearch Service.

Todos los dominios de OpenSearch servicio toman instantáneas automatizadas, pero la frecuencia varía de las siguientes maneras:

- En el caso de los dominios que ejecutan Elasticsearch 5.3 OpenSearch o versiones posteriores, OpenSearch Service toma instantáneas automatizadas cada hora y las conserva hasta 336 durante 14 días. Las instantáneas por hora son menos disruptivas, debido a su naturaleza progresiva. También proporcionan un punto de recuperación más reciente en caso de problemas del dominio.

- En el caso de los dominios que ejecutan Elasticsearch 5.1 o versiones anteriores, OpenSearch Service toma instantáneas automatizadas a diario durante la hora que especifiques, conserva hasta 14 de ellas y no conserva ningún dato de las instantáneas durante más de 30 días.

Si el clúster pasa al estado rojo, se produce un error en todas las instantáneas automatizadas mientras persiste el estado del clúster. Si no corrige el problema en dos semanas, puede perder de forma permanente los datos del clúster. Para ver los pasos de solución de problemas, consulte the section called “Estado rojo del clúster”.

Requisitos previos

Para crear instantáneas manualmente, debe trabajar con IAM y Amazon S3. Compruebe que cumple los siguientes requisitos previos antes de intentar tomar una instantánea:

Requisito previo	Description (Descripción)
Bucket de S3	Cree un depósito de S3 para almacenar las instantáneas manuales de su dominio de servicio. OpenSearch Para obtener instrucciones, consulte Creación de un bucket de propósito general en la Guía del usuario de Amazon Simple Storage Service. Recuerde el nombre del bucket para utilizarlo en los siguientes lugares: • La instrucción `Resource` de la política de IAM que se adjunta al rol de IAM • El cliente Python utilizado para registrar un repositorio de instantáneas (si utiliza este método) **Important** No aplique una regla de ciclo de vida de Amazon Glacier a este bucket. Las instantáneas manuales no admiten la clase de almacenamiento de Amazon Glacier.
rol de IAM	Cree un rol de IAM para delegar los permisos al OpenSearch Servicio. Para obtener instrucciones, consulte Creación de roles de IAM (consola) en la Guía del

Requisito previo	Description (Descripción)
	usuario de IAM. En el resto de este capítulo, este rol se denomina `TheSnapshotRole` . Adjuntar una política de IAM Adjunte la siguiente política a `TheSnapshotRole` para permitir el acceso al bucket de S3: `{ "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 /*" ] } ] }` Para obtener instrucciones sobre cómo adjuntar una política a un rol, consulte Agregar permisos de identidad de IAM (consola) en la Guía del usuario de IAM. Modificar la relación de confianza

Requisito previo	Description (Descripción)

	Edite la relación de confianza de `TheSnapshotRole` para especificar el OpenSearch servicio en la `Principal` declaración, como se muestra en el siguiente ejemplo:
	
	`{`
	`  "Version":"2012-10-17",`
	`  "Statement": [{`
	`  "Sid": "",`
	`  "Effect": "Allow",`
	`  "Principal": {`
	`  "Service": "es.amazonaws.com"`
	`  },`
	`  "Action": "sts:AssumeRole"`
	`  }]`
	
	`}`
	
	Para obtener instrucciones a fin de editar la relación de confianza, consulte Actualización de una política de confianza de rol en la Guía del usuario de IAM.

Requisito previo	Description (Descripción)

Permisos	Para registrar el repositorio de instantáneas, debe poder pasarlo `TheSnapshotRole` a OpenSearch Service. También necesita tener acceso a la acción `es:ESHttpPut`. Para conceder estos dos permisos, asocie la siguiente política al rol de IAM cuyas credenciales se utilicen para firmar la solicitud:
	
	`{ "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 /*" } ] }`
	
	Si su usuario o rol no tiene permisos de `iam:PassRole` para transferir `TheSnapshotRole`, puede que se produzca el siguiente error común cuando intente registrar un repositorio en el siguiente paso:
	
	`$ 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 "}`

Registrar un repositorio de instantáneas manuales

Debe registrar un repositorio de instantáneas en OpenSearch Service antes de poder realizar instantáneas de indexación manuales. Esta operación única requiere que firme la AWS solicitud con las credenciales a las que se permite el accesoTheSnapshotRole, tal y como se describe enthe section called “Requisitos previos”.

Paso 1: Asigne la función de captura de pantalla en los OpenSearch paneles (si utiliza un control de acceso detallado)

El control de acceso detallado presenta un paso adicional al registrar un repositorio. Incluso si utiliza autenticación HTTP básica para todos los demás fines, debe asignar el rol `manage_snapshots` al rol de IAM que tenga permisos `iam:PassRole` para transferir `TheSnapshotRole`.

1. Navegue hasta el complemento OpenSearch Dashboards de su dominio de servicio. OpenSearch Puedes encontrar el punto de conexión de Dashboards en el panel de control de tu dominio, en la consola de OpenSearch servicio.
2. En el menú principal, seleccione Seguridad, Roles y seleccione el rol manage_snapshots.
3. Seleccione Usuarios asignados, Administrar mapeo.
4. Agregue el ARN del rol que tenga permisos para transferir `TheSnapshotRole`. Coloca el rol en ARNs Roles de backend.

   ```
   arn:aws:iam::123456789123:role/role-name
   ```

5. Seleccione Asignar y confirme que el usuario o el rol aparecen en Usuarios asignados.

Paso 2: registrar un repositorio

La siguiente pestaña Instantáneas muestra cómo registrar un directorio de instantáneas. Para ver las opciones específicas para cifrar una instantánea manual y registrarla después de migrarla a un nuevo dominio, consulte las pestañas correspondientes.

Snapshots

Para registrar un repositorio de instantáneas, envíe una solicitud PUT al punto final del dominio OpenSearch de servicio. Puede usar curl, el cliente Python de muestra, Postman o algún otro método para enviar una solicitud firmada para registrar el repositorio de instantáneas. Tenga en

cuenta que no puede usar una solicitud PUT en la consola de OpenSearch Dashboards para registrar el repositorio.

La solicitud tiene el siguiente 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**
>
> Los nombres de repositorio no pueden comenzar por “cs-”. Además, no debe escribir en el mismo repositorio desde varios dominios. Solo un dominio debe tener acceso de escritura al repositorio.

Si el dominio reside en una nube privada virtual (VPC), la computadora debe estar conectada a la VPC para que la solicitud registre correctamente el repositorio de instantáneas. El acceso a una VPC depende de la configuración de red, pero probablemente implica conectarse a una VPN o a una red corporativa. Para comprobar que puede acceder al dominio del OpenSearch servicio, navegue `https://your-vpc-domain.region.es.amazonaws.com` en un navegador web y compruebe que recibe la respuesta JSON predeterminada.

Cuando su bucket de Amazon S3 se encuentre en un dominio Región de AWS que no sea su OpenSearch dominio, añada el parámetro `"endpoint": "s3.amazonaws.com"` a la solicitud.

Encrypted snapshots

Actualmente no puede usar claves AWS Key Management Service (KMS) para cifrar instantáneas manuales, pero puede protegerlas mediante el cifrado del lado del servidor (SSE).

Para activar el SSE con claves administradas de S3 para el bucket que utiliza como repositorio de instantáneas, agregue `"server_side_encryption": true` al bloque `"settings"` de la solicitud PUT. Para obtener más información, consulte Uso del cifrado del servidor con claves

administradas por Amazon S3 (SSE-S3) en la Guía del usuario de Amazon Simple Storage Service.

Como alternativa, puede usar AWS KMS claves para el cifrado del lado del servidor en el bucket de S3 que utiliza como repositorio de instantáneas. Si utiliza este enfoque, asegúrese de dar `TheSnapshotRole` permiso a la AWS KMS clave utilizada para cifrar el depósito de S3. Para más información, consulte Políticas de claves en AWS KMS.

Domain migration

El registro de un repositorio de instantáneas es una operación que se realiza una vez. Sin embargo, para migrar de un dominio a otro, debe registrar el mismo repositorio de instantáneas en el dominio antiguo y en el nuevo. El nombre del repositorio es arbitrario.

Tenga en cuenta las siguientes pautas al migrar a un nuevo dominio o registrar el mismo repositorio con varios dominios:

- Al registrar el repositorio en el nuevo dominio, agregue `"readonly": true` al bloque `"settings"` de la solicitud PUT. Esta configuración impide sobrescribir accidentalmente datos del dominio antiguo. Solo un dominio debe tener acceso de escritura al repositorio.
- Si migra datos a un dominio en una Región de AWS diferente (por ejemplo, desde un dominio antiguo y un bucket ubicado en us-east-2 a un nuevo dominio en us-west-2), reemplace `"region": "region"` por `"endpoint": "s3.amazonaws.com"` en la instrucción PUT y vuelva a enviar la solicitud.

Utilizar el cliente de Python de ejemplo

El cliente de Python es más fácil de automatizar que una simple solicitud HTTP y tiene una mejor reutilización. Si elige utilizar este método para registrar un repositorio de instantáneas, guarde el siguiente código de muestra de Python como archivo Python, por ejemplo `register-repo.py`. El cliente requiere los paquetes AWS SDK para Python (Boto3), requests y requests-aws4auth. El cliente contiene ejemplos de otras operaciones de instantánea que están marcados como comentarios.

Actualice las variables siguientes en el código de muestra: `host`, `region`, `path` y `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)
```

Tomar instantáneas manuales

Las instantáneas no son inmediatas. Tardan tiempo en completarse y no representan point-in-time vistas perfectas del clúster. Mientras que una instantánea está en curso, puede indexar los documentos y realizar otras solicitudes al clúster, pero los nuevos documentos y las actualizaciones de los existentes no suelen estar incluidos en la instantánea. La instantánea incluye los fragmentos principales tal como estaban cuando se OpenSearch inició la instantánea. En función del tamaño del grupo de subprocesos de la instantánea, se podrían incluir diferentes particiones en la instantánea a horas ligeramente diferentes. Para ver las prácticas recomendadas para instantáneas, consulte the section called “Cómo mejorar el rendimiento de las instantáneas”.

> **Warning**
>
> Las instantáneas manuales no incluyen los datos almacenados en los niveles de almacenamiento en frío UltraWarm ni en los niveles de almacenamiento en frío. Si su dominio utiliza UltraWarm almacenamiento en frío, migre esos índices a un almacenamiento activo antes de realizar una instantánea manual si necesita conservar esos datos.

Almacenamiento y rendimiento de instantáneas

OpenSearch las instantáneas son incrementales, lo que significa que solo almacenan los datos que han cambiado desde la última instantánea correcta. Esta naturaleza progresiva significa que la diferencia en la utilización de disco entre instantáneas frecuentes e infrecuentes suele ser mínima. En otras palabras, tomar instantáneas cada hora durante una semana (lo que equivale a un total de 168 instantáneas) no usaría mucho más espacio en disco que una sola instantánea al final de la semana. Además, cuanto mayor sea la frecuencia con la que toma las instantáneas, menor será el tiempo que lleve completarse. Por ejemplo, las instantáneas diarias pueden tardar entre 20 y 30 minutos en completarse, mientras que las instantáneas por hora pueden completarse en pocos minutos. Algunos OpenSearch usuarios toman instantáneas con una frecuencia de hasta media hora.

Tome una instantánea

Cuando crea una instantánea, especifica la siguiente información:

- El nombre del repositorio de instantáneas
- Un nombre de la instantánea

En los ejemplos que aparecen en este capítulo, se utiliza curl, un cliente HTTP común, por motivos de comodidad y brevedad. Para pasar un nombre de usuario y una contraseña a su solicitud de curl, consulte el tutorial de introducción.

Si las políticas de acceso especifican usuarios o roles, debe firmar las solicitudes de instantáneas. Para curl, puede usar la opción `--aws-sigv4` con la versión 7.75.0 o posterior. También puede utilizar los ejemplos comentados en el cliente Python de ejemplo para realizar solicitudes HTTP firmadas para los mismos puntos de conexión que utilizan los comandos curl.

Para realizar una instantánea manual, siga estos pasos:

1. No puede tomar una instantánea si hay alguna en curso. Para comprobarlo, ejecute el siguiente comando:

```
curl -XGET 'domain-endpoint/_snapshot/_status'
```

2. Ejecute el siguiente comando para tomar una instantánea manual:

```
curl -XPUT 'domain-endpoint/_snapshot/repository-name/snapshot-name'
```

   Para incluir o excluir ciertos índices y especificar otras configuraciones, agregue un cuerpo de la solicitud. Para conocer la estructura de la solicitud, consulte Tomar instantáneas en la OpenSearch documentación.

> **Note**
>
> El tiempo necesario para tomar una instantánea aumenta con el tamaño del dominio del OpenSearch servicio. Cuando una operación de instantánea tarda mucho tiempo en ejecutarse, en ocasiones se genera el siguiente error: `504 GATEWAY_TIMEOUT`. Normalmente, puede hacer caso omiso de estos errores y esperar a que la operación se complete correctamente. Ejecute el siguiente comando para comprobar el estado de todas las instantáneas del dominio:
>
> ```
> curl -XGET 'domain-endpoint/_snapshot/repository-name/_all?pretty'
> ```

Restaurar instantáneas

Antes de restaurar una instantánea, asegúrese de que el dominio de destino no utilice Multi-AZ con modo de espera. Si el modo de espera está activado, se produce un error en la operación de restauración.

> **Warning**
>
> Si utiliza alias de índice, debe detener las solicitudes de escritura a un alias, o cambiar el alias a otro índice, antes de borrar su índice. Al detener las solicitudes de escritura ayuda a evitar la siguiente situación:

1. Elimina un índice, de modo que también elimina su alias.
2. Una solicitud de escritura errante al alias que se ha eliminado crea un nuevo índice con el mismo nombre que el alias.
3. Ya no puede usar el alias debido a un conflicto de nombres con el nuevo índice. Si ha cambiado el alias a otro índice, especifique `"include_aliases": false` al realizar la restauración a partir de una instantánea.

Para restablecer una instantánea

1. Identifique la instantánea que desea restaurar. Asegúrese de que todos los ajustes de este índice, como los paquetes de analizadores personalizados o los ajustes de requisitos de asignación, sean compatibles con el dominio. Para ver todos los repositorios de instantáneas, ejecute el siguiente comando:

   ```
   curl -XGET 'domain-endpoint/_snapshot?pretty'
   ```

   Cuando identifique el repositorio, ejecute el siguiente comando para ver todas las instantáneas:

   ```
   curl -XGET 'domain-endpoint/_snapshot/repository-name/_all?pretty'
   ```

   > **Note**
   >
   > La mayoría de las instantáneas automatizadas se almacenan en el repositorio `cs-automated`. Si su dominio cifra datos en reposo, se almacenan en el repositorio `cs-automated-enc`. Si no ve el repositorio de instantáneas manuales que busca, asegúrese de que lo haya registrado en el dominio.

2. (Opcional) Elimine o cambie el nombre de uno o más índices del dominio de OpenSearch servicio si hay conflictos de nomenclatura entre los índices del clúster y los índices de la instantánea. No puede restaurar una instantánea de los índices en un OpenSearch clúster que ya contenga índices con los mismos nombres.

   Puede seleccionar las siguientes opciones si tiene conflictos de nomenclatura de índice:

   - Elimine los índices del dominio de OpenSearch servicio existente y, a continuación, restaure la instantánea.

- Cambie el nombre de los índices a medida que los restaure desde la instantánea y vuelva a indexarlos más tarde. Para obtener información sobre cómo cambiar el nombre de los índices, consulte esta solicitud de ejemplo en la documentación. OpenSearch
- Restaure la instantánea en un dominio de OpenSearch servicio diferente (solo es posible con instantáneas manuales).

El siguiente comando elimina todos los índices existentes en un dominio:

```
curl -XDELETE 'domain-endpoint/_all'
```

Sin embargo, si no piensa restaurar todos los índices, solo puede eliminar uno:

```
curl -XDELETE 'domain-endpoint/index-name'
```

3. Para restaurar una instantánea, ejecute el siguiente comando:

```
curl -XPOST 'domain-endpoint/_snapshot/repository-name/snapshot-name/_restore'
```

Debido a los permisos especiales en los OpenSearch paneles y a los índices de control de acceso detallados, los intentos de restaurar todos los índices pueden fallar, especialmente si se intenta restaurar desde una instantánea automática. En el siguiente ejemplo, se restaura un solo índice, `my-index`, a partir de `2020-snapshot` en el repositorio de instantáneas `cs-automated`:

```
curl -XPOST 'domain-endpoint/_snapshot/cs-automated/2020-snapshot/_restore' \
-d '{"indices": "my-index"}' \
-H 'Content-Type: application/json'
```

También puede que quiera restaurar todos los índices excepto los índices de control de acceso detallado y paneles:

```
curl -XPOST 'domain-endpoint/_snapshot/cs-automated/2020-snapshot/_restore' \
-d '{"indices": "-.kibana*,-.opendistro*"}' \
-H 'Content-Type: application/json'
```

> **Note**
>
> En función de la OpenSearch versión, es posible que también sea necesario excluir otros índices del sistema, como y. `-.opensearch-observability*` `-.plugins-ml-config*` La operación de restauración puede fallar si estos índices ya existen en el dominio de destino. Para excluirlos, agréguelos a la lista de `indices` exclusiones. Por ejemplo: `"indices": "-.kibana*,-.opendistro*,-.opensearch-observability*,-.plugins-ml-config*"`.

Puede restaurar una instantánea sin eliminar sus datos mediante los parámetros `rename_pattern` y `rename_replacement`. Para obtener más información sobre estos parámetros, consulte los campos de solicitud de la API Restore Snapshot y la solicitud de ejemplo en la OpenSearch documentación.

> **Note**
>
> Si no están disponibles todas las particiones principales para los índices implicados, es posible que una instantánea tenga `state` establecido como `PARTIAL`. Este valor indica que los datos de al menos una partición no se han almacenado correctamente. Puede continuar con la restauración a partir de una instantánea parcial, pero es posible que tenga que utilizar instantáneas anteriores para restaurar los índices que falten.

Eliminar instantáneas manuales

Ejecute el siguiente comando para tomar una instantánea manual:

```
DELETE _snapshot/repository-name/snapshot-name
```

Automatizar instantáneas con la administración de instantáneas

Puede configurar una política de administración de instantáneas (SM) en los OpenSearch paneles para automatizar la creación y eliminación periódicas de instantáneas. SM puede tomar una instantánea de un grupo de índices, mientras que Administración de estados de índice solo puede tomar una instantánea por índice. Para utilizar SM in OpenSearch Service, debe registrar su propio

repositorio de Amazon S3. Para ver instrucciones sobre cómo registrar su repositorio, consulte Registrar un repositorio de instantáneas manuales.

> **Note**
>
> A partir de la OpenSearch versión 2.5, se admite la administración de instantáneas. Los dominios que ejecutan versiones anteriores no admiten esta función.

Antes de SM, OpenSearch Service ofrecía una función de instantáneas automática y gratuita que todavía está activada de forma predeterminada. Esta característica envía las instantáneas al repositorio mantenido por el servicio`cs-*`. Para desactivar la característica, póngase en contacto con Soporte.

Para obtener más información sobre la función SM, consulte la administración de instantáneas en la OpenSearch documentación.

Actualmente, SM no admite la creación de instantáneas en varios tipos de índices. Por ejemplo, si intenta crear una instantánea en varios índices con `*` y algunos índices están en el nivel semiactivo, la creación de la instantánea producirá un error. Si necesita que la instantánea contenga varios tipos de índices, utilice la acción de instantánea de ISM hasta que SM admita esta opción.

Configuración de permisos

Si vas a actualizar a la versión 2.5 desde una versión anterior del dominio de OpenSearch servicio, es posible que los permisos de seguridad de administración de instantáneas no estén definidos en el dominio. Los usuarios que no sean administradores deben estar asignados a este rol para usar la administración de instantáneas en los dominios mediante un control de acceso detallado. Para crear el rol de administración de instantáneas de forma manual, siga estos pasos:

1. En los OpenSearch paneles de control, vaya a Seguridad y elija Permisos.
2. Seleccione Crear grupo de acciones y configure los siguientes grupos:

Nombre del grupo	Permisos
`snapshot_managemen`	• `cluster:admin/opensearch/snapshot_management/*` • `cluster:admin/opensearch/notifications/feature/publish`

Nombre del grupo	Permisos
`t_full_ac cess`	• `cluster:admin/repository/*` • `cluster:admin/snapshot/*`
`snapshot_ managemen t_read_ac cess`	• `cluster:admin/opensearch/snapshot_management/ policy/get` • `cluster:admin/opensearch/snapshot_management/ policy/search` • `cluster:admin/opensearch/snapshot_management/ policy/explain` • `cluster:admin/repository/get` • `cluster:admin/snapshot/get`

3. Seleccione Roles y, a continuación, Crear rol.
4. Asigne el nombre snapshot_management_role al rol.
5. Para Permisos de clúster, seleccione `snapshot_management_full_access` o `snapshot_management_read_access`.
6. Seleccione Crear.
7. Después de crear el rol, debe mapearlo a cualquier rol de usuario o backend que administre instantáneas.

Consideraciones

Tenga en cuenta lo siguiente al configurar la administración de instantáneas:

- Se permite una política por repositorio.
- Se permiten hasta 400 instantáneas para una política.
- Esta característica no se ejecutará si su dominio tiene un estado rojo, está sometido a una presión elevada por la JVM (un 85 % o más) o tiene una característica de captura bloqueada. Si el rendimiento general de indexación y búsqueda de su clúster se ve afectado, es posible que SM también se vea afectado.
- Una operación de instantánea solo comienza una vez finalizada la operación anterior, por lo que una política no activa ninguna operación de instantánea simultánea.

- Varias políticas con la misma programación pueden provocar un pico de recursos. Si los índices instantáneos de las políticas se superponen, las operaciones de instantáneas a nivel de partición solo se pueden ejecutar de forma secuencial, lo que puede provocar un problema de rendimiento en cascada. Si las políticas comparten un repositorio, habrá un pico de operaciones de escritura en ese repositorio.
- Le recomendamos que programe la automatización de sus operaciones instantáneas para que no se realice más de una vez por hora, a menos que se trate de un caso de uso especial.

Automatizar instantáneas con la administración de estado de índice

Puede utilizar la operación de administración de estado de índice (ISM) para activar automáticamente instantáneas de índices en función de los cambios en su antigüedad, tamaño o número de documentos. ISM es la mejor opción cuando se necesita una instantánea por índice. Si necesita hacer una instantánea de un grupo de índices, consulte Automatizar instantáneas con la administración de instantáneas.

Para utilizar SM in OpenSearch Service, debe registrar su propio repositorio de Amazon S3. Para obtener un ejemplo de política ISM con la operación `snapshot`, consulte Políticas de muestra.

Utilizar Curator para instantáneas

Si ISM no funciona para la administración de índices e instantáneas, puede utilizar Curator en su lugar. Ofrece una funcionalidad avanzada de filtrado que ayuda a simplificar tareas de administración en clústeres complejos. Utilice pip para instalar Curator:

```
pip install elasticsearch-curator
```

Puede utilizar Curator como una interfaz de línea de comandos (CLI) o una API de Python. Si utiliza la API de Python, debe utilizar la versión 7.13.4 o anterior del cliente elasticsearch-py heredado. No admite el cliente opensearch-py.

Si utiliza la CLI, exporte sus credenciales en la línea de comandos y configure `curator.yml` como se indica a continuación:

```
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
```

Actualización de los dominios OpenSearch de Amazon Service

> **Note**
>
> OpenSearch y las actualizaciones de la versión de Elasticsearch son diferentes de las actualizaciones del software del servicio. Para obtener información sobre cómo actualizar el software de servicio para su dominio OpenSearch de servicio, consulte. the section called "Actualizaciones del software del servicio"

Amazon OpenSearch Service ofrece actualizaciones locales para dominios que ejecutan la OpenSearch versión 1.0 o una versión posterior, o Elasticsearch 5.1 o una versión posterior. Si utilizas servicios como Amazon Data Firehose o Amazon CloudWatch Logs para transmitir datos a OpenSearch Service, comprueba que estos servicios sean compatibles con la versión más reciente de OpenSearch antes de realizar la migración.

Rutas de actualización admitidas

Actualmente, el OpenSearch servicio admite las siguientes rutas de actualización:

Desde la versión	Hasta la versión
OpenSearch 1.3, 2. x o 3. x	OpenSearch 3. x Si está ejecutando OpenSearch 1.3 o 2.x, primero debe actualizar a la OpenSearch 2.19 antes de actualizar a la 3.x. OpenSearch Al actualizar de la versión OpenSearch 2.19 a la OpenSearch 3.x, es posible que tengas que corregir las siguientes configuraciones de índice incompatibles.

Desde la versión	Hasta la versión
	Quedaron en desuso a principios de la versión 2.x y, si no se eliminan, provocará n un error en la validación de la comprobación de la actualización. • index.knn.algo_param.ef_construction • index.knn.algo_param.m • index.knn.space_type • index.store.hybrid.mmap.extensions Para obtener más información sobre la configuración de índice obsoleta y los pasos para corregirla, consulte Configuración de índices. Para ver una lista completa de los cambios más importantes de la versión OpenSearch 3.x, consulte la sección 3.0.0. Si tu dominio contiene índices creados en OpenSearch 1.3, Elasticsearch 7.10 o versiones anteriores, debes volver a indexarlos antes de actualizar a la versión 3.x. OpenSearch OpenSearch La versión 3.x no admite estos índices antiguos, independientemente de si están almacenados en caliente o en frío. UltraWarm Para volver a indexar índices incompatibles UltraWarm o inactivos: 1. Migre los índices a un almacenamiento activo. 2. Vuelva a indexar los datos. 3. Vuelva a migrar los índices a un almacenamiento en caliente o en frío. Como alternativa, elimine los índices si ya no los necesita.

Desde la versión	Hasta la versión
OpenSearch 1.3 o 2. x	OpenSearch 2. x OpenSearch 2.17 habilitará la búsqueda simultánea de segmentos de forma predeterminada con el modo automático si el dominio cumple las siguientes condiciones: • No se ha establecido explícitamente ninguna configuración de búsqueda simultánea previa. • Todas las instancias de datos (calientes y cálidos) son del tipo 2.xl o superior. • El uso medio de la CPU p90 en las instancias de datos (calientes y cálidos) durante más de una semana es inferior al 45 %. Para obtener más información sobre la configuración de búsqueda de segmentos simultáneos aquí, consulte Búsqueda de segmentos simultáneos. El control de versiones 2.3 incluye los siguientes cambios importantes: • El `type` parámetro se eliminó de todos los puntos finales de la OpenSearch API en la versión 2.0. Para más información, consulte Cambios bruscos. • Si tu dominio contiene índices (activos o fríos) que se crearon originalmente en Elasticsearch 6.8, esos índices no son compatibles con la 2.3. UltraWarm OpenSearch Antes de actualizar a la versión 2.3, debe volver a indexar los índices incompatibles. En el caso de índices incompatibles UltraWarm o en frío, migre los datos a un almacenamiento activo, vuelva a indexar los datos y, a continuación, vuelva a migrarlos a un almacenamiento caliente o frío. También puede quitar índices si ya no los necesita. Si actualiza accidentalmente su dominio al control de versiones 2.3 sin realizar primero estos pasos, no podrá migrar los índices incompatibles fuera de su nivel de almacenamiento actual. La única opción es eliminarlos.

Desde la versión	Hasta la versión
OpenSearch 1. x	OpenSearch 1. x
Elasticsearch 7.x	Elasticsearch 7. x o 1. OpenSearch x **Important** OpenSearch 1. x introduce numerosos cambios importantes. Para obtener más información, consulte Cambio de nombre OpenSearch de Amazon Service.
Elasticsearch 6.8	Elasticsearch 7. x o 1. OpenSearch x **Important** Elasticsearch 7.0 y OpenSearch 1.0 incluyen numerosos cambios importantes. Antes de iniciar una actualización local, recomendamos realizar una instantánea manual de la versión 6. x dominio, restaurándolo en una prueba 7. x o OpenSearch 1. dominio x y usar ese dominio de prueba para identificar posibles problemas de actualización. Para ver los cambios más importantes en la OpenSearch versión 1.0, consulteCambio de nombre OpenSearch de Amazon Service. Al igual que Elasticsearch 6.x, los índices solo pueden contener un tipo de mapeo, pero ese tipo debe denominarse `_doc`. Como resultado , algunos ya APIs no requieren un tipo de mapeo en el cuerpo de la solicitud (como la `_bulk` API). Para los índices nuevos, Elasticsearch 7 se hospedó automáticamente. x y 1. OpenSearch x tienen un recuento de fragmentos predeterminado de uno. OpenSearch Dominios de servicio en Elasticsearch 7. x y las versiones posteriores conservan el valor predeterminado anterior de cinco.

Desde la versión	Hasta la versión
Elasticsearch 6.x	Elasticsearch 6.x
Elasticsearch 5.6	Elasticsearch 6.x **Important** Los índices creados en la versión 6.x ya no admiten varios tipos de mapeo. Los índices creados en la versión 5.x siguen admitiendo varios tipos de mapeo cuando se restauran en un clúster 6.x. Compruebe que el código de cliente solo cree un único tipo de mapeo por índice. Para minimizar el tiempo de inactividad durante la actualización de Elasticsearch 5.6 a 6. x, OpenSearch Service vuelve a indexar el `.kibana` índice, lo elimina `.kibana-6 .kibana`, crea un nombre `.kibana` de alias y asigna el nuevo índice al nuevo alias.
Elasticsearch 5.x	Elasticsearch 5.x

El proceso de actualización consta de tres pasos:

1. Comprobaciones previas a la actualización: el OpenSearch servicio comprueba si hay problemas que puedan bloquear una actualización y no pasa al siguiente paso a menos que estas comprobaciones se realicen correctamente.
2. Instantánea: el OpenSearch servicio toma una instantánea del clúster OpenSearch o del clúster de Elasticsearch y no pasa al siguiente paso a menos que la instantánea se realice correctamente. Si se produce un error en la actualización, OpenSearch Service usa esta instantánea para restaurar el clúster a su estado original. Para obtener más información, consulte the section called “No se puede cambiar a una versión anterior después de una actualización”.
3. Actualización: el OpenSearch servicio inicia la actualización, que puede tardar entre 15 minutos y varias horas en completarse. OpenSearch Es posible que los paneles no estén disponibles durante parte de la actualización o durante toda la actualización.

Actualización de un dominio (consola)

El proceso de actualización es irreversible y no se puede pausar ni cancelar. Durante una actualización no puede realizar cambios de configuración en el dominio. Antes de comenzar una actualización, compruebe que desea continuar. Puede utilizar estos mismos pasos para realizar la verificación de actualización sin iniciar realmente una actualización.

Si el clúster tiene nodos maestros dedicados, OpenSearch las actualizaciones se completan sin tiempo de inactividad. De lo contrario, el clúster podría no responder durante varios segundos después de la actualización mientras elige un nodo maestro.

Para actualizar un dominio a una versión posterior de OpenSearch Elasticsearch

1. Tome una instantánea manual de su dominio. Esta instantánea sirve como copia de seguridad que puede restaurar en un dominio nuevo si desea volver a usar la versión anterior OpenSearch .
2. Ve a Iniciar sesión en la consola https://aws.amazon.comy selecciona Iniciar sesión.
3. En Analytics, selecciona Amazon OpenSearch Service.
4. En el panel de navegación, en Dominios, seleccione el dominio que desea actualizar.
5. Seleccione Acciones y Actualizar.
6. Seleccione la version a la que actualizar. Si vas a actualizar a una OpenSearch versión, aparecerá la opción Habilitar el modo de compatibilidad. Si habilitas esta configuración, indicará que OpenSearch su versión es 7.10 para permitir que los clientes y complementos de Elasticsearch OSS, como Logstash, sigan funcionando con Amazon Service. OpenSearch Podrá desactivar esta configuración más tarde
7. Seleccione Actualizar.
8. Verifique el Estado en el panel del dominio para monitorear el estado de la actualización.

Actualización de un dominio (CLI)

Puedes usar las siguientes operaciones para identificar la versión correcta de OpenSearch Elasticsearch para tu dominio, iniciar una actualización inmediata, realizar la comprobación previa a la actualización y ver el progreso:

- `get-compatible-versions` (`GetCompatibleVersions`)
- `upgrade-domain` (`UpgradeDomain`)

- `get-upgrade-status` (`GetUpgradeStatus`)
- `get-upgrade-history` (`GetUpgradeHistory`)

Para obtener más información, consulta la referencia de comandos de AWS CLI y la referencia de la API de Amazon OpenSearch Service.

Actualización de un dominio (SDK)

En este ejemplo, se utiliza el cliente Python de OpenSearchServicebajo nivel del AWS SDK para Python (Boto) para comprobar si un dominio es apto para la actualización a una versión específica, lo actualiza y comprueba continuamente el estado de la actualización.

```
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()
```

Solución de errores de validación

Cuando inicias una actualización de una versión OpenSearch o de Elasticsearch, OpenSearch Service primero realiza una serie de comprobaciones de validación para garantizar que tu dominio

sea apto para una actualización. Si se produce un error en alguna de estas comprobaciones, recibirá una notificación con los problemas específicos que debe corregir antes de actualizar su dominio. Para obtener una lista de posibles problemas y los pasos para resolverlos, consulte the section called "Solución de errores de validación".

Solución de problemas de una actualización

Las actualizaciones locales requieren dominios en buen estado. Es posible que su dominio no pueda optar por una actualización o que no se actualice por diversas razones. La siguiente tabla muestra los problemas más comunes.

Problema	Description (Descripción)
No se admite el complemento opcional	Cuando actualizas un dominio con complementos opcionales, OpenSearch Service también actualiza automáticamente los complemen tos. Por lo tanto, la versión de destino de su dominio también debe admitir estos complementos opcionales. Si el dominio tiene instalado un complemento opcional que no está disponible para la versión de destino, se produce un error en la solicitud de actualización.
Demasiadas particion es por nodo	OpenSearch, así como 7. x versiones de Elasticsearch, tienen una configuración predeterminada de no más de 1000 particiones por nodo. Si un nodo de tu clúster actual supera esta configuración, el OpenSearc h servicio no te permite realizar la actualización. Consulte the section called "Límite máximo de fragmentos superado" para obtener opciones de solución de problemas.
Dominio en procesamiento	El dominio está en medio de un cambio de configuración. Compruebe si es posible optar a la actualización una vez que se complete la operación.
Estado rojo del clúster	Uno o varios índices en el clúster aparecen en rojo. Para ver los pasos de solución de problemas, consulte the section called "Estado rojo del clúster".
Tasa de errores alta	El clúster devuelve un gran número de errores 5xx al intentar procesar solicitudes. Este problema suele ser el resultado de demasiadas solicitud es de lectura o escritura simultáneas. Considere la posibilidad de reducir el tráfico hacia el clúster o el escalado del dominio.

Problema	Description (Descripción)
Cerebro dividido	Cerebro dividido significa que el clúster tiene más de un nodo maestro y se ha dividido en dos clústeres que nunca se volverán a unir por sí mismos. Puede evitar un cerebro dividido con el número recomendado de nodos maestros dedicados. Para obtener ayuda a fin de recuperar un cerebro dividido, póngase en contacto con Soporte.
Nodo maestro no encontrado	OpenSearch El servicio no encuentra el nodo principal del clúster. Si el dominio utiliza Multi-AZ, es posible que un error de una zona de disponibilidad haya hecho que el clúster pierda el quorum y no pueda elegir un nuevo nodo maestro. Si el problema no se resuelve por sí solo, póngase en contacto con Soporte.
Demasiadas tareas pendientes	El nodo maestro está bajo una gran carga y tiene muchas tareas pendientes. Considere la posibilidad de reducir el tráfico hacia el clúster o el escalado del dominio.
Volumen de almacenamiento dañado	El volumen de disco de uno o más nodos no funciona correctamente. Este problema a menudo ocurre junto con otros problemas, como una alta tasa de error o demasiadas tareas pendientes. Si se produce aisladamente y no se resuelve por sí solo, póngase en contacto con Soporte.
Problema de clave de KMS	La clave de KMS que se utiliza para cifrar el dominio está inaccesible o no está. Para más información, consulte the section called “Monitorear dominios que cifran los datos en reposo”.
Snapshot en proceso de creación	El dominio está tomando una instantánea en este momento. Compruebe si es posible optar a la actualización una vez que se complete la instantánea. Compruebe también si puede hacer una lista de los repositorios de instantáneas manuales, una lista de instantáneas dentro de dichos repositorios y tomar instantáneas manuales. Si el OpenSearch servicio no puede comprobar si una instantánea está en curso, las actualizaciones pueden fallar.

Problema	Description (Descripción)
Error en instantánea o tiempo de espera agotado	La instantánea previa a la actualización ha tardado demasiado tiempo en completarse o ha generado un error. Compruebe el estado del clúster y vuelva a intentarlo. Si el problema sigue sin resolverse, póngase en contacto con Soporte.
Índices incompatibles	Uno o varios índices son incompatibles con la versión de destino. Este problema puede producirse si migraste los índices desde una versión anterior de OpenSearch Elasticsearch. Reindexe los índices e inténtelo de nuevo.
Alto uso del disco	El uso del disco para el clúster es superior al 90 %. Elimine los datos o escale el dominio, e inténtelo de nuevo.
Alto uso de JVM	La presión de memoria de JVM es superior al 75 %. Reduzca el tráfico hacia el clúster o amplíe el dominio, e inténtelo de nuevo.
OpenSearch Problema con el alias de los paneles	`.dashboards` ya está configurado como un alias y se asigna a un índice incompatible, probablemente uno de una versión anterior de OpenSearch Dashboards. Reindexe e inténtelo de nuevo.
Estado rojo de Dashboards	OpenSearch El estado del panel es rojo. Intente utilizar Dashboards cuando se complete la actualización. Si el estado rojo persiste, resuélval o manualmente, e inténtelo de nuevo.
Compatibilidad entre clústeres	Solo puede actualizar si se mantiene la compatibilidad entre clústeres entre los dominios de origen y destino después de la actualización. Durante el proceso de actualización, se identifican las conexiones incompatibles. Para continuar, actualice el dominio remoto o elimine las conexiones incompatibles. Tenga en cuenta que si la replicación está activa en el dominio, no podrá reanudarla una vez que elimine la conexión.
Otro problema con OpenSearch el servicio	Los problemas con el propio OpenSearch Servicio pueden hacer que tu dominio aparezca como no apto para una actualización. Si a su dominio no se aplica ninguna de las condiciones anteriores y el problema persiste durante más de un día, póngase en contacto con Soporte.

Uso de una instantánea para migrar datos

Las actualizaciones locales son la forma más fácil, rápida y confiable de actualizar un dominio a una versión posterior OpenSearch o a una versión de Elasticsearch. Las instantáneas son una buena opción si tiene que migrar de una versión previa a la 5.1 de Elasticsearch o si desea migrar a un clúster totalmente nuevo.

La siguiente tabla muestra cómo usar las instantáneas para migrar datos a un dominio que usa una versión diferente OpenSearch o de Elasticsearch. Para obtener información sobre cómo tomar instantáneas y restaurarlas, consulte the section called “Crear instantáneas de índice”.

Desde la versión	Hasta la versión	Proceso de migración

OpenSearch 1.3 o 2. x	OpenSearch 2. x	1. Revise los cambios más recientes de la versión OpenSearch 2.3 para ver si necesita realizar ajustes en sus índices o aplicaciones.
		2. Cree una instantánea manualmente del dominio 1.3 o 2.x.
		3. Cree un dominio 2.x que sea una versión superior al dominio 1.3 o 2.x original.
		4. Restaure la instantánea desde el dominio original al dominio 2.x. Durante la operación, es probable que tenga que restaurar el índice `.opensearch` con un nuevo nombre:
		`POST _snapshot/ <repository-name> /<snapshot-name>/_restore`
		`{`
		`  "indices": "*",`
		`  "ignore_unavailable": true,`
		`  "rename_pattern": ".opensearch",`
		`  "rename_replacement": ".backup-opensearch"`
		`}`
		A continuación, puede reindexar `.backup-opensearch` en el nuevo dominio y asignarlo a `.opensearch`. Tenga en cuenta que la llamada

Desde la versión	Hasta la versión	Proceso de migración
		REST `_restore` no incluye `include_global_state` porque la entrada predeterminada en `_restore` es falsa. Como resultado, el dominio de prueba no incluirá ninguna plantilla de índice y no tendrá el estado completo de la copia de seguridad. 5. Si ya no necesita el dominio original, elimínelo. De lo contrario, el dominio seguirá generando costos.

Desde la versión	Hasta la versión	Proceso de migración

OpenSearch 1. x	OpenSearch 1. x	1. Cree una instantánea manual del dominio 1.x.
		2. Cree un dominio 1.x que sea una versión superior al dominio 1.x original.
		3. Restaure la instantánea desde el dominio original al nuevo dominio 1.x. Durante la operación, es probable que tenga que restaurar el índice `.opensearch` con un nuevo nombre:
		`POST _snapshot/ <repository-name> /<snapshot-name>/_restore`
		`{`
		`  "indices": "*",`
		`  "ignore_unavailable": true,`
		`  "rename_pattern": ".opensearch",`
		`  "rename_replacement": ".backup-opensearch"`
		`}`
		A continuación, puede reindexar `.backup-opensearch` en el nuevo dominio y asignarlo a `.opensearch`. Tenga en cuenta que la llamada REST `_restore` no incluye `include_global_state` porque la entrada predeterminada en `_restore` es falsa. Como resultado, el dominio de prueba no incluirá ninguna plantilla de índice y no tendrá el estado completo de la copia de seguridad.
		4. Si ya no necesita el dominio original, elimínelo. De lo contrario, el dominio seguirá generando costos.

Desde la versión	Hasta la versión	Proceso de migración

Elasticsearch 6.x o 7.x	OpenSearch 1. x	1. Revise los cambios más recientes de la OpenSearch versión 1.0 para ver si necesita realizar ajustes en sus índices o aplicaciones.
		2. Cree una instantánea manualmente del dominio de Elasticsearch 7.x o 6.x.
		3. Cree un OpenSearch 1. dominio x.
		4. Restaure la instantánea del dominio de Elasticsearch al OpenSearch dominio. Durante la operación, es probable que tenga que restaurar el índice `.elasticsearch` con un nuevo nombre:
		`POST _snapshot/ <repository-name> /<snapshot-name>/_restore`
		`{`
		`  "indices": "*",`
		`  "ignore_unavailable": true,`
		`  "rename_pattern": ".elasticsearch",`
		`  "rename_replacement": ".backup-opensearch"`
		`}`
		A continuación, puede reindexar `.backup-opensearch` en el nuevo dominio y asignarlo a `.elasticsearch`. Tenga en cuenta que la llamada REST `_restore` no incluye `include_global_state` porque la entrada predeterminada en `_restore` es falsa. Como resultado, el dominio de prueba no incluirá ninguna plantilla de índice y no tendrá el estado completo de la copia de seguridad.
		5. Si ya no necesita el dominio original, elimínelo. De lo contrario, el dominio seguirá generando costos.

Desde la versión	Hasta la versión	Proceso de migración

Elasticsearch 6.x	Elasticsearch 7.x	1. Consulte los cambios importantes en 7.0 para saber si necesita realizar ajustes en los índices o en la aplicación.
		2. Cree una instantánea manual del dominio 6.x.
		3. Cree un dominio de la versión 7x.
		4. Restaure la instantánea desde el dominio original al dominio 7x. Durante la operación, es probable que tenga que restaurar el índice `.opensearch` con un nuevo nombre:
		`POST _snapshot/ <repository-name> /<snapshot-name>/_restore`
		`{`
		`  "indices": "*",`
		`  "ignore_unavailable": true,`
		`  "rename_pattern": ".elasticsearch",`
		`  "rename_replacement": ".backup-elasticsearch"`
		`}`
		A continuación, puede reindexar `.backup-elasticsearch` en el nuevo dominio y asignarlo a `.elasticsearch`. Tenga en cuenta que la llamada REST `_restore` no incluye `include_global_state` porque la entrada predeterminada en `_restore` es falsa. Como resultado, el dominio de prueba no incluirá ninguna plantilla de índice y no tendrá el estado completo de la copia de seguridad.
		5. Si ya no necesita el dominio original, elimínelo. De lo contrario, el dominio seguirá generando costos.

Desde la versión	Hasta la versión	Proceso de migración
Elasticsearch 6.x	Elasticsearch 6.8	1. Cree una instantánea manual del dominio 6.x. 2. Cree un dominio de la versión 6.8. 3. Restaure la instantánea desde el dominio original al dominio 6.8. 4. Si ya no necesita el dominio original, elimínelo. De lo contrario, el dominio seguirá generando costos.
Elasticsearch 5.x	Elasticsearch 6.x	1. Consulte los cambios importantes en 6.0 para saber si necesita realizar ajustes en los índices o en la aplicación. 2. Cree una instantánea manual del dominio 5.x. 3. Cree un dominio de la versión 6.x. 4. Restaure la instantánea desde el dominio original al dominio 6.x. 5. Si ya no necesita el dominio 5.x, elimínelo. De lo contrario, el dominio seguirá generando costos.
Elasticsearch 5.x	Elasticsearch 5.6	1. Cree una instantánea manual del dominio 5.x. 2. Cree un dominio de la versión 5.6. 3. Restaure la instantánea desde el dominio original al dominio 5.6. 4. Si ya no necesita el dominio original, elimínelo. De lo contrario, el dominio seguirá generando costos.

Desde la versión	Hasta la versión	Proceso de migración
Elasticsearch 2.3	Elasticsearch 6.x	Las instantáneas de Elasticsearch 2.3 no son compatibl es con la versión 6.x. Para migrar sus datos directamente de la versión 2.3 a 6.x, debe volver a crear manualmente sus índices en el nuevo dominio. Como alternativa, puede seguir los pasos del cambio de la versión 2.3 a 5.x de esta tabla, realizar las operaciones de `_reindex` en el nuevo dominio 5.x para convertir sus índices 2.3 en índices 5.x y después seguir los pasos del cambio de la versión 5.x a 6.x.
Elasticsearch 2.3	Elasticsearch 5.x	1. Consulte los cambios importantes en 5.0 para saber si necesita realizar ajustes en los índices o en la aplicación. 2. Cree una instantánea manualmente del dominio 2.3. 3. Cree un dominio de la versión 5.x. 4. Restaure la instantánea desde el dominio 2.3 al dominio 5.x. 5. Si ya no necesita el dominio 2.3, elimínelo. De lo contrario, el dominio seguirá generando costos.

Desde la versión	Hasta la versión	Proceso de migración
Elasticsearch 1.5	Elasticsearch 5.x	Las instantáneas de Elasticsearch 1.5 no son compatibles con la versión 5.x. Para migrar sus datos de la versión 1.5 a 5.x, debe volver a crear manualmente sus índices en el nuevo dominio. **Important** Las instantáneas de la versión 1.5 son compatibles con la versión 2.3, pero los dominios del OpenSearch Servicio 2.3 no admiten esta operación. `_reindex` Dado que no puede reindexarlos, los índices que se originaron en un dominio 1.5 siguen sin restaurar instantáneas de 2.3 a dominios 5x.
Elasticsearch 1.5	Elasticsearch 2.3	1. Utilice el complemento de migración para saber si puede hacer directamente la actualización a la versión 2.3. Es posible que tenga que realizar cambios en los datos antes de la migración. a. En un navegador web, abra `http://`*`domain-endpoint`*` /_plugin/migration/` . b. Seleccione Realizar comprobaciones ahora. c. Revise los resultados y, si es necesario, siga las instrucciones para realizar cambios en los datos. 2. Cree una instantánea manualmente del dominio 1.5. 3. Cree un dominio de la versión 2.3. 4. Restaure la instantánea desde el dominio 1.5 al dominio 2.3. 5. Si ya no necesita el dominio 1.5, elimínelo. De lo contrario, el dominio seguirá generando costos.

Crear un punto de conexión personalizado para Amazon OpenSearch Service

Al crear un punto de conexión personalizado para tu dominio de Amazon OpenSearch Service, te resultará más fácil consultar tus OpenSearch paneles OpenSearch URLs y los de control. Puedes incluir la marca de tu empresa o simplemente usar un easier-to-remember punto final más corto que el estándar.

Si alguna vez necesita cambiar a un nuevo dominio, simplemente actualice el DNS para que apunte a la nueva URL y continúe utilizando el mismo punto de conexión anterior.

Para proteger los puntos finales personalizados, puede generar un certificado en AWS Certificate Manager (ACM) o importar uno propio.

Puntos de conexión personalizados para nuevos dominios

Puede habilitar un punto final personalizado para un nuevo dominio de OpenSearch servicio mediante la consola de OpenSearch servicio o la AWS CLI API de configuración.

Cómo personalizar el punto de conexión (consola)

1. En la consola de OpenSearch servicio, selecciona Crear dominio y proporciona un nombre para el dominio.
2. En Punto de conexión personalizado, seleccione Habilitar punto de conexión personalizado.
3. En Nombre de anfitrión personalizado, ingrese el nombre de anfitrión del punto de conexión personalizado preferido. El nombre de anfitrión debe ser un nombre de dominio completo (FQDN), como www.yourdomain.com o example.yourdomain.com.

   > **Note**
   >
   > Si no dispone de un certificado comodín, es posible que necesite obtener un nuevo certificado para los subdominios del punto de conexión personalizado.

4. En Certificado de AWS , seleccione el certificado SSL que desee utilizar con el dominio. Si no hay certificados disponibles, puede importar uno a ACM o utilizar ACM para aprovisionar uno. Para más información, consulte Emisión y administración de certificados en la Guía del usuario de Certificate Manager de AWS .

> **Note**
>
> El certificado debe tener el nombre de punto final personalizado y estar en la misma cuenta que su dominio OpenSearch de servicio. El estado del certificado debe ser EMITIDO.

- Siga el resto de pasos para crear el dominio y elija Crear.
- Seleccione el dominio, cuando se haya terminado de procesar, para ver el punto de conexión personalizado.

Para utilizar la CLI o la API de configuración, utilice las operaciones `CreateDomain` y `UpdateDomainConfig`. Para obtener más información, consulta la Referencia de AWS CLI comandos y la Referencia de la API de Amazon OpenSearch Service.

Puntos de conexión personalizados para dominios existentes

Para añadir un punto de conexión personalizado a un dominio de OpenSearch servicio existente, selecciona Editar y realiza los pasos 2 y 4 anteriores.

Asignación CNAME

Después de habilitar un punto de enlace personalizado para su dominio de OpenSearch servicio, puede crear un mapeo de CNAME en Amazon Route 53 (o en su proveedor de servicios de DNS preferido). La creación de una asignación CNAME le permitirá dirigir el tráfico a su punto de conexión personalizado y sus subdominios. Sin esta asignación, no podrá dirigir el tráfico a su punto de conexión personalizado. A fin de conocer los pasos para crear esta asignación en Route 53, consulte Configuring DNS routing for a new domain y Creating a new hosted zone for a subdomain. Para otros proveedores, consulte su documentación.

Cree un registro CNAME en el que el punto de conexión personalizado apunte al punto de conexión del dominio generado automáticamente. Si el dominio es de doble pila, puede apuntar su registro CNAME a cualquiera de los dos puntos de conexión generados por el servicio. La capacidad de doble pila del punto de conexión personalizado depende del punto de conexión generado por el servicio al que apunte el registro CNAME. El nombre de host del punto de conexión personalizado es el nombre del registro CNAME y el nombre de host del punto de conexión del dominio es el valor del registro CNAME.

Si utilizas la autenticación SAML para los OpenSearch paneles, debes actualizar tu IDP con la nueva URL del SSO.

Puede utilizar Amazon Route 53 a fin de crear un tipo de registro de alias a fin de dirigir el punto de conexión personalizado de su dominio a un punto de conexión de búsqueda de doble pila. Para crear un tipo de registro de alias, debe configurar su dominio para que utilice el tipo de dirección IP de doble pila. Para ello, puede utilizar la API de Route 53.

Para crear un tipo de registro de alias mediante la API de Route 53, especifique el alias de destino de su dominio. Puedes encontrar el alias de destino de tu dominio en el campo Zona alojada (pila doble) de la sección de terminales personalizados de la consola de OpenSearch servicio o bien, mediante la `DescribeDomain` API y copiando el valor del. `DomainEndpointV2HostedZoneId`

Auto-Tune para Amazon Service OpenSearch

Auto-Tune in Amazon OpenSearch Service utiliza las métricas de rendimiento y uso del OpenSearch clúster para sugerir cambios de configuración relacionados con la memoria, incluidos los tamaños de las colas y la caché y la configuración de la máquina virtual Java (JVM) en los nodos. Estos cambios opcionales mejoran la velocidad y la estabilidad del clúster.

Algunos cambios se implementan inmediatamente, mientras que otros se programan durante la ventana de menor actividad de su dominio. Puedes volver a la configuración predeterminada OpenSearch del servicio en cualquier momento. A medida que Auto-Tune recopila y analiza las métricas de rendimiento de tu dominio, puedes ver sus recomendaciones en la consola de OpenSearch servicio, en la página de notificaciones.

Auto-Tune está disponible en formato comercial Regiones de AWS en dominios que ejecutan cualquier OpenSearch versión, o Elasticsearch 6.7 o posterior, con un tipo de instancia compatible.

Tipos de cambios

La función de ajuste automático cuenta con dos amplias categorías de cambios:

- Cambios no disruptivos que aplica a medida que se ejecuta el clúster.
- Cambios que requieren una implementación azul/verde, que se aplica durante el período de menor actividad del dominio.

Según las métricas de rendimiento del dominio, la función de ajuste automático puede sugerir realizar ajustes en las siguientes configuraciones:

Cambio de tipo	Categoría	Description (Descripción)
Tamaño de la pila de JVM	Azul/verde	De forma predeterminada, OpenSearch Service utiliza el 50% de la RAM de una instancia para el montón de JVM, hasta un tamaño de pila de 32 GiB. Si se aumenta este porcentaje, se obtiene OpenSearch más memoria, pero se deja menos para el sistema operativo y otros procesos. Los valores más altos pueden disminuir el número de pausas de recolección de basura, pero pueden aumentar la amplitud de esas pausas.
Configuración de JVM de nueva generación	Azul/verde	La configuración de JVM de “nueva generación” afecta la frecuencia de recolecciones de basura menores. Las recolecciones menores más frecuentes pueden disminuir el número de recolecciones principales y pausas.
Tamaño de la cola	No disruptivo	De forma predeterminada, el tamaño de la cola de búsqueda es `1000` y el tamaño de la cola de escritura es `10000`. La función de ajuste automático escala automáticamente las colas de búsqueda y escritura si hay una pila adicional disponible para encargarse de las solicitudes.
Tamaño de caché	No disruptivo	La caché de campo monitorea las estructuras de datos en la pila, por lo que es importante monitorear el uso de la caché. La función de ajuste automático escala el tamaño de la caché de datos de campo para evitar problemas de memoria y con los interruptores de circuito. La caché de solicitudes de partición se administra a nivel del nodo y tiene un tamaño máximo predeterminado del 1 % de la pila. La función de ajuste automático escala el tamaño de la caché de solicitudes de partición para aceptar más solicitudes de búsqueda e índice de las que puede manejar el clúster configurado.
Solicitar tamaño	No disruptivo	De forma predeterminada, cuando el tamaño total de las solicitudes en curso supera el 10% del total de la JVM (el 2% para los tipos

Cambio de tipo	Categoría	Description (Descripción)
		de `t2` instancia y el 1% para las solicitudes`t3.small`), limita todas las `_bulk` solicitudes nuevas `_search` y OpenSearch posteriores hasta que se completen las existentes. Auto-Tune ajusta automáticamente este umbral, normalmente entre el 5-15 %, basándose en la cantidad de JVM que está actualmente ocupada en el sistema. Por ejemplo, si la presión de la memoria de la JVM es alta, el ajuste automático podría reducir el umbral al 5 %, momento en el que podría ver más rechazos hasta que el clúster se estabilice y el umbral aumente.

Habilitar o deshabilitar la función de ajuste automático

OpenSearch El servicio habilita el ajuste automático de forma predeterminada en los dominios nuevos. Para habilitar o deshabilitar la función de ajuste automático en dominios existentes, recomendamos utilizar la consola, lo cual simplifica el proceso. La activación de Auto-Tune no provoca una blue/green implementación.

Actualmente no se puede habilitar o desactivar la función de ajuste automático mediante AWS CloudFormation.

Consola

Cómo habilitar el ajuste automático en un dominio existente

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación, en Dominios, seleccione el nombre del dominio para abrir la configuración del clúster.
3. Seleccione Activar si el ajuste automático aún no está activado.
4. Si lo desea, seleccione el período de menor actividad para programar las optimizaciones que requieran un blue/green despliegue durante el período de menor actividad configurado por el dominio. Para obtener más información, consulte the section called “Programación de mejoras de ajuste automático”.
5. Seleccione Guardar cambios.

CLI

Para habilitar el ajuste automático mediante el AWS CLI, envíe una solicitud: UpdateDomainConfig

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --auto-tune-options DesiredState=ENABLED
```

Programación de mejoras de ajuste automático

Antes del 16 de febrero de 2023, Auto-Tune utilizaba períodos de mantenimiento para programar los cambios que requerían una blue/green implementación. Las ventanas de mantenimiento ahora están en desuso en favor del período de menor actividad, que es un bloque de tiempo diario de 10 horas durante el cual su dominio suele experimentar poco tráfico. Puede modificar la hora de inicio predeterminada para la ventana de menor actividad, pero no puede modificar la duración.

Todos los dominios que tenían habilitados los períodos de mantenimiento de ajuste automático antes de la introducción de las ventanas de menor actividad el 16 de febrero de 2023 pueden seguir utilizando las ventanas de mantenimiento antiguas sin interrupciones. Sin embargo, le recomendamos que migre sus dominios actuales para utilizar en su lugar la ventana de menor actividad para el mantenimiento de los dominios. Para obtener instrucciones, consulte the section called “Migración desde los intervalos de mantenimiento de ajuste automático”.

Consola

Cómo programar las acciones de ajuste automático en las ventanas de menor actividad

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación, en Dominios, seleccione el nombre del dominio para abrir la configuración del clúster.
3. Vaya a la pestaña Ajuste automático y seleccione Editar.
4. Seleccione Activar si el ajuste automático aún no está activado.
5. En Programar optimizaciones durante la ventana de menor actividad, seleccione Ventana de menor actividad.
6. Seleccione Guardar cambios.

CLI

Para configurar tu dominio para programar las acciones de Auto-Tune durante el período de menor actividad configurado, incluye `UseOffPeakWindow` en la UpdateDomainConfigsolicitud:

```
aws opensearch update-domain-config \
  --domain-name my-domain \
  --auto-tune-options
 DesiredState=ENABLED,UseOffPeakWindow=true,MaintenanceSchedules=null
```

Supervisión de cambios de ajuste automático

Puede supervisar las estadísticas de ajuste automático en Amazon CloudWatch. Para obtener una lista completa de las métricas, consulte the section called “Métricas de ajuste automático”.

OpenSearch El servicio envía los eventos de Auto-Tune a Amazon EventBridge. Puede usarlo EventBridge para configurar reglas que envíen un correo electrónico o realicen una acción específica cuando se recibe un evento. Para ver el formato de cada evento de Auto-Tune enviado a EventBridge, consultethe section called “Eventos de ajuste automático”.

Etiquetado de dominios de Amazon OpenSearch Service

Las etiquetas te permiten asignar información arbitraria a un dominio de Amazon OpenSearch Service para que puedas categorizar y filtrar esa información. Una etiqueta es un par clave-valor que se define y se asocia a un OpenSearch dominio de servicio. Puede usar estas etiquetas para realizar un seguimiento de los costos agrupando los gastos de los recursos etiquetados de manera similar. AWS no aplica ningún significado semántico a tus etiquetas. Las etiquetas se interpretan estrictamente como cadenas de caracteres. Todas las etiquetas presentan los siguientes elementos:

Elemento de etiqueta	Description (Descripción)	Obligatorio
Clave de etiqueta	La clave de la etiqueta es el nombre de la etiqueta. La clave debe ser exclusiva del dominio del OpenSearch servicio al que está asociada. Para obtener una lista de restricciones básicas relativas a las claves y valores de las etiquetas, consulte Restricciones de las etiquetas definidas por el usuario.	Sí

Elemento de etiqueta	Description (Descripción)	Obligatorio
Valor de etiqueta	El valor de etiqueta es un valor de cadena en la etiqueta. Los valores de etiquetas pueden ser `null` y no tienen que ser únicos dentro de un conjunto de etiquetas. Por ejemplo, puede tener un par clave-valor en un conjunto de etiquetas de. project/Trinity and cost-center/Trinity Para obtener una lista de restricciones básicas relativas a las claves y valores de las etiquetas, consulte Restricciones de las etiquetas definidas por el usuario.	No

Cada dominio OpenSearch de servicio tiene un conjunto de etiquetas, que contiene todas las etiquetas asignadas a ese dominio de OpenSearch servicio. AWS no asigna automáticamente ninguna etiqueta a los dominios OpenSearch de servicio. Un conjunto de etiquetas puede contener entre 0 y 50 etiquetas. Si agrega una etiqueta a un dominio con la misma clave que una etiqueta existente en un recurso, el nuevo valor sobrescribirá el antiguo.

Ejemplos de etiquetas

Puede utilizar una clave para definir una categoría, y el valor podría ser un elemento dentro de esa categoría. Por ejemplo, puede definir una clave de etiqueta `project` y un valor de etiqueta de`Salix`, que indiquen que el dominio de OpenSearch servicio está asignado al proyecto de Salix. También puede usar etiquetas para designar los dominios de OpenSearch servicio que se utilizan para pruebas o producción mediante una clave como `environment=test` o`environment=production`. Intente utilizar un conjunto coherente de claves de etiquetas para facilitar el seguimiento de los metadatos asociados a los dominios OpenSearch de servicio.

También puedes usar etiquetas para organizar tu AWS factura y reflejar tu propia estructura de costos. Para ello, regístrate para recibir tu Cuenta de AWS factura con los valores clave de las etiquetas incluidos. A continuación, para ver los costos de los recursos combinados, organice la información de facturación de acuerdo con los recursos que tienen los mismos valores de clave de etiqueta. Por ejemplo, puedes etiquetar varios dominios de OpenSearch servicio con pares clave-valor y, a continuación, organizar la información de facturación para ver el coste total de cada dominio en varios servicios. Para obtener más información, consulte Uso de etiquetas de asignación de costos en la documentación de administración de costos y facturación de AWS .

> **Note**
>
> Las etiquetas se almacenan en caché con fines de autorización. Por este motivo, es posible que las adiciones y actualizaciones de las etiquetas OpenSearch de los dominios de servicio tarden varios minutos en estar disponibles.

Etiquetado de dominios (consola)

La consola es la forma más sencilla de etiquetas de un dominio.

Para crear una etiqueta (consola)

1. Vaya a https://aws.amazon.com, y seleccione Sign In to the Console (Inicie sesión en la consola).
2. En Analytics, selecciona Amazon OpenSearch Service.
3. Seleccione el dominio al que desee agregar etiquetas y vaya a la pestaña Tags (Etiquetas).
4. Seleccione Administrar y Add new tag (Agregar nueva etiqueta).
5. Introduzca una clave de etiqueta y un valor opcional.
6. Seleccione guardar.

Para eliminar una etiqueta, siga los mismos pasos y elija Quitar en la página Administrar etiquetas.

Para obtener más información sobre cómo utilizar la consola para trabajar con etiquetas, consulte Tag Editor en la Guía de introducción a la consola de administración de AWS .

Etiquetado de dominios (AWS CLI)

Puede crear etiquetas de recursos mediante el AWS CLI --add-tags comando with.

Sintaxis

```
add-tags --arn=<domain_arn> --tag-list Key=<key>,Value=<value>
```

Parámetro	Description (Descripción)
`--arn`	Nombre de recurso de Amazon para el dominio de OpenSearch servicio al que se adjunta la etiqueta.

Parámetro	Description (Descripción)
`--tag-list`	Conjunto de pares de clave-valor separados por espacios con el siguiente formato: `Key=<key>,Value=<value>`

Ejemplo

En el siguiente ejemplo se crean dos etiquetas para el dominio registros:

```
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
```

Puede eliminar etiquetas de un dominio OpenSearch de servicio mediante el --remove-tags comando.

Sintaxis

```
remove-tags --arn=<domain_arn> --tag-keys Key=<key>,Value=<value>
```

Parámetro	Description (Descripción)
`--arn`	Nombre de recurso de Amazon (ARN) para el dominio de OpenSearch servicio al que está asociada la etiqueta.
`--tag-keys`	Conjunto de pares clave-valor separados por espacios que desea eliminar del dominio del servicio. OpenSearch

Ejemplo

En el siguiente ejemplo se eliminan dos etiquetas del dominio logs que se crearon en el ejemplo anterior:

```
aws opensearch remove-tags --arn arn:aws:es:us-east-1:379931976431:domain/logs --tag-
keys service instances
```

Puede ver las etiquetas existentes de un dominio de OpenSearch servicio con el comando: --list-tags

Sintaxis

```
list-tags --arn=<domain_arn>
```

Parámetro	Description (Descripción)
`--arn`	Nombre de recurso de Amazon (ARN) para el dominio de OpenSearch servicio al que se adjuntan las etiquetas.

Ejemplo

En el siguiente ejemplo se muestran todas las etiquetas de recursos del dominio logs:

```
aws opensearch list-tags --arn arn:aws:es:us-east-1:379931976431:domain/logs
```

Etiquetado de dominios (AWS SDKs)

Las AWS SDKs (excepto Android e iOS SDKs) admiten todas las acciones definidas en la referencia de la API de Amazon OpenSearch Service, incluidas las `AddTags` `RemoveTags` operaciones`ListTags`, y. Para obtener más información sobre la instalación y el uso de AWS SDKs, consulte los kits de desarrollo de AWS software.

Python

En este ejemplo, se utiliza el cliente Python de OpenSearchServicebajo nivel del SDK de AWS para Python (Boto) para añadir una etiqueta a un dominio, enumerar la etiqueta adjunta al dominio y eliminar una etiqueta del dominio. Debe proporcionar valores para `DOMAIN_ARN`, `TAG_KEY` y `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
```

Realizar acciones administrativas en los dominios OpenSearch de Amazon Service

Amazon OpenSearch Service ofrece varias opciones administrativas que proporcionan un control detallado si necesitas solucionar problemas con tu dominio. Estas opciones incluyen la posibilidad de reiniciar el OpenSearch proceso en un nodo de datos y la posibilidad de reiniciar un nodo de datos.

OpenSearch El servicio monitorea los parámetros de estado de los nodos y, cuando hay anomalías, toma medidas correctivas para mantener la estabilidad de los dominios. Con las opciones administrativas para reiniciar el OpenSearch proceso en un nodo y reiniciar el propio nodo, usted tiene el control de algunas de estas acciones de mitigación.

Puede usar el Consola de administración de AWS AWS CLI, o el AWS SDK para realizar estas acciones. En las siguientes secciones, se explica cómo realizar estas acciones con la consola.

Reiniciar el OpenSearch proceso en un nodo de Amazon Service OpenSearch

Para reiniciar el OpenSearch proceso en un nodo

1. Diríjase a la consola de OpenSearch servicio enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Dominios. Seleccione el nombre del dominio con el que desea trabajar.
3. Cuando se abra la página de detalles del dominio, vaya a la pestaña Estado de la instancia.
4. En Nodos de datos, seleccione el botón situado junto al nodo en el que desea reiniciar el proceso.
5. Seleccione el menú desplegable Acciones y elija Reiniciar el OpenSearch/Elasticsearch proceso.
6. Seleccione Confirmar en el modal.
7. Para ver el estado de la acción que ha iniciado, seleccione el nombre del nodo. Cuando se abra la página de detalles del nodo, seleccione la pestaña Eventos situada debajo del nombre del nodo para ver una lista de los eventos asociados a ese nodo.

Reiniciar un nodo de datos en Amazon Service OpenSearch

Cómo reiniciar un nodo de datos

1. Diríjase a la consola de OpenSearch servicio enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Dominios. Seleccione el nombre del dominio con el que desea trabajar.
3. Cuando se abra la página de detalles del dominio, vaya a la pestaña Estado de la instancia.
4. En Nodos de datos, seleccione el botón situado junto al nodo en el que desea reiniciar el proceso.
5. Seleccione el menú desplegable Acciones y elija Nodo de reinicio.
6. Seleccione Confirmar en el modal.
7. Para ver el estado de la acción que ha iniciado, seleccione el nombre del nodo. Cuando se abra la página de detalles del nodo, seleccione la pestaña Eventos situada debajo del nombre del nodo para ver una lista de los eventos asociados a ese nodo.

Reiniciar el proceso de OpenSearch paneles en un nodo de Amazon Service OpenSearch

Puedes reiniciar el proceso de los OpenSearch paneles (anteriormente Kibana) para recuperarte de problemas como una interfaz bloqueada, errores de carga o visualizaciones que no respondan. La opción de reiniciar los OpenSearch paneles solo está disponible para el nodo que ejecuta activamente el proceso de paneles de control. En la mayoría OpenSearch de los dominios de servicio, este proceso se ejecuta en nodos coordinadores dedicados, no en nodos de datos. Como resultado, al abrir el menú desplegable Acciones en la consola, la opción normalmente aparece solo para los nodos coordinadores. Para obtener más información, consulte the section called “Nodos coordinadores dedicados”.

Este comportamiento depende de cómo esté configurado el dominio.

- Dominios con nodos coordinadores dedicados: el proceso de Dashboards se ejecuta exclusivamente en esos nodos y solo en ellos se muestra la opción de reinicio.
- Dominios sin nodos coordinadores dedicados: en configuraciones más simples o antiguas, es posible que los Dashboards se ejecuten en un nodo de datos y allí aparezca la opción de reinicio.
- Nodos maestros: estos nodos sirven únicamente para administrar el estado y las elecciones del clúster. No ejecutan Dashboards y nunca muestran la opción de reinicio.

Para determinar qué nodo ejecuta el proceso de Dashboards, vaya a la sección de configuración del clúster de su dominio y revise las funciones de los nodos. La opción de reiniciar solo está disponible para el nodo que aloja el proceso de Dashboards.

Para reiniciar el proceso de Dashboards en un nodo

1. Diríjase a la consola de OpenSearch servicio en https://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Dominios. Seleccione el nombre del dominio con el que desea trabajar.
3. Cuando se abra la página de detalles del dominio, vaya a la pestaña Estado de la instancia.
4. En la sección para los nodos que están ejecutando el proceso de Dashboards, seleccione el botón al lado del nodo en el que usted desea reiniciar el proceso.
5. Seleccione el menú desplegable Acciones y elija Reiniciar el Dashboard/Kibana proceso.
6. Seleccione Confirmar en el modal.

7. Para ver el estado de la acción que ha iniciado, seleccione el nombre del nodo. Cuando se abra la página de detalles del nodo, seleccione la pestaña Eventos situada debajo del nombre del nodo para ver una lista de los eventos asociados a ese nodo.

Limitaciones

Las opciones administrativas presentan las siguientes limitaciones:

- Las opciones administrativas se admiten en las versiones 7.x y superiores de Elasticsearch.
- Las opciones administrativas no admiten dominios con Multi-AZ con modo de espera activado.
- El reinicio del proceso de Elasticsearch OpenSearch y el reinicio del nodo de datos se admiten en dominios con tres o más nodos de datos.
- La compatibilidad con los procesos de Dashboards y Kibana se admite en dominios con dos o más nodos de datos.
- Para reiniciar el OpenSearch proceso en un nodo o reiniciar un nodo, el dominio no debe estar en estado rojo y todos los índices deben tener las réplicas configuradas.

Trabajar con consultas directas OpenSearch de Amazon Service

Utilice Amazon OpenSearch Service Direct Query para analizar los datos de Amazon CloudWatch Logs, Amazon S3 y Amazon Security Lake sin crear canalizaciones de ingestión. Esta integración sin ETL le permite consultar datos in situ mediante OpenSearch SQL o PPL y explorarlos en Discover.

Para empezar, configure la fuente de datos en la OpenSearch consola de servicio. Para Amazon S3, utilice conexiones de dominio y cree tablas con SQL en Query Workbench. CloudWatch Los registros y Security Lake utilizan fuentes y AWS Glue Data Catalog tablas preconfiguradas.

Precios de las consultas directas

Cuando utilice consultas directas del OpenSearch Servicio, incurrirá en cargos separados por el OpenSearch Servicio y los recursos utilizados para procesar y almacenar sus datos en Amazon S3, Amazon CloudWatch Logs y Amazon Security Lake. A medida que realice consultas directas, verá los cargos por unidades de OpenSearch cómputo (OCUs) por hora, que figuran como tipo de uso de DirectQuery OCU en su factura.

Las consultas directas son de dos tipos: consultas de vista indexada e interactivas.

- Las consultas interactivas se utilizan para rellenar el selector de datos y realizar análisis de los datos en S3, CloudWatch Logs o Security Lake.

  En el caso de las consultas directas de Amazon S3, cuando ejecuta una nueva consulta desde Discover, el OpenSearch servicio inicia una nueva sesión que dura un mínimo de tres minutos. OpenSearch El servicio mantiene esta sesión activa para garantizar que las consultas posteriores se ejecuten rápidamente.

  En el CloudWatch caso de las consultas de Logs y Security Lake, OpenSearch Service gestiona cada consulta con un trabajo preajustado independiente, sin necesidad de mantener una sesión prolongada.
- Las consultas de vistas indexadas utilizan el procesamiento para mantener las vistas indexadas en Service. OpenSearch Estas consultas suelen tardar más porque ingieren una cantidad de datos variable en un índice con nombre. Al indexar los datos, puede hacer que las futuras consultas interactivas se ejecuten más rápido o desbloquear capacidades de análisis avanzadas, como paneles o alertas, que requieren un índice como referencia.

En el caso de los orígenes de datos de Amazon S3, los datos indexados se almacenan en un dominio en función del tipo de instancia adquirido. En el caso de las fuentes de datos conectadas a CloudWatch Logs y Security Lake, los datos indexados se almacenan en una colección OpenSearch sin servidor, en la que se cobra por los datos indexados (indexingOCU), los datos buscados (SearchOCU) y los datos almacenados en GB.

Para obtener más información, consulta las secciones Direct Query y Serverless de Amazon OpenSearch Service Pricing.

Limitaciones de las consultas directas

Limitaciones generales

Las siguientes limitaciones se aplican a las consultas directas del OpenSearch Servicio.

- Algunos tipos de datos no son compatibles. Los tipos de datos admitidos se limitan a Parquet, CSV y JSON.
- Si la estructura de los datos cambia con el tiempo, tendrá que actualizar las vistas indexadas o out-of-the-box las integraciones para tener en cuenta los cambios en la estructura de los datos.
- AWS CloudFormation las plantillas aún no son compatibles.
- OpenSearch Las sentencias SQL y OpenSearch PPL tienen limitaciones diferentes cuando se trabaja con OpenSearch índices en comparación con el uso de consultas directas. Direct Query admite comandos avanzados JOINs, como subconsultas y búsquedas, mientras que la compatibilidad con estos comandos en los OpenSearch índices es limitada o inexistente. Para obtener más información, consulte the section called “Comandos SQL y PPL compatibles”.

Limitaciones de Amazon S3

Si consulta datos directamente en Amazon S3, se aplican las siguientes limitaciones adicionales:

- La consulta directa para S3 solo está disponible en los dominios de OpenSearch servicio que ejecutan la OpenSearch versión 2.13 o posterior y requiere acceso a. AWS Glue Data Catalog AWS Glue Data Catalog Las tablas existentes se deben volver a crear con SQL en OpenSearch Query Workbench.

- La consulta directa para S3 requiere que se especifique un bucket de puntos de control en Amazon S3. Este bucket mantiene el estado de las vistas indexadas, incluida la hora de la última actualización y los datos ingeridos más recientemente.
- Su OpenSearch dominio y AWS Glue Data Catalog deben estar en el mismo dominio. Cuenta de AWS Su bucket de S3 puede estar en una cuenta diferente (es necesario añadir una condición a su política de IAM), pero debe estar en la Región de AWS misma cuenta que su dominio.
- OpenSearch Las consultas directas de servicio con S3 solo admiten las tablas de Spark generadas desde Query Workbench. Las tablas generadas dentro de Athena AWS Glue Data Catalog o Athena no son compatibles con la transmisión de Spark, que es necesaria para mantener las vistas indexadas.
- OpenSearch los tipos de instancias tienen limitaciones de carga útil de red de 10 MiB o 100 MiB, según el tipo de instancia específico que elijas.

Limitaciones de Amazon CloudWatch Logs

Si realiza consultas directas de datos en CloudWatch Logs, se aplican las siguientes limitaciones adicionales:

- La integración de consultas directas con CloudWatch los registros solo está disponible en las colecciones de OpenSearch servicios y en la interfaz OpenSearch de usuario.
- OpenSearch Las colecciones sin servidor tienen limitaciones de carga útil de red de 100 MiB.
- CloudWatch Logs admite el flujo de VPC y las integraciones de AWS WAF paneles instaladas desde la consola. CloudTrail

Limitaciones de Amazon Security Lake

Si consulta datos directamente en Security Lake, se aplican las siguientes limitaciones adicionales:

- La integración de consultas directas con Security Lake solo está disponible en las colecciones de OpenSearch servicios y en la interfaz de OpenSearch usuario.
- OpenSearch Las colecciones sin servidor tienen limitaciones de carga útil de red de 100 MiB.
- La administración de las tablas de Security Lake se realiza en Lake Formation.
- Security Lake solo admite vistas materializadas como vistas indexadas. No se admiten índices de cobertura.

Recomendaciones para el uso de consultas directas en Amazon OpenSearch Service

En esta página se proporcionan recomendaciones para utilizar las consultas directas de Amazon OpenSearch Service para analizar los datos de CloudWatch Logs, Amazon S3 y Amazon Security Lake. Estas prácticas recomendadas lo ayudan a optimizar el rendimiento y a garantizar una consulta eficiente sin necesidad de ingerir ni duplicar datos.

Temas

- Recomendaciones generales
- Recomendaciones para Amazon S3
- CloudWatch Registra recomendaciones
- Recomendaciones para Security Lake

Recomendaciones generales

Se recomienda que lleve a cabo las siguientes acciones al utilizar la consulta directa:

- Utilice la función `COALESCE SQL` para administrar las columnas que faltan y garantizar que se devuelvan los resultados.

- Utilice límites en sus consultas para asegurarse de no recuperar demasiados datos.
- Si tiene pensado analizar el mismo conjunto de datos varias veces, cree una vista indexada para asimilar e indexar todos los datos OpenSearch y suéltelos cuando haya completado el análisis.
- Elimine los índices y las tareas de aceleración cuando ya no sean necesarios.
- No se admiten consultas que contengan nombres de campo idénticos pero que solo difieran en mayúsculas y minúsculas (como `field1` y `FIELD1`).

  Por ejemplo, no se admiten las siguientes consultas:

```
Select AWSAccountId, AwsAccountId from LogGroup
Select a.@LogStream, b.@logStream from Table A INNER Join Table B ona.id = b.id
```

  Sin embargo, se admite la siguiente consulta porque el nombre del campo (@logStream) es idéntico en ambos grupos de registros:

```
Select a.@logStream, b.@logStream from Table A INNER Join Table B on a.id = b.id
```

- Las funciones y expresiones deben funcionar con los nombres de los campos y formar parte de una sentencia `SELECT` con un grupo de registros especificado en la cláusula `FROM`.

  Por ejemplo, no se admite esta consulta:

  ```
  SELECT cos(10) FROM LogGroup
  ```

  Se admite esta consulta:

  ```
  SELECT cos(field1) FROM LogGroup
  ```

Recomendaciones para Amazon S3

Si utiliza Amazon OpenSearch Service para realizar consultas directas de datos en Amazon S3, también le recomendamos lo siguiente:

- Ingiera datos en Amazon S3 mediante formatos de partición de año, mes, día y hora para acelerar las consultas.
- Cuando cree índices de omisión, utilice los filtros Bloom para los campos con una cardinalidad alta y los min/max índices para los campos con rangos de valores grandes. Para los campos de cardinalidad alta, considere la posibilidad de utilizar un enfoque basado en valores para mejorar la eficiencia de las consultas.
- Utilice Administración de estados de índice para mantener el almacenamiento de las vistas materializadas y los índices de cobertura.

CloudWatch Registra recomendaciones

Si utilizas Amazon OpenSearch Service para realizar consultas directas de datos en CloudWatch los registros, también te recomendamos lo siguiente:

- Al buscar varios grupos de registros en una consulta, utilice la sintaxis adecuada. Para obtener más información, consulte the section called “Funciones de grupos de registros múltiples”.
- Cuando utilice comandos SQL o PPL, encierre determinados campos entre comillas invertidas para consultarlos correctamente. Las comillas invertidas son necesarias para los campos

con caracteres especiales (no alfabéticos ni numéricos). Por ejemplo, encierre `@message`, `Operation.Export,` y `Test::Field` entre comillas invertidas. No es necesario incluir las columnas con nombres exclusivamente alfabéticos entre comillas invertidas.

Ejemplo de consulta con campos sencillos:

```
SELECT SessionToken, Operation, StartTime  FROM `LogGroup-A`
LIMIT 1000;
```

Consulta similar con comillas invertidas agregadas:

```
SELECT `@SessionToken`, `@Operation`, `@StartTime` FROM `LogGroup-A`
LIMIT 1000;
```

Recomendaciones para Security Lake

Si utilizas Amazon OpenSearch Service para realizar consultas directas de datos en Security Lake, también te recomendamos lo siguiente:

- Compruebe el estado de Security Lake y asegúrese de que funcione sin problemas. Para ver los pasos detallados de solución de problemas, consulte Solución de problemas del estado del lago de datos en la Guía del usuario de Amazon Security Lake.
- Compruebe el acceso a su consulta:
  - Si realiza consultas en Security Lake desde una cuenta diferente a la cuenta de administrador delegado de Security Lake, configure un suscriptor con acceso de consulta en Security Lake.
  - Si realiza consultas en Security Lake desde la misma cuenta, compruebe si hay algún mensaje en Security Lake sobre el registro de los buckets de S3 gestionados. LakeFormation
- Explore las plantillas de consulta y los paneles prediseñados para iniciar su análisis con rapidez.
- Familiarícese con Open Cybersecurity Schema Framework (OCSF) y Security Lake:
  - Revise los ejemplos de mapeo de esquemas de las AWS fuentes en el repositorio de OCSF GitHub
  - Aprenda a realizar consultas de Security Lake de forma eficaz consultando las consultas de Security Lake para la versión 2 del AWS código fuente (OCSF 1.1.0)
  - Mejore el rendimiento de las consultas mediante el uso de particiones: `accountid`, `region` y `time_dt`.

- Familiarícese con la sintaxis de SQL, que Security Lake admite para realizar consultas. Para obtener más información, consulte the section called “Comandos SQL compatibles”.

Cuotas de las consultas directas

Su cuenta tiene las siguientes cuotas relacionadas con las consultas directas OpenSearch del Servicio.

Cuotas para Amazon S3

Cada vez que inicia una consulta a una fuente de datos de Amazon S3, OpenSearch Service abre una sesión y la mantiene activa durante al menos tres minutos. Esto reduce la latencia de las consultas al eliminar el tiempo de inicio de la sesión en las consultas posteriores.

Description (Descripción)	Máximo	Puede anular
Conexiones por dominio	10	Sí
Origen de datos por dominio	20	Sí
Índices por dominio	5	Sí
Sesiones simultáneas por origen de datos	10	Sí
OCU máxima por consulta	60	Sí
Tiempo máximo de ejecución de consulta (minutos)	30	Sí
Máximo OCUs por aceleración	20	Sí
Almacenamiento efímero máximo	20	Sí

Cuotas de CloudWatch registros

> **Note**
>
> Si desea realizar consultas directas con CloudWatch Logs Insights, asegúrese de consultarthe section called “CloudWatch Registra Insights”.

Description (Descripción)	Valor	¿El límite es flexible?	Notas
Límite de TPS a nivel de cuenta en todas las consultas directas APIs	3 TPS	Sí	
Número máximo de orígenes de datos	20	Sí	El límite es por. Cuenta de AWS
Número máximo de índices o vistas materializadas que se actualizan automáticamente	30	Sí	El límite es por origen de datos.
Máximo de consultas simultáneas	30	Sí	El límite es por fuente de datos y se aplica a las consultas en estado pendiente o en ejecución. Incluye consultas interactivas (por ejemplo, comandos de recuperación de datos`SELECT`) y consultas de índice (por ejemplo, operaciones como`CREATE/ALTER`).
OCU simultáneas máximas por consulta	512	Sí	OpenSearch Unidades de cómputo (OCU). El límite se basa en 15 ejecutores y 1 controlador, cada uno con 16 vCPU y 32 GB de memoria.

Description (Descripción)	Valor	¿El límite es flexible?	Notas
			Representa la potencia de procesamiento simultáneo.
Tiempo máximo de ejecución de consulta en minutos	60	No	El límite se aplica a las consultas de OpenSearch PPL/SQL en Logs Insights. CloudWatch
Periodo para eliminar una consulta obsoleta IDs	90 días	Sí	Este es el período de tiempo después del cual el OpenSearch Servicio purga los metadatos de las consultas de las entradas más antiguas. Por ejemplo, se llama GetDirectQuery o GetDirect QueryResult se produce un error en el caso de consultas de más de 90 días.

Cuotas para Security Lake

Description (Descripción)	Valor	¿El límite es flexible?	Notas
Límite de TPS a nivel de cuenta en todas las consultas directas APIs	3 TPS	Sí	
Número máximo de orígenes de datos	20	Sí	El límite es por. Cuenta de AWS
Número máximo de índices o vistas materializadas que se actualizan automáticamente	30	Sí	Se aplica un límite por origen de datos. Solo incluye índices y vistas materializadas (MVs) con la actualización automática establecida en true.

Description (Descripción)	Valor	¿El límite es flexible?	Notas
Máximo de consultas simultáneas	30	Sí	El límite se aplica a las consultas en estado pendiente o en ejecución. Incluye consultas interactivas (por ejemplo, comandos de recuperación de datos como `SELECT`) y consultas de índice (por ejemplo, operaciones como `CREATE`, `ALTER` y `DROP`).
OCU simultáneas máximas por consulta	512	Sí	OpenSearch Unidades de cómputo (OCU). El límite se basa en 15 ejecutores y 1 controlador, cada uno con 16 vCPU y 32 GB de memoria. Representa la potencia de procesamiento simultáneo.
Tiempo máximo de ejecución de consulta en minutos	30	No	Se aplica solo a consultas interactivas (por ejemplo, comandos de recuperación de datos como `SELECT`). En el caso de las consultas `REFRESH`, el límite es de 6 horas.
Período para purgar una consulta obsoleta IDs	90 días	Sí	Este es el período de tiempo después del cual el OpenSearch Servicio purga los metadatos de las consultas de las entradas más antiguas. Por ejemplo, se llama GetDirectQuery o GetDirect QueryResult se produce un error en el caso de consultas de más de 90 días.

Compatible Regiones de AWS

Regiones de AWS Se admite lo siguiente para las consultas directas de OpenSearch servicio en Amazon S3, CloudWatch Logs y Security Lake:

Disponible Regiones de AWS para Amazon S3

- Asia-Pacífico (Hong Kong)
- Asia-Pacífico (Mumbai)
- Asia-Pacífico (Seúl)
- Asia-Pacífico (Singapur)
- Asia-Pacífico (Sídney)
- Asia-Pacífico (Tokio)
- Canadá (centro)
- Europa (Fráncfort)
- Europa (Irlanda)
- Europa (Estocolmo)
- Este de EE. UU. (Norte de Virginia)
- Este de EE. UU. (Ohio)
- Oeste de EE. UU. (Oregón)

Disponible Regiones de AWS para CloudWatch registros

- Asia-Pacífico (Mumbai)
- Asia-Pacífico (Hong Kong)
- Asia-Pacífico (Osaka)
- Asia-Pacífico (Seúl)
- Asia-Pacífico (Singapur)
- Asia-Pacífico (Sídney)
- Asia-Pacífico (Tokio)
- Canadá (centro)
- Europa (Fráncfort)
- Europa (Irlanda)
- Europa (Estocolmo)
- Europa (Milán)

- Europa (España)
- Este de EE. UU. (Norte de Virginia)
- Este de EE. UU. (Ohio)
- Oeste de EE. UU. (Oregón)
- Oeste de EE. UU. (Norte de California)
- Europa (París)
- Europa (Londres)
- América del Sur (São Paulo)

Disponible Regiones de AWS para Security Lake

- Asia-Pacífico (Mumbai)
- Asia-Pacífico (Singapur)
- Asia-Pacífico (Sídney)
- Asia-Pacífico (Tokio)
- Canadá (centro)
- Europa (Fráncfort)
- Europa (Irlanda)
- Europa (Estocolmo)
- Este de EE. UU. (Norte de Virginia)
- Este de EE. UU. (Ohio)
- Oeste de EE. UU. (Oregón)
- Europa (París)
- Europa (Londres)
- América del Sur (São Paulo)

Consulta directa de datos de Amazon S3 en Service OpenSearch

En esta sección, se explica el proceso de creación y configuración de la integración de una fuente de datos en Amazon OpenSearch Service, lo que le permitirá consultar y analizar de forma eficiente los datos almacenados en Amazon S3.

En las siguientes páginas, aprenderá a configurar una fuente de datos de consulta directa de Amazon S3, a conocer los requisitos previos necesarios y a seguir los step-by-step procedimientos tanto con la API como con la API de Consola de administración de AWS servicio. OpenSearch También se describen los próximos pasos importantes, como la asignación de AWS Glue Data Catalog funciones y la configuración de los controles de acceso en los paneles de control. OpenSearch

Temas

- Creación de una integración de fuente de datos de Amazon S3 en OpenSearch Service
- Configuración y consulta de una fuente de datos de S3 en los paneles OpenSearch

Creación de una integración de fuente de datos de Amazon S3 en OpenSearch Service

Puede crear una nueva fuente de datos de consulta directa de Amazon S3 para OpenSearch Service a través de la API Consola de administración de AWS o la API. Cada nueva fuente de datos la utiliza AWS Glue Data Catalog para administrar las tablas que representan los buckets de Amazon S3.

Temas

- Requisitos previos
- Procedimiento
- Siguientes pasos
- Asigne la función AWS Glue Data Catalog
- Recursos adicionales

Requisitos previos

Antes de comenzar, asegúrese de que ha revisado la siguiente documentación:

- the section called “Limitaciones de Amazon S3”
- the section called “Recomendaciones para Amazon S3”
- the section called “Cuotas para Amazon S3”

Antes de poder crear una fuente de datos, debe disponer de los siguientes recursos: Cuenta de AWS

- Un OpenSearch dominio con la versión 2.13 o posterior. Esta es la base para configurar la integración de consultas directas. Para obtener instrucciones acerca de cómo configurarlo, consulte the section called “Creación de dominios OpenSearch de servicio”.
- Uno o varios buckets de S3. Deberá especificar los buckets que contienen los datos que desea consultar y un bucket en el cual almacenar los puntos de control de las consultas. Para obtener instrucciones acerca de cómo crear un bucket de S3, consulte Crear un bucket en la Guía del usuario de Amazon S3.
- (Opcional) Una o más AWS Glue tablas.La consulta de datos en Amazon S3 requiere que tenga tablas configuradas en AWS Glue Data Catalog para que apunten a los datos de S3. Debe crear las tablas con OpenSearch Query Workbench. Las tablas de Hive existentes no son compatibles.

  Si es la primera vez que configura un origen de datos de Amazon S3, debe crear un origen de datos de administrador para configurar todas las tablas de AWS Glue Data Catalog . Puede hacerlo instalando OpenSearch out-of-the-box integraciones o utilizando OpenSearch Query Workbench para crear tablas SQL personalizadas para casos de uso avanzados. Para ver ejemplos sobre la creación de tablas para registros de VPC y AWS WAF, consulte la documentación sobre GitHub VPC y. CloudTrail CloudTrailAWS WAF Tras crear las tablas, puede crear nuevos orígenes de datos de Amazon S3 y restringir el acceso a tablas limitadas.
- (Opcional) Un rol de IAM creado manualmente. Puede usar este rol para administrar el acceso al origen de datos. Como alternativa, puede hacer que OpenSearch Service le cree un rol automáticamente con los permisos necesarios. Si decide utilizar un rol de IAM creado manualmente, siga las instrucciones que se indican en the section called “Permisos necesarios para los roles de IAM creados manualmente”.

Procedimiento

Puede configurar una fuente de datos de consulta directa en un dominio con la API de OpenSearch servicio Consola de administración de AWS o con la API.

Para configurar una fuente de datos mediante la Consola de administración de AWS

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Dominios.
3. Seleccione el dominio para el que desea configurar un nuevo origen de datos. Se abrirá la página de detalles del dominio.

4. Seleccione la pestaña Conexiones que aparece debajo de los detalles generales del dominio y busque la sección Consulta directa.
5. Seleccione Configurar origen de datos.
6. Escriba un nombre y una descripción opcional para el nuevo origen de datos.
7. Elija Amazon S3 con AWS Glue Data Catalog.
8. En Configuración del permiso de acceso de IAM, elija cómo administrar el acceso.

    a. Si desea crear automáticamente un rol para este origen de datos, siga estos pasos:

        i. Seleccione Crear un nuevo rol.

        ii. Ingrese un nombre para el rol de IAM.

        iii. Seleccione uno o más buckets de S3 que contengan datos que desee consultar.

        iv. Seleccione un bucket de S3 de puntos de control en el cual almacenar los puntos de control de las consultas.

        v. Seleccione una o más AWS Glue bases de datos o tablas para definir qué datos se pueden consultar. Si aún no se han creado las tablas, proporcione acceso a la base de datos predeterminada.

    b. Si desea utilizar un rol existente que usted administra, siga estos pasos:

        i. Seleccione Usar un rol existente.

        ii. Seleccione un rol existente en el menú desplegable.

    > **Note**
    >
    > Cuando utilice su propio rol, debe asegurarse de que tiene todos los permisos requeridos. Para ello, adjunte las políticas necesarias desde la consola de IAM. Para obtener más información, consulte el ejemplo de política en the section called “Permisos necesarios para los roles de IAM creados manualmente”.

9. Elija Configurar. Esto abre la pantalla de detalles de la fuente de datos con una URL del OpenSearch panel de control. Puede navegar hasta esta URL para completar los siguientes pasos.

OpenSearch API de servicio

Utilice la operación de la AddDataSourceAPI para crear una nueva fuente de datos en su 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"
}
```

Siguientes pasos

Visite los OpenSearch paneles

Después de crear una fuente de datos, OpenSearch Service le proporciona un enlace a los OpenSearch paneles de control. Puede usarlo para configurar el control de acceso, definir tablas, instalar out-of-the-box integraciones y consultar sus datos.

Para obtener más información, consulte the section called “Configuración de un origen de datos de S3”.

Asigne la función AWS Glue Data Catalog

Si ha habilitado un control de acceso detallado después de crear una fuente de datos, debe asignar a los usuarios que no sean administradores a una función de IAM con AWS Glue Data Catalog acceso para poder ejecutar consultas directas. Para crear manualmente un rol `glue_access` de backend que pueda asignar al rol de IAM, siga estos pasos:

> **Note**
>
> Los índices se utilizan para cualquier consulta realizada en el origen de datos. Un usuario con acceso de lectura al índice de solicitudes de un origen de datos determinado puede leer todas las consultas realizadas en ese origen de datos. Un usuario con acceso de lectura

al índice de resultados puede leer los resultados de todas las consultas realizadas en ese origen de datos.

1. En el menú principal de los OpenSearch paneles, seleccione Seguridad, Funciones y Crear funciones.
2. Asigne el rol glue_access.
3. En Permisos de clúster, seleccione `indices:data/write/bulk*`, `indices:data/read/scroll`, `indices:data/read/scroll/clear`.
4. En Índice, introduzca los siguientes índices a los que quiere conceder acceso al usuario con el rol:
   - `.query_execution_request_<name of data source>`
   - `query_execution_result_<name of data source>`
   - `.async-query-scheduler`
   - `flint_*`
5. En Permisos de índice, seleccione `indices_all`.
6. Seleccione Crear.
7. Seleccione Usuarios asignados, Administrar mapeo.
8. En Roles de backend, agregue el ARN del rol de AWS Glue que necesita permiso para llamar a su dominio.

   ```
   arn:aws:iam::account-id:role/role-name
   ```
9. Seleccione Asignar y confirme que el rol aparece en Usuarios asignados.

Para obtener más información sobre la asignación de roles, consulte the section called “Mapear roles a usuarios”.

Recursos adicionales

Permisos necesarios para los roles de IAM creados manualmente

Al crear un origen de datos para su dominio, elige un rol de IAM para administrar el acceso a sus datos. Tiene dos opciones:

1. Crear un nuevo rol de IAM automáticamente

2. Utilizar un rol de IAM creado manualmente

Si utiliza un rol creado manualmente, debe asociar los permisos correctos al rol. Los permisos deben permitir el acceso a la fuente de datos específica y permitir que OpenSearch Service asuma la función. Esto es necesario para que el OpenSearch Servicio pueda acceder a sus datos e interactuar con ellos de forma segura.

El siguiente ejemplo de política muestra los permisos con privilegios mínimos necesarios para crear y administrar un origen de datos. Si tiene permisos más generales, como `s3:*` o la política de `AdminstratorAccess`, estos engloban los permisos con privilegios mínimos de la política de ejemplo.

En el siguiente ejemplo de política, *`placeholder text`* sustitúyalo por su propia información.

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/*"
         ]
      }
   ]
}
```

Para admitir buckets de Amazon S3 en diferentes cuentas, tendrá que incluir una condición en la política de Amazon S3 y agregar la cuenta adecuada.

En el siguiente ejemplo de estado, *placeholder text* sustitúyalo por su propia información.

```
"Condition": {
               "StringEquals": {
                   "aws:ResourceAccount": "{{accountId}}"
               }
```

El rol debe tener la siguiente política de confianza, que especifica el ID de destino.

JSON

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

```
            "Effect":"Allow",
            "Principal":{
               "Service": "directquery.opensearchservice.amazonaws.com"
            },
            "Action":"sts:AssumeRole"
         }
      ]
}
```

Para obtener instrucciones sobre cómo crear los roles, consulte Creación de un rol mediante políticas de confianza personalizadas.

Si tiene activado el control de acceso detallado en el OpenSearch Servicio, se creará automáticamente un nuevo rol de control de acceso OpenSearch detallado para su fuente de datos. El nombre del nuevo rol de control de acceso detallado será `AWS OpenSearchDirectQuery <name of data source>`.

De forma predeterminada, el rol solo tiene acceso a los índices de orígenes de datos de consulta directa. Si bien puede configurar el rol para limitar o conceder el acceso a su origen de datos, se recomienda no ajustar el acceso de este rol. Si elimina el origen de datos, se eliminará este rol. Esto eliminará el acceso de los demás usuarios si están asignados al rol.

Configuración y consulta de una fuente de datos de S3 en los paneles OpenSearch

Ahora que ha creado su origen de datos, puede configurar los ajustes de seguridad, definir las tablas de Amazon S3 o configurar la indexación acelerada de datos. En esta sección, se explican varios casos de uso de la fuente de datos en los OpenSearch paneles de control antes de realizar consultas en los datos.

Para configurar las siguientes secciones, primero debe navegar hasta su fuente de datos en los OpenSearch paneles. En el menú de navegación izquierdo, en Administración, elija Origen de datos. En Administrar orígenes de datos, seleccione el nombre del origen de datos que creó en la consola.

Creación de tablas de Spark con Query Workbench

Las consultas directas de OpenSearch Service a Amazon S3 utilizan las tablas de Spark incluidas en AWS Glue Data Catalog. Puede crear tablas desde Query Workbench sin tener que salir de los OpenSearch paneles.

Para administrar las bases de datos y tablas existentes en el origen de datos, o para crear tablas nuevas en las que desee utilizar consultas directas, seleccione Query Workbench en el menú de navegación de la izquierda y elija el origen de datos de Amazon S3 en el menú desplegable de orígenes de datos.

Para configurar una tabla para los registros de flujo de VPC almacenados en S3 en formato Parquet, ejecute la siguiente consulta:

```
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"
```

Tras crear la tabla, ejecute la siguiente consulta para asegurarse de que es compatible con las consultas directas:

```
MSCK REPAIR TABLE  datasourcename.databasename.vpclogstable
```

Configure integraciones para los tipos de registro más populares AWS

Puede integrar los tipos de AWS registro almacenados en Amazon S3 con OpenSearch Service. Utilice los OpenSearch paneles para instalar integraciones que creen AWS Glue Data Catalog tablas, consultas guardadas y paneles. Estas integraciones utilizan vistas indexadas para mantener los paneles actualizados.

Para obtener instrucciones sobre cómo instalar una integración, consulte Instalación de un activo de integración en la documentación. OpenSearch

Cuando seleccione una integración, asegúrese de que tenga la etiqueta `S3 Glue`.

Al configurar la integración, especifique Conexión a S3 para el tipo de conexión. A continuación, seleccione el origen de datos para la integración, la ubicación de Amazon S3 de los datos, el punto

de control para administrar la indexación de la aceleración y los activos necesarios para su caso de uso.

> **Note**
>
> Asegúrese de que el bucket de S3 de su punto de control tenga permisos de escritura para la ubicación del punto de control. Sin estos permisos, se producirá un error en las aceleraciones de la integración.

Cómo configurar el control de acceso

En la página de detalles de su origen de datos, busque la sección Controles de acceso y elija Editar. Si el dominio tiene habilitado un control de acceso detallado, elija Restringido y seleccione las funciones a las que desee proporcionar acceso al nuevo origen de datos. También puede elegir Solo administrador si solo quiere que el administrador tenga acceso al origen de datos.

> **Important**
>
> Los índices se utilizan para cualquier consulta realizada en el origen de datos. Un usuario con acceso de lectura al índice de solicitudes de un origen de datos determinado puede leer todas las consultas realizadas en ese origen de datos. Un usuario con acceso de lectura al índice de resultados puede leer los resultados de todas las consultas realizadas en ese origen de datos.

Consulta de datos de S3 en Discover OpenSearch

Después de configurar las tablas y configurar la aceleración de consultas opcional deseada, ya puede empezar a analizar sus datos. Para consultar sus datos, seleccione el origen de datos en el menú desplegable. Si utiliza Amazon S3 y OpenSearch Dashboards, vaya a Discover y seleccione el nombre de la fuente de datos.

Si utiliza un índice de omisión o no ha creado uno, puede utilizar SQL o PPL para consultar los datos. Si ha configurado una vista materializada o un índice de cobertura, ya tiene un índice y puede usar el lenguaje de consultas de Dashboards (DQL) en todo Dashboards. También puede usar PPL con el complemento Observabilidad y SQL con el complemento Query Workbench. Actualmente, solo los complementos Observabilidad y Query Workbench admiten PPL y SQL. Para consultar datos mediante la API de OpenSearch servicio, consulte la documentación de la API asíncrona.

> **Note**
>
> No se admiten todas las instrucciones, comandos y funciones de SQL y PPL. Para obtener una lista de los comandos admitidos, consulte the section called “Comandos SQL y PPL compatibles”.
> Si ha creado una vista materializada o un índice de cobertura, puede utilizar DQL para consultar sus datos, siempre que los haya indexado dentro.

Resolución de problemas

Puede haber casos en los que los resultados no devuelvan los resultados esperados. Si tiene algún problema, asegúrese de seguir las the section called “Recomendaciones”.

Consulta directa de los datos de Amazon CloudWatch Logs en el servicio OpenSearch

En esta sección, se explica el proceso de creación y configuración de la integración de una fuente de datos en Amazon OpenSearch Service, lo que le permitirá consultar y analizar de forma eficaz los datos almacenados en CloudWatch Logs.

En las siguientes páginas, aprenderá a configurar una fuente de datos de consulta directa de CloudWatch Logs, a conocer los requisitos previos necesarios y a seguir los step-by-step procedimientos utilizando el. Consola de administración de AWS

Temas

- Creación de una integración de fuentes de datos de Amazon CloudWatch Logs en OpenSearch Service
- Configuración y consulta de una fuente de datos CloudWatch de Logs en OpenSearch los paneles

Creación de una integración de fuentes de datos de Amazon CloudWatch Logs en OpenSearch Service

Si utiliza Amazon OpenSearch Serverless para sus necesidades de observabilidad, ahora puede analizar sus Amazon CloudWatch Logs sin copiar ni ingerir los datos en Service. OpenSearch Esta

capacidad aprovecha la consulta directa para consultar datos, de forma similar al análisis de datos en Amazon S3 desde OpenSearch Service. Para empezar, puede crear una nueva fuente de datos conectada desde la consola de AWS administración.

Puede crear una nueva fuente de datos para analizar los datos de CloudWatch Logs sin tener que crear Amazon OpenSearch Serverless para consultar directamente los registros operativos de CloudWatch Logs. Esto le permite analizar los datos operativos a los que ha accedido y que se encuentran fuera del OpenSearch Servicio. Al realizar consultas entre el OpenSearch servicio y CloudWatch los registros, puede empezar a analizar los registros de los CloudWatch registros y, después, volver a supervisar las fuentes de datos OpenSearch sin tener que cambiar de herramienta.

Para utilizar esta función, debe crear una fuente de datos de consulta directa de CloudWatch Logs para OpenSearch Service a través de la consola AWS de administración.

Temas

- Requisitos previos
- Procedimiento
- Siguientes pasos
- Recursos adicionales

Requisitos previos

Antes de comenzar, asegúrese de que ha revisado la siguiente documentación:

- the section called “Limitaciones de Amazon CloudWatch Logs”
- the section called “CloudWatch Registra recomendaciones”
- the section called “Cuotas de CloudWatch registros”

Antes de poder crear una fuente de datos, debe tener los siguientes recursos en su cuenta Cuenta de AWS:

- Habilite CloudWatch los registros. Configure CloudWatch los registros para recopilar registros en el Cuenta de AWS mismo lugar que su OpenSearch recurso. Para obtener instrucciones, consulta Cómo empezar a usar CloudWatch Logs en la guía del usuario de Amazon CloudWatch Logs.
- Uno o más grupos de CloudWatch registros.Puede especificar los grupos de registros que contienen los datos que desea consultar. Para obtener instrucciones sobre cómo crear un grupo

de registros, consulte Crear un grupo de CloudWatch registros en Logs en la guía del usuario de Amazon CloudWatch Logs.

- (Opcional) Un rol de IAM creado manualmente.Puede usar este rol para administrar el acceso al origen de datos. Como alternativa, puede hacer que OpenSearch Service cree un rol para usted automáticamente con los permisos necesarios. Si decide utilizar un rol de IAM creado manualmente, siga las instrucciones que se indican en the section called “Permisos necesarios para los roles de IAM creados manualmente”.

Procedimiento

Puede configurar un origen de datos de consulta de nivel de colección con la Consola de administración de AWS.

Para configurar una fuente de datos a nivel de recopilación mediante el Consola de administración de AWS

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación de la izquierda, vaya a Administración central y seleccione Orígenes de datos conectados.
3. Elija Conectar.
4. Elija CloudWatchel tipo de fuente de datos.
5. Elija Siguiente.
6. En Detalles de la conexión de datos, escriba un nombre y una descripción opcional.
7. En Rol de IAM, elija cómo administrar el acceso a los grupos de registros.
    a. Si desea crear automáticamente un rol para este origen de datos, siga estos pasos:
        i. Seleccione Crear un nuevo rol.
        ii. Ingrese un nombre para el rol de IAM.
        iii. Seleccione uno o más grupos de registros para definir qué datos se pueden consultar.
    b. Si desea utilizar un rol existente que usted administra, siga estos pasos:
        i. Seleccione Usar un rol existente.
        ii. Seleccione un rol existente en el menú desplegable.

> **Note**
>
> Cuando utilice su propio rol, debe asegurarse de que tiene todos los permisos requeridos. Para ello, adjunte las políticas necesarias desde la consola de IAM. Para obtener más información, consulte the section called “Permisos necesarios para los roles de IAM creados manualmente”.

8. (Opcional) En Política de acceso, configure una política de acceso para la fuente de datos. Las políticas de acceso controlan si se acepta o rechaza una solicitud a la fuente de datos de consulta directa del OpenSearch Servicio. Si no configura una política de acceso, solo el propietario de la fuente de datos tiene acceso. Puede configurar la política de acceso para habilitar el acceso entre cuentas y permitir que los directores de otras cuentas accedan Cuentas de AWS a la fuente de datos.

   Puede crear una política de acceso mediante el editor visual o proporcionando un documento de política JSON. Con el editor visual, puede permitir o denegar el acceso especificando un Cuenta de AWS ID principal, un ARN de cuenta, un ARN de usuario de IAM, un ARN de rol de IAM, una dirección IP de origen o un bloque de CIDR. El editor visual admite hasta 10 elementos. Para definir una política con más de 10 elementos, utilice el editor JSON.

   También puede elegir Política de importación para importar una política de acceso existente desde otra fuente de datos.
9. (Opcional) En Etiquetas, agregue etiquetas al origen de datos.
10. Elija Siguiente.
11. En Configurar OpenSearch, elige cómo configurarlo OpenSearch.

    a. Utilice la configuración predeterminada:

       - Revise los nombres de los recursos y la configuración de retención de datos predeterminados. Le sugerimos que utilice nombres personalizados.

         Si utilizas la configuración predeterminada, se crea una nueva OpenSearch aplicación y un espacio de trabajo de Essentials sin coste adicional. OpenSearch le permite analizar varias fuentes de datos. Incluye espacios de trabajo, que proporcionan experiencias personalizadas para casos de uso populares. Los espacios de trabajo admiten el

control de acceso, lo que le permite crear espacios privados para sus casos de uso y compartirlos solo con sus colaboradores.

b. Use ajustes personalizados:

   i. Elija Personalizar.

   ii. Edite el nombre de la colección y la configuración de retención de datos según sea necesario.

   iii. Seleccione la OpenSearch aplicación y el espacio de trabajo que desee utilizar.

12. Elija Siguiente.
13. Revise sus opciones y seleccione Editar si necesita hacer algún cambio.
14. Seleccione Conector para configurar el origen de datos. Permanezca en esta página mientras se crea el origen de datos. Cuando esté listo, accederá a la página de detalles del origen de datos.

Siguientes pasos

Visite los OpenSearch paneles

Después de crear una fuente de datos, OpenSearch Service le proporciona una URL de OpenSearch Dashboards. Se usa para configurar el control de acceso, definir tablas, configurar paneles basados en tipos de registro para los tipos de registro más populares y consultar los datos mediante SQL o PPL.

Para obtener más información, consulte the section called “Configuración de una fuente de datos de CloudWatch Logs”.

Recursos adicionales

Permisos necesarios para los roles de IAM creados manualmente

Al crear un origen de datos, se elige un rol de IAM para administrar el acceso a los datos. Tiene dos opciones:

1. Crear un nuevo rol de IAM automáticamente
2. Utilizar un rol de IAM creado manualmente

Si utiliza un rol creado manualmente, debe asociar los permisos correctos al rol. Los permisos deben permitir el acceso a la fuente de datos específica y permitir que OpenSearch Service asuma la

función. Esto es necesario para que el OpenSearch Servicio pueda acceder a sus datos e interactuar con ellos de forma segura.

El siguiente ejemplo de política muestra los permisos con privilegios mínimos necesarios para crear y administrar un origen de datos. Si tiene permisos más generales, como `logs:*` o la política de `AdminstratorAccess`, estos engloban los permisos con privilegios mínimos de la política de ejemplo.

En el siguiente ejemplo de política, *`placeholder text`* sustitúyalo por su propia información.

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:*"
            ]
        }
    ]
}
```

El rol debe tener la siguiente política de confianza, que especifica el ID de destino.

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"
                }
            }
        }
    ]
}
```

Para obtener instrucciones sobre cómo crear los roles, consulte Creación de un rol mediante políticas de confianza personalizadas.

De forma predeterminada, el rol solo tiene acceso a los índices de orígenes de datos de consulta directa. Si bien puede configurar el rol para limitar o conceder el acceso a su origen de datos, se recomienda no ajustar el acceso de este rol. Si elimina el origen de datos, se eliminará este rol. Esto eliminará el acceso de los demás usuarios si están asignados al rol.

Ejemplo de política de acceso para una fuente de datos de consulta directa

Las políticas de acceso a las fuentes de datos de consulta directa siguen la sintaxis de las políticas de IAM. El documento de política debe estar en un formato JSON válido. El siguiente ejemplo de política otorga un Cuenta de AWS acceso específico a la fuente de datos de la consulta directa.

En el siguiente ejemplo de política, sustituya la *placeholder text* política por su propia información.

```
{
```

```
 "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"
   }
 ]
}
```

Si no configura una política de acceso, solo el propietario de la fuente de datos tiene acceso a la fuente de datos.

Configuración y consulta de una fuente de datos CloudWatch de Logs en OpenSearch los paneles

Ahora que ha creado su fuente de datos, puede empezar a utilizarla OpenSearch en los paneles. En esta sección, se explican varios casos de uso de la fuente de datos en los OpenSearch paneles de control.

Consulte grupos de registros desde la página Discover

En la página OpenSearch Descubrir, puede usar la nueva fuente de datos de consulta directa que configuró para consultar sus grupos de CloudWatch registros. Para ello, elija Explorar registro y, a continuación, utilice la barra de búsqueda para crear la consulta mediante SQL o PPL. Puede filtrar, ordenar y visualizar los datos devueltos por sus grupos de registros. Para saber qué instrucciones, comandos y limitaciones se admiten en la integración de CloudWatch los registros, consultethe section called “Comandos SQL y PPL compatibles”.

Crear una vista de panel para el origen de datos

Al utilizar el OpenSearch Servicio, puede analizar rápidamente los tipos de AWS registros más populares mediante plantillas de panel prediseñadas. Para CloudWatch los registros, hay plantillas

para los registros de VPC CloudTrail AWS WAF, y Network Firewall. Estas plantillas le permiten crear rápidamente un panel adaptado a sus datos específicos. Incluyen paneles diseñados para ese tipo de registro específico. Esto le permite empezar a analizar rápidamente estas populares fuentes de AWS registro, sin tener que compilarlo todo desde cero.

> **Note**
>
> Los paneles utilizan vistas indexadas, que incorporan los datos de los CloudWatch registros mediante unidades de OpenSearch cálculo de consulta directa (OCUs), así como la indexaciónOCUs, la búsqueda y el almacenamiento de colecciones sin servidor. OCUs

Siga estos pasos para crear un panel con una de estas plantillas prediseñadas, de modo que pueda empezar a explorar y analizar sus datos de inmediato.

Cómo crear un panel

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Administración central y, a continuación, Orígenes de datos conectados.
3. Seleccione el origen de datos para abrir la página de detalles.
4. Elija Crear panel.
5. Elija el tipo de panel que desea crear.
6. Introduzca un nombre para su panel de control.
7. Escriba una descripción opcional para el panel de control.
8. Seleccione uno o más grupos de registros para verlos en el panel de control.
9. Elija la frecuencia con la que quiere actualizar los datos del panel.
10. Elige qué OpenSearch espacio de trabajo quieres usar.
    a. Para crear un nuevo espacio de trabajo, seleccione Crear nuevo espacio de trabajo y escriba un nombre.
    b. Para usar un espacio de trabajo existente, seleccione Seleccionar espacio de trabajo existente.
11. Elija Crear panel.

Consulta de datos CloudWatch de registros en Discover OpenSearch

Para consultar sus datos, seleccione el origen de datos en el menú desplegable. Si usa CloudWatch Logs, vaya a Discover desde su espacio de trabajo de Essentials y comience a consultar datos mediante OpenSearch SQL o el lenguaje de procesamiento canalizado (PPL). Para obtener una lista de los comandos admitidos, consulte the section called “Comandos SQL y PPL compatibles”.

> **Note**
>
> Si ha creado una vista materializada, puede usar DQL para consultar sus datos, siempre que los haya indexado dentro.

Resolución de problemas

Puede haber casos en los que los resultados no devuelvan los resultados esperados. Si tiene algún problema, asegúrese de seguir las the section called “Recomendaciones”.

Consulta directa de los datos de Amazon Security Lake en Service OpenSearch

En esta sección, se explica el proceso de creación y configuración de una integración de fuentes de datos en Amazon OpenSearch Service, lo que le permitirá consultar y analizar de forma eficiente los datos almacenados en Security Lake.

En las páginas siguientes, aprenderá a configurar una fuente de datos de consulta directa de Security Lake, a analizar los requisitos previos necesarios y a seguir los step-by-step procedimientos utilizando el. Consola de administración de AWS

Temas

- Creación de una integración de fuentes de datos de Amazon Security Lake en OpenSearch Service
- Configuración y consulta de una fuente de datos de Security Lake en OpenSearch los paneles

Creación de una integración de fuentes de datos de Amazon Security Lake en OpenSearch Service

Puede usar Amazon OpenSearch Serverless para consultar directamente los datos de seguridad en Amazon Security Lake. Para ello, debe crear una fuente de datos que le permita utilizar funciones OpenSearch sin ETL en los datos de Security Lake. Al crear un origen de datos, puede buscar directamente los datos almacenados en Security Lake, obtener información sobre ellos y analizarlos. Puede acelerar el rendimiento de sus consultas y utilizar OpenSearch análisis avanzados en determinados conjuntos de datos de Security Lake mediante la indexación bajo demanda.

Temas

- Requisitos previos
- Procedimiento
- Siguientes pasos
- Recursos adicionales

Requisitos previos

Antes de comenzar, asegúrese de que ha revisado la siguiente documentación:

- the section called “Limitaciones de Amazon Security Lake”
- the section called “Recomendaciones para Security Lake”
- the section called “Cuotas para Security Lake”

Antes de poder crear un origen de datos, haga lo siguiente en Security Lake:

- Habilitar Security Hub. Configure Security Lake para recopilar registros en el mismo lugar Región de AWS que su OpenSearch recurso. Para obtener instrucciones, consulte Introducción a Amazon Security Lake en la guía del usuario de Amazon Security Lake.
- Configure los permisos de Security Lake. Asegúrese de haber aceptado los permisos de los roles vinculados al servicio para la administración de recursos y de que la consola no muestre ningún problema en la página Problemas. Para obtener más información, consulte Rol vinculado al servicio para Security Lake en la guía del usuario de Amazon Security Lake.
- Comparta los orígenes de datos de Security Lake. Cuando acceda OpenSearch desde la misma cuenta que Security Lake, asegúrese de que no aparezca ningún mensaje para registrar

sus depósitos de Security Lake con Lake Formation en la consola de Security Lake. Para el OpenSearch acceso entre cuentas, configure un suscriptor de consultas de Lake Formation en la consola de Security Lake. Utilice la cuenta asociada a su OpenSearch recurso como suscriptor. Para obtener más información, consulte Administración de suscriptores en Security Lake en la guía del usuario de Amazon Security Lake.

Además, también debe tener los siguientes recursos en su Cuenta de AWS:

- (Opcional) Un rol de IAM creado manualmente. Puede usar este rol para administrar el acceso al origen de datos. Como alternativa, puede hacer que OpenSearch Service cree un rol para usted automáticamente con los permisos necesarios. Si decide utilizar un rol de IAM creado manualmente, siga las instrucciones que se indican en the section called “Permisos necesarios para los roles de IAM creados manualmente”.

Procedimiento

Puede configurar un origen de datos para que se conecte a una base de datos de Security Lake desde la Consola de administración de AWS.

Para configurar una fuente de datos mediante el Consola de administración de AWS

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación de la izquierda, vaya a Administración central y seleccione Orígenes de datos conectados.
3. Elija Conectar.
4. Elija Security Lake como tipo de origen de datos.
5. Elija Siguiente.
6. En Detalles de la conexión de datos, escriba un nombre y una descripción opcional.
7. En Configuración del permiso de acceso de IAM, elija cómo administrar el acceso a su origen de datos.

    a. Si desea crear automáticamente un rol para este origen de datos, siga estos pasos:

        i. Seleccione Crear un nuevo rol.

        ii. Ingrese un nombre para el rol de IAM.

        iii. Seleccione una o más AWS Glue tablas para definir qué datos se pueden consultar.

b. Si desea utilizar un rol existente que usted administra, siga estos pasos:

i. Seleccione Usar un rol existente.

ii. Seleccione un rol existente en el menú desplegable.

> **Note**
>
> Cuando utilice su propio rol, debe asegurarse de que tiene todos los permisos requeridos. Para ello, adjunte las políticas necesarias desde la consola de IAM. Para obtener más información, consulte the section called “Permisos necesarios para los roles de IAM creados manualmente”.

8. (Opcional) En Etiquetas, agregue etiquetas al origen de datos.
9. Elija Siguiente.
10. En Configurar OpenSearch, elija cómo desea OpenSearch configurarlo.

- Revise los nombres de los recursos y la configuración de retención de datos predeterminados.

  Si utilizas la configuración predeterminada, se crea una nueva OpenSearch aplicación y un espacio de trabajo de Essentials sin coste adicional. OpenSearch le permite analizar varias fuentes de datos. Incluye espacios de trabajo, que proporcionan experiencias personalizadas para casos de uso populares. Los espacios de trabajo admiten el control de acceso, lo que le permite crear espacios privados para sus casos de uso y compartirlos solo con sus colaboradores.

11. Use ajustes personalizados:

a. Elija Personalizar.

b. Edite el nombre de la colección y la configuración de retención de datos según sea necesario.

c. Seleccione la OpenSearch aplicación y el espacio de trabajo que desee utilizar.

12. Elija Siguiente.
13. Revise sus opciones y seleccione Editar si necesita hacer algún cambio.
14. Seleccione Conector para configurar el origen de datos. Permanezca en esta página mientras se crea el origen de datos. Cuando esté listo, accederá a la página de detalles del origen de datos.

Siguientes pasos

Visite los OpenSearch paneles y cree un panel

Después de crear una fuente de datos, OpenSearch Service le proporciona una URL de OpenSearch Dashboards. Puede utilizar esto para consultar los datos mediante SQL o PPL. La integración de Security Lake incluye plantillas de consulta preempaquetadas para SQL y PPL para que pueda empezar a analizar sus registros.

Para obtener más información, consulte the section called “Configuración de un origen de datos de Security Lake”.

Recursos adicionales

Permisos necesarios para los roles de IAM creados manualmente

Al crear un origen de datos, se elige un rol de IAM para administrar el acceso a los datos. Tiene dos opciones:

1. Crear un nuevo rol de IAM automáticamente
2. Utilizar un rol de IAM creado manualmente

Si utiliza un rol creado manualmente, debe asociar los permisos correctos al rol. Los permisos deben permitir el acceso a la fuente de datos específica y permitir que el OpenSearch Servicio asuma la función para que el OpenSearch Servicio pueda acceder a sus datos e interactuar con ellos de forma segura. Además, conceda LakeFormation permisos al rol para cualquier base de datos y tablas que desee consultar. Otorgue `DESCRIBE` permisos al rol en SecurityLake las bases de datos que desee consultar desde la conexión de consulta directa. Otorgue al menos los permisos `SELECT and DESCRIBE` al rol de origen de datos para las tablas de la base de datos.

El siguiente ejemplo de política muestra los permisos con privilegios mínimos necesarios para crear y administrar un origen de datos. Si tiene permisos más generales, como la política `AdminstratorAccess`, estos engloban los permisos con privilegios mínimos de la política de ejemplo.

En el siguiente ejemplo de política, *`placeholder text`* sustitúyalo por su propia información.

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": [
                "*"
            ]
        }
    ]
```

```
}
```

El rol debe tener la siguiente política de confianza, que especifica el ID de destino.

JSON

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

Para obtener instrucciones sobre cómo crear los roles, consulte Creación de un rol mediante políticas de confianza personalizadas.

De forma predeterminada, el rol solo tiene acceso a los índices de orígenes de datos de consulta directa. Si bien puede configurar el rol para limitar o conceder el acceso a su origen de datos, se recomienda no ajustar el acceso de este rol. Si elimina el origen de datos, se eliminará este rol. Esto eliminará el acceso de los demás usuarios si están asignados al rol.

Consulta de datos de Security Lake que están cifrados con una clave administrada por el cliente

Si el depósito de Security Lake asociado a la conexión de datos se cifra mediante cifrado del lado del servidor y gestionado por un cliente AWS KMS key, debe añadir la función de LakeFormation servicio a la política de claves. Esto permite que el servicio acceda a los datos de sus consultas y los lea.

En el siguiente ejemplo de política, sustitúyalo *placeholder text* por su propia información.

```
{
    "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": "*"
}
```

Configuración y consulta de una fuente de datos de Security Lake en OpenSearch los paneles

Ahora que ha creado su fuente de datos, puede configurarla en OpenSearch los paneles.

En esta sección, se explican varios casos de uso de la fuente de datos en los OpenSearch paneles de control antes de realizar consultas en los datos. Para empezar, debe navegar hasta su fuente de datos en los OpenSearch paneles. En el menú de navegación izquierdo, en Administración, elija Origen de datos. A continuación, seleccione el nombre de la fuente de datos que creó anteriormente en la consola de OpenSearch servicio.

Consulte las tablas de Security Lake de Discover

Si ha creado tablas basadas en sus registros de Security Lake, ahora puede consultarlas directamente desde OpenSearch Discover. Esto le permite acceder y analizar sin problemas los datos almacenados en Security Lake, directamente desde la conocida interfaz Discover. Al consultar Security Lake directamente desde Discover, puede evitar la necesidad de extraer, transformar y cargar manualmente los datos en un índice de búsqueda independiente. Para empezar rápidamente a analizar sus registros, Discover incluye un conjunto de consultas guardadas de PPL y SQL.

Comience por seleccionar el origen de datos que ha configurado. Seleccione la base de datos y la tabla asociadas que desee consultar y, a continuación, utilice la barra de búsqueda para escribir consultas en las tablas. Para saber qué sentencias, comandos y limitaciones son compatibles con la integración de Security Lake, consulte the section called “Comandos SQL y PPL compatibles”.

Para aprovechar las consultas prediseñadas que están disponibles para Security Lake, vaya a ... en la parte superior derecha de Discover, seleccione Abrir consulta y, a continuación, seleccione

Plantillas. Hay muchas consultas prediseñadas disponibles para los orígenes de registros compatibles con Security Lake. Busque las plantillas que coincidan con su caso de uso, copie la consulta para utilizarla en la barra de búsqueda y sustituya los campos con plantillas (como región y acción) por su propia información.

Acelere los datos de Discover

Para mejorar el rendimiento y agilizar las consultas y los análisis posteriores OpenSearch, puede incorporar los resultados de la consulta de Discover en una vista OpenSearch indexada.

Cómo crear una vista indexada

1. En Discover, elija Crear vista indexada.
2. En el editor de consultas, escriba la consulta que prefiera. Puede crear una nueva consulta aquí o utilizar una existente de sus búsquedas anteriores.
3. Especifique un nombre para la nueva vista indexada. Elija un nombre descriptivo que lo ayude a identificar la vista más tarde.
4. Configure los ajustes de retención de datos para la vista indexada. Puede especificar durante cuánto tiempo deben mantenerse los datos en el índice, lo que le permite equilibrar el rendimiento con los costos de almacenamiento.
5. Cree la vista indexada. Una vez creada, la vista indexada estará disponible para realizar consultas y análisis más rápidos.

Si ya ha creado vistas indexadas, puede acceder a ellas desde Discover.

Cómo usar una vista indexada existente

1. En Discover, elija Seleccionar vista indexada para ver una lista de las vistas indexadas existentes de Security Lake.
2. Elija la vista indexada que desea usar. Esto aplicará la vista a la consulta actual, lo que podría acelerar considerablemente la recuperación y el análisis de los datos.

Crear una vista de panel para el origen de datos

Al utilizar OpenSearch Service, puede analizar los tipos de AWS registro más populares mediante plantillas de panel prediseñadas. Para Security Lake, hay plantillas para registros de VPC y WAF. CloudTrail Estas plantillas le permiten crear un panel de control adaptado a sus datos específicos.

Incluyen consultas y paneles prediseñados que están diseñados para ese tipo de registro específico. Esto le permite empezar a analizar rápidamente estas populares fuentes de AWS registro, sin tener que compilarlo todo desde cero.

> **Note**
>
> Los Dashboards utilizan vistas indexadas, que ingieren datos de Security Lake y contribuyen a la computación directa de consultas y colecciones.

Siga estos pasos para crear un panel con una de estas plantillas prediseñadas, de modo que pueda empezar a explorar y analizar sus datos de inmediato.

Cómo crear un panel

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Administración central y, a continuación, Orígenes de datos conectados.
3. Seleccione el origen de datos para abrir la página de detalles.
4. Elija Crear panel.
5. Elija el tipo de panel que desea crear.
6. Introduzca un nombre para su panel de control.
7. Escriba una descripción opcional para el panel de control.
8. Selecciona una o más tablas de AWS Glue para verlas en tu panel de control.
9. Elija la frecuencia con la que quiere actualizar los datos del panel.
10. Elige qué OpenSearch espacio de trabajo quieres usar.
    a. Seleccione Crear un espacio de trabajo para crear un espacio de trabajo nuevo.
    b. Para usar un espacio de trabajo existente, seleccione Seleccionar espacio de trabajo existente.
11. Escriba un nombre para su espacio de trabajo.
12. Elija Crear panel.

Resolución de problemas

Puede haber casos en los que los resultados no devuelvan los resultados esperados. Si tiene algún problema, asegúrese de seguir las the section called “Recomendaciones”.

Administrar una fuente de datos en Amazon OpenSearch Service

La gestión de la fuente de datos es una parte importante del mantenimiento de la fiabilidad, la disponibilidad y el rendimiento de las fuentes de datos de consultas directas y del resto de sus AWS soluciones. AWS proporciona las siguientes herramientas para supervisar, informar cuando algo va mal y tomar medidas automáticas cuando sea necesario.

Temas

- Supervisión con CloudWatch métricas y fuentes de datos
- Activación y desactivación de orígenes de datos
- Supervisión con AWS presupuesto
- Eliminación de un origen de datos

Supervisión con CloudWatch métricas y fuentes de datos

Puede monitorear la consulta directa usando CloudWatch. CloudWatch recopila datos sin procesar y los procesa para convertirlos en métricas legibles y prácticamente en tiempo real. Estas estadísticas se mantienen durante 15 meses, de forma que pueda obtener acceso a información histórica y disponer de una mejor perspectiva sobre el rendimiento de su aplicación web o servicio.

También puede establecer alarmas para supervisar determinados umbrales y enviar notificaciones o llevar a cabo acciones cuando se cumplan dichos umbrales. Para obtener más información, consulta Qué es Amazon CloudWatch.

Amazon S3 notifica las siguientes métricas:

Métrica	Description (Descripción)
AsyncQueryCreateAPI	El número total de solicitudes hechas a la API para crear consultas asíncronas. Estadísticas pertinentes: promedio, máximo, suma

Métrica	Description (Descripción)
	Dimensiones: `ClientId`, `DomainName` Frecuencia: 60 segundos
AsyncQueryGetApiRequestCount	El número total de solicitudes hechas a la API para recuperar resultados de consultas asíncronas. Estadísticas pertinentes: promedio, máximo, suma Dimensiones: `ClientId`, `DomainName` Frecuencia: 60 segundos
AsyncQueryCancelApiRequestCount	El número total de solicitudes hechas a la API para cancelar consultas asíncronas. Estadísticas pertinentes: promedio, máximo, suma Dimensiones: `ClientId`, `DomainName` Frecuencia: 60 segundos
AsyncQueryGetApiFailedRequestCusErrCount	El número de solicitudes fallidas al recuperar los resultados de una consulta asíncrona debido a errores relacionados con el cliente (por ejemplo, un ID de consulta no válido). Estadísticas pertinentes: promedio, máximo, suma Dimensiones: `ClientId`, `DomainName` Frecuencia: 60 segundos

Métrica	Description (Descripción)
AsyncQueryCancelApiFailedRequestCusErrCount	El número de solicitudes fallidas al recuperar los resultados de una consulta asíncrona debido a errores relacionados con el cliente (por ejemplo, un ID de consulta no válido). Estadísticas pertinentes: promedio, máximo, suma Dimensiones: `ClientId`, `DomainName` Frecuencia: 60 segundos
AsyncQueryCancelApiFailedRequestSysErrCount	El número de solicitudes fallidas al crear consultas asíncronas debido a errores relacionados con el cliente. Estadísticas pertinentes: promedio, máximo, suma Dimensiones: `ClientId`, `DomainName` Frecuencia: 60 segundos
AsyncQueryGetApiFailedRequestSysErrCount	El número de solicitudes fallidas al recuperar los resultados de una consulta asíncrona debido a errores relacionados con el sistema. Estadísticas pertinentes: promedio, máximo, suma Dimensiones: `ClientId`, `DomainName` Frecuencia: 60 segundos

CloudWatch Logs y Security Lake informan de las siguientes métricas:

Métrica	Description (Descripción)
DirectQueryRate	La tasa de solicitudes realizadas en relación con los orígenes de datos. Estadísticas pertinentes: mínimo, máximo, promedio, suma

Métrica	Description (Descripción)
	Dimensiones: `DataSourceName` Frecuencia: 60 segundos
DirectQueryLatency	La latencia observada al ejecutar consultas en los orígenes de datos. Estadísticas pertinentes: mínimo, máximo, promedio, suma, P90, P99 Dimensiones: `DataSourceName` Frecuencia: 60 segundos
FailedDirectQueries	El número total de errores de consulta que se observan en las consultas del origen de datos. Estadísticas pertinentes: mínimo, máximo, promedio, suma Dimensiones: `DataSourceName` Frecuencia: 60 segundos
DirectQueryConsumedOCU	La cantidad OCUs que se consume para ejecutar las consultas en las fuentes de datos. Estadísticas pertinentes: mínimo, máximo, promedio, suma, P90, P99 Dimensiones: `DataSourceName` Frecuencia: 60 segundos

Activación y desactivación de orígenes de datos

> **Note**
>
> La siguiente información solo se aplica a los orígenes de datos de Amazon S3.

Si desea detener el uso de consultas directas para un origen de datos, puede optar por deshabilitarlo. El deshabilitar un origen de datos, finalizará la ejecución de las consultas existentes e impedirá que el usuario ejecute todas las consultas nuevas.

Las aceleraciones configuradas para mejorar el rendimiento de las consultas, como los índices de omisión, las vistas materializadas o el índice de cobertura, pasarán a ser manuales una vez que se desactive el origen de datos. Una vez que un origen de datos se establece como activo después de deshabilitarlo, las consultas de los usuarios se ejecutarán según lo previsto. Las aceleraciones que anteriormente estaban configuradas como manuales deberán configurarse manualmente para volver a ejecutarse según una programación.

Supervisión con AWS presupuesto

Amazon OpenSearch Service está rellenando los datos de uso de la OCU a nivel de cuenta en el Cost Explorer de Billing and Cost Management. Los clientes pueden contabilizar el uso de las OCU por cuenta y establecer umbrales y alertas cuando estos se superen.

El formato del tipo de uso que se va a filtrar en Cost Explorer es el siguiente RegionCode: DirectQuery OCU (OCU-Hours). Si quieres recibir una notificación cuando el uso de DirectQuery OCU (ocu-horas) alcance tu límite, puedes crear una cuenta de AWS Budgets y configurar una alerta en función del umbral que establezcas. De forma opcional, en Amazon S3 se puede optar por configurar un tema de Amazon SNS, que desactivará un origen de datos en caso de que se cumpla un criterio límite.

> **Note**
>
> Los datos de uso de AWS Budgets no están disponibles en tiempo real y pueden retrasarse hasta 8 horas.

Eliminación de un origen de datos

Cuando eliminas una fuente de datos, Amazon OpenSearch Service la elimina de tu dominio o colección. OpenSearch El servicio también elimina los índices asociados a la fuente de datos. Los datos transaccionales no se eliminan del otro Servicio de AWS, pero el otro Servicio de AWS no envía nuevos datos al OpenSearch Servicio.

Puede eliminar la integración de una fuente de datos mediante la API Consola de administración de AWS o la API del OpenSearch servicio.

Consola de administración de AWS

Para eliminar un origen de datos de Amazon S3

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Dominios.
3. Seleccione el dominio del que desea eliminar un origen de datos. Se abrirá la página de detalles del dominio. Elija la pestaña Conexiones que se encuentra debajo de la información general, y busque la sección Consulta directa.
4. Seleccione el origen de datos que desea eliminar, elija Eliminar y confirme la eliminación.

Para eliminar una fuente de datos de CloudWatch Logs o Security Lake

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Administración central y, a continuación, Orígenes de datos conectados.
3. Seleccione el origen de datos que desea eliminar, elija Eliminar y confirme la eliminación.

OpenSearch API de servicio

Para eliminar una fuente de datos de Amazon S3, utilice la operación de DeleteDataSourceAPI.

```
POST https://es.region.amazonaws.com/2021-01-01/opensearch/domain/domain-name/
dataSource/data-source-name
```

Para eliminar una fuente de datos de CloudWatch Logs o Security Lake, utilice la operación de DeleteDirectQueryDataSourceAPI.

Optimización del rendimiento de las consultas para las fuentes OpenSearch de datos de Amazon Service

El rendimiento de las consultas en Amazon OpenSearch Service puede ralentizarse cuando accedes a fuentes de datos externas. Esto puede deberse a factores como la latencia de la red, la transformación de datos o los grandes volúmenes de datos. Para mejorar el rendimiento, considere la posibilidad de indexar determinadas cantidades de datos según el caso de uso:

- Cómo acelerar las consultas directas en Amazon S3 (índice de saltos)
- Creación de visualizaciones de paneles en Security Lake (vistas materializadas)
- Ingesta de los resultados de las consultas mediante vistas indexadas para revisarlos sin conexión o mejorar el rendimiento en Security Lake (vistas materializadas)

Para obtener la documentación completa sobre las consultas aceleradas, incluidas las consultas de ejemplo, consulte Optimizar el rendimiento de las consultas mediante la OpenSearch indexación en la documentación de código abierto.

Temas

- Índices de omisión
- Vistas materializadas
- Índices de cobertura

Índices de omisión

Un índice de omisión ingiere solo los metadatos de los datos almacenados en Amazon S3. Cuando consulta una tabla con un índice de omisión, el planificador de consultas hace referencia al índice y vuelve a escribir la consulta para localizar los datos de manera eficiente, en lugar de escanear todas las particiones y archivos. Este enfoque ayuda a delimitar la ubicación exacta de los datos almacenados.

Hay dos formas de crear un índice de omisión. La primera forma consiste en generar automáticamente el índice de omisiones a partir de los detalles del origen de datos. La segunda consiste en utilizar Query Workbench para crear manualmente el índice de saltos mediante una sentencia SQL.

Para generar automáticamente un índice de omisión a partir del origen de datos, vaya a Administración de paneles y a Acelerar datos y, a continuación, seleccione la base de datos y la tabla (puede que tenga que actualizarlas para obtener las bases de datos y tablas más recientes). A continuación, puede elegir Generar para generar automáticamente un índice de omisión o seleccionar manualmente cada campo que desee indexar y especificar la aceleración (tipo de índice de omisión). Por último, elija Crear aceleración para crear una tarea recurrente que rellene el nuevo índice de saltos.

Los índices de omisión solo se admite para orígenes de datos de Amazon S3.

Para obtener más información sobre cómo configurar los índices de omisión mediante Query Workbench, consulte Omitir índices en la documentación. OpenSearch

Vistas materializadas

Las vistas materializadas utilizan consultas complejas, como las agregaciones, para admitir las visualizaciones de los paneles. OpenSearch Ingieren un subconjunto de sus datos en función de la consulta y lo almacenan en un índice. OpenSearch Puede utilizar este índice para crear visualizaciones.

Las vistas materializadas son compatibles con los orígenes de datos de Amazon S3 y Security Lake.

Para obtener más información sobre la configuración de vistas materializadas mediante Query Workbench, consulte Vistas materializadas en la documentación. OpenSearch

Índices de cobertura

Un índice de cobertura ingiere datos de una columna específica de una tabla y OpenSearch crea un nuevo índice basado en estos datos. Puede usar este nuevo índice para visualizaciones y otras OpenSearch funciones, como la detección de anomalías o el análisis geoespacial.

Los índices de cobertura solo se admiten para orígenes de datos de Amazon S3.

Para obtener más información sobre cómo configurar los índices de cobertura, consulte Cubrir índices en la documentación. OpenSearch

Comandos SQL y PPL compatibles

OpenSearch SQL y OpenSearch Pipeline Processing Language (PPL) son lenguajes para consultar, analizar y procesar datos en CloudWatch Logs Insights y OpenSearch Security Lake. Puede usar

OpenSearch SQL y OpenSearch PPL en OpenSearch Discover para consultar datos en CloudWatch Logs, Amazon S3 o Security Lake. CloudWatch Logs Insights también es compatible con los lenguajes de consulta OpenSearch PPL y OpenSearch SQL, además de Logs Insights QL, un lenguaje de consultas diseñado específicamente para analizar los registros. CloudWatch

- OpenSearch SQL: OpenSearch SQL ofrece una opción conocida si está acostumbrado a trabajar con bases de datos relacionales. OpenSearch SQL ofrece un subconjunto de funciones de SQL, lo que lo convierte en una buena opción para realizar consultas ad hoc y tareas de análisis de datos. Con OpenSearch SQL, puede usar comandos como SELECT, FROM, WHERE, GROUP BY, HAVING y varios otros comandos y funciones de SQL disponibles en SQL. Puede ejecutar JOINs en tablas (o grupos de registros), correlacionar datos entre tablas (o grupos de registros) mediante subconsultas y utilizar el amplio conjunto de funciones JSON, matemáticas, de cadena, condicionales y otras funciones de SQL para realizar análisis eficaces de los datos de registro y seguridad.
- OpenSearch PPL (lenguaje de procesamiento canalizado): con OpenSearch PPL, puede recuperar, consultar y analizar datos mediante comandos agrupados, lo que facilita la comprensión y la redacción de consultas complejas. Su sintaxis se basa en las canalizaciones de Unix y permite encadenar comandos para transformar y procesar datos. Con el PPL, puede filtrar y agregar datos y utilizar comandos como JOINs subconsultas, LOOKUP y un amplio conjunto de funciones matemáticas, de cadenas, de fecha, condicionales y de otro tipo para el análisis.

Si bien la mayoría de los comandos de los lenguajes de consulta OpenSearch PPL y OpenSearch SQL son comunes en todos CloudWatch los registros OpenSearch, existen algunas diferencias en cuanto al conjunto de comandos y funciones que admite cada uno de estos servicios. Consulte la siguiente tabla para obtener más detalles.

- the section called “Comandos SQL compatibles”
  - the section called “CloudWatch Registra Insights”
  - the section called “Restricciones SQL generales”
- the section called “Comandos compatibles PPL”
  - the section called “Información adicional para CloudWatch los usuarios de Logs Insights que utilizan OpenSearch PPL”

Comandos y funciones de OpenSearch SQL compatibles

En las siguientes tablas de referencia, se muestran los comandos de SQL compatibles con OpenSearch Discover para consultar datos en Amazon S3, Security Lake o CloudWatch Logs, y los comandos de SQL que se admiten en CloudWatch Logs Insights. La sintaxis SQL admitida en CloudWatch Logs Insights y la admitida en OpenSearch Discover para consultar CloudWatch los registros son las mismas y se hace referencia a ellas como CloudWatch registros en las siguientes tablas.

> **Note**
>
> OpenSearch también es compatible con SQL para consultar los datos que se ingieren OpenSearch y almacenan en índices. Este dialecto de SQL es diferente del SQL utilizado en las consultas directas y se denomina SQL en los índices. OpenSearch

Temas

- Comandos
- Funciones
- Restricciones SQL generales
- Información adicional para los usuarios de CloudWatch Logs Insights que utilizan SQL OpenSearch

Comandos

> **Note**
>
> En la columna de comandos del ejemplo, reemplace *<tableName/logGroup>* según sea necesario en función del origen de datos que esté consultando.
>
> - Comando de ejemplo: `SELECT Body , Operation FROM` **`<tableName/logGroup>`**
> - Si está consultando en Amazon S3 o Security Lake, utilice: `SELECT Body , Operation FROM` **`table_name`**
> - Si está consultando los CloudWatch registros, utilice: `SELECT Body , Operation FROM` **`` `LogGroupA` ``**

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Cláusula SELECT”	Muestra los valores proyecta os.	compatib e	compatib e	compatib e	`SELECT method, status FROM <tableName/logGroup>`
the section called “Cláusula WHERE”	Filtra los eventos del registro en función de los criterios de campo proporcio nados.	compatib e	compatib e	compatib e	`SELECT * FROM <tableName/logGroup> WHERE status = 100`
the section called “Cláusula GROUP BY”	Los grupos registran los eventos según la categorí y encuentr n el promedio	compatib e	compatib e	compatib e	`SELECT method, status, COUNT(*) AS request_count, SUM(bytes) AS total_bytes FROM <tableName/logGroup> GROUP BY method, status`

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
	en función de las estadísti cas.				

the section called “Cláusula HAVING”	Filtra los resultado s en función de las condicio es de agrupaci n.	compatib e	compatib e	compatib e	`SELECT`
					`  method,`
					`  status,`
					`  COUNT(*) AS request_count,`
					`  SUM(bytes) AS total_bytes`
					`FROM`
					`  <tableName/logGroup>`
					`GROUP BY`
					`  method,`
					`  status`
					`HAVING`
					`  COUNT(*) > 5`

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Cláusula ORDER BY”	Ordena los resultado s en función de los campos de la cláusula de orden. Puede especific ar un orden tanto ascende e como descend te.	compatib e	compatib e	compatib e	`SELECT * FROM <tableName/logGroup> ORDER BY status DESC`

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Cláusula JOIN” (INNER | CROSS | LEFT OUTER)	Une los resultado s de dos tablas en función de los campos en común.	compatib e (se deben usar Inner y Left Outer como palabras clave para unir; solo se admite una operació JOIN en una sentenci SELECT	compatib e (se deben usar las palabras clave Inner, Left Outer y Cross para unir)	compatib e (se deben usar las palabras clave Inner, Left Outer y Cross para unir)	SELECT A.Body, B.Timestamp FROM <tableNameA/logGroupA> AS A INNER JOIN <tableNameB/logGroupB> AS B ON A.`requestId` = B.`requestId`
the section called “Cláusula LIMIT”	Restring los resultado s a las N primeras filas.	compatib e	compatib e	compatib e	SELECT * FROM <tableName/logGroup> LIMIT 10

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Cláusula CASE”	Evalúa las condicion es y devuelve un valor cuando se cumple la primera condició .	compatib e	compatib e	compatib e	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	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:

the section called “Expresio nes de tabla comunes”	Crea un conjunto de resultado s temporal con nombre en una sentenci SELECT INSERT, UPDATE DELETE o MERGE.	no compatib e	compatib e	compatib e	`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`

the section called “EXPLAIN	Muestra el plan de ejecució de una sentenci SQL sin ejecutarl a.	no compatib e	compatib e	compatib e	`EXPLAIN`
					`SELECT`
					`  k,`
					`  SUM(v)`
					`FROM`
					`  VALUES`
					`  (1, 2),`
					`  (1, 3) AS t(k, v)`
					`GROUP BY`
					`  k`

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Cláusula LATERAL SUBQUEF	Permite que una subcons ta de la cláusula FROM haga referenci a a las columna de los element anteriore s de la misma cláusula FROM.	no compatib e	compatib e	compatib e	`SELECT * FROM tableName LATERAL ( SELECT * FROM t2 WHERE t1.c1 = t2.c1 )`

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Cláusula LATERAL VIEW”	Genera una tabla virtual mediante la aplicació n de una función generad a de tablas a cada fila de una tabla base.	no compatib e	compatib e	compatib e	`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 “Predicad o LIKE”	Hace coincidir una cadena con un patrón mediante caracter es comodín	compatib e	compatib e	compatib e	`SELECT method, status, request, host FROM <tableName/logGroup> WHERE method LIKE 'D%'`

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “OFFSET”	Especific a el número de filas que se van a omitir antes de empezar a devolver las filas de la consulta	compatib e cuando se utiliza junto con una cláusula LIMIT en una consulta Por ejemplo: • Compa e: SELEC * FROM Table LIMIT 100 OFFSE 10 • No compa e: SELEC *	compatib e	compatib e	SELECT method, status, bytes, datetime FROM <tableName/logGroup> ORDER BY datetime OFFSET 10

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
		FROM Table OFFSE 10			
the section called “Cláusula PIVOT”	Transfor a las filas en columna rotando los datos de un formato basado en filas a un formato basado en columna	no compatib e	compatib e	compatib e	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 (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Operador es de estableci miento”	Combina los resultado s de dos o más sentenci s SELECT (p. ej., UNION, INTERSI , EXCEPT	compatib e	compatib e	compatib e	`SELECT method, status, bytes FROM <tableName/logGroup> WHERE status = '416' UNION SELECT method, status, bytes FROM <tableName/logGroup> WHERE bytes > 20000`
the section called “Cláusula SORT BY”	Especific a el orden en el que se deben devolver los resultado s de la consulta	compatib e	compatib e	compatib e	`SELECT method, status, bytes FROM <tableName/logGroup> SORT BY bytes DESC`

Comando	Descripti on (Descrip ión)	CloudWa h Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called "UNPIVOT	Transfor a las columna en filas, rotando los datos de un formato basado en columna a un formato basado en filas.	no compatib e	compatib e	compatib e	`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`

Funciones

> **Note**
>
> En la columna de comandos del ejemplo, reemplace *<tableName/logGroup>* según sea necesario en función del origen de datos que esté consultando.
>
> - Comando de ejemplo: `SELECT Body , Operation FROM` **`<tableName/logGroup>`**
> - Si está consultando en Amazon S3 o Security Lake, utilice: `SELECT Body , Operation FROM` **`table_name`**

- Si estás consultando CloudWatch registros, usa: `SELECT Body , Operation FROM` **`` `LogGroupA` ``**

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s de cadena	Funciones integradas que pueden manipular y transform ar cadenas y datos de texto en consultas SQL. Por ejemplo, convertir mayúsculas y minúscula s, combinar cadenas, extraer partes y limpiar el texto.	compa e	compa e	compa e	`SELECT UPPER(method) AS upper_method, LOWER(host) AS lower_hos t FROM <tableName/logGroup>`
the section called “Funcic s de fecha	Funciones integrada s para administrar y transform ar los datos	compa e	compa e	compa e	`SELECT TO_TIMESTAMP(datetime) AS timestamp, TIMESTAMP_SECONDS( UNIX_TIMESTAMP(datetime)) AS from_seconds,`

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
y hora”	de fecha y hora en las consultas. Por ejemplo, date_add, date_form at, datediff y current_d ate.				UNIX_TIMESTAMP(datetime) AS to_unix, FROM_UTC_TIMESTAMP (datetime, 'PST') AS to_pst, TO_UTC_TIMESTAMP(d atetime, 'EST') AS from_est FROM <tableName/logGroup>
the section called “Funcic s de agrega n”	Funciones integradas que realizan cálculos en varias filas para generar un único valor resumido. Por ejemplo, suma, recuento, promedio, máximo y mínimo.	compa e	compa e	compa e	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>

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s condici ales”	Funciones integradas que realizan acciones en función de condicion es específic as o que evalúan las expresion es de forma condicional. Por ejemplo, CASE e IF.	compa e	compa e	compa e	`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>`

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s JSON”	Funciones integrada s para analizar, extraer, modificar y consultar datos con formato JSON en consultas SQL (por ejemplo, from_json , to_json, get_json_ object, json_tupl e) que permiten manipular las estructur as JSON en los conjuntos de datos.	compa e	compa e	compa e	`SELECT FROM_JSON( @message, 'STRUCT< host: STRING, user-identifier: STRING, datetime: STRING, method: STRING, status: INT, bytes: INT >' ) AS parsed_json FROM <tableName/logGroup>`

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s de matriz”	Funciones integrada s para trabajar con columnas de tipo matriz en consultas SQL, lo que permite realizar operacion es como acceder, modificar y analizar los datos de la matriz (p. ej., size, explode, array_con tains).	compa e	compa e	compa e	`SELECT scores, size(scores) AS length, array_contains(scores, 90) AS has_90 FROM <tableName/logGroup>`

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s de ventan	Funciones integradas que realizan cálculos en un conjunto específic o de filas relacionadas con la fila (ventana) actual, lo que permite realizar operacion es como la clasificación, los totales acumulado s y los promedios móviles (por ejemplo, ROW_NUMBE R, RANK, LAG, LEAD).	compa e	compa e	compa e	`SELECT field1, field2, RANK() OVER (ORDER BY field2 DESC) AS field2Rank FROM <tableName/logGroup>`

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s de conver n”	Funciones integrada s para convertir datos de un tipo a otro en las consultas SQL, lo que permite la transform ación de los tipos de datos y las conversio nes de formato (por ejemplo, CAST, TO_DATE, TO_TIMEST AMP, BINARY)	compa e	compa e	compa e	`SELECT CAST('123' AS INT) AS converted_number, CAST(123 AS STRING) AS converted_string FROM <tableName/logGroup>`

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s de predica s”	Funciones integrada s que evalúan las condiciones y devuelven valores booleanos (verdader o/falso) en función de criterios o patrones específic os (por ejemplo, IN, LIKE, BETWEEN, IS NULL, EXISTS)	compa e	compa e	compa e	`SELECT * FROM <tableName/logGroup> WHERE id BETWEEN 50000 AND 75000`

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s de mapa”	Aplica una función específic a a cada elemento de una colección y transform a los datos a un nuevo conjunto de valores.	no compat e	compat e	compat e	`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 “Funcic s matem as”	Realiza operaciones matemátic as con datos numéricos , como calcular promedios , sumas o valores trigonomé tricos.	compat e	compat e	compat e	`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>`

Gramá SQL disponi e	Description (Descripc ión)	CloudV h Registr	Amazo S3	Securit Lake	Comando de ejemplo:
the section called “Funcic s de grupos de registro múltiple ”	Permite a los usuarios especific ar varios grupos de registros en una sentencia SQL SELECT	compat e	No aplicab	No aplicab	`SELECT lg1.Column1, lg1.Column2 FROM `logGroups(logGrou pIdentifier: ['LogGroup1', 'LogGroup2'])` AS lg1 WHERE lg1.Column3 = "Success"`
the section called “Funcic s genera as”	Crea un objeto iterador que produce una secuencia de valores, lo que permite un uso eficiente de la memoria en conjuntos de datos de gran tamaño.	no compat e	compat e	compat e	`SELECT explode(array(10, 20))`

Restricciones SQL generales

Se aplican las siguientes restricciones al usar OpenSearch SQL with CloudWatch Logs, Amazon S3 y Security Lake.

1. Solo puede usar una operación JOIN en una sentencia SELECT.
2. Solo se admite un nivel de subconsultas anidadas.
3. No se admiten consultas de varias sentencias separadas por punto y coma.
4. No se admiten consultas que contengan nombres de campo idénticos pero que solo difieran en mayúsculas y minúsculas (como field1 y FIELD1).

   Por ejemplo, no se admiten las siguientes consultas:

   ```
   Select AWSAccountId, awsaccountid from LogGroup
   ```

   Sin embargo, la siguiente consulta se debe a que el nombre del campo (@logStream) es idéntico en ambos grupos de registros:

   ```
   Select a.`@logStream`, b.`@logStream` from Table A INNER Join Table B on a.id = b.id
   ```

5. Las funciones y expresiones deben funcionar con los nombres de los campos y formar parte de una sentencia SELECT con un grupo de registros especificado en la cláusula FROM.

   Por ejemplo, no se admite esta consulta:

   ```
   SELECT cos(10) FROM LogGroup
   ```

   Se admite esta consulta:

   ```
   SELECT cos(field1) FROM LogGroup
   ```

Información adicional para los usuarios de CloudWatch Logs Insights que utilizan SQL OpenSearch

CloudWatch Logs admite consultas OpenSearch SQL en la consola, la API y la CLI de Logs Insights. Es compatible con la mayoría de los comandos, incluidos SELECT, FROM, WHERE, GROUP BY, HAVING, JOINS y consultas anidadas, además de funciones JSON, matemáticas, de cadena y condicionales. Sin embargo, CloudWatch Logs solo admite operaciones de lectura, por lo que no permite sentencias DDL o DML. Consulte las tablas de las secciones anteriores para obtener una lista completa de los comandos y funciones compatibles.

Funciones de grupos de registros múltiples

CloudWatch Logs Insights admite la posibilidad de consultar varios grupos de registros. Para abordar este caso de uso en SQL, puede usar el comando `logGroups`. Este comando es específico para consultar datos en CloudWatch Logs Insights que involucran uno o más grupos de registros. Use esta sintaxis para consultar varios grupos de registros especificándolos en el comando, en lugar de escribir una consulta para cada uno de los grupos de registros y combinarlos con un comando `UNION`.

Sintaxis:

```
`logGroups(
    logGroupIdentifier: ['LogGroup1','LogGroup2', ...'LogGroupn']
)
```

En esta sintaxis, puede especificar hasta 50 grupos de registros en el parámetro `logGroupIndentifier`. Para hacer referencia a los grupos de registros de una cuenta de supervisión, utilícelo ARNs en lugar de `LogGroup` nombres.

Consulta de ejemplo:

```
SELECT LG1.Column1, LG1.Column2 from `logGroups(
    logGroupIdentifier: ['LogGroup1', 'LogGroup2']
)` as LG1
WHERE LG1.Column1 = 'ABC'
```

Al consultar los registros, no se admite la siguiente sintaxis que implica varios grupos de CloudWatch registros después de la `FROM` sentencia:

```
SELECT Column1, Column2 FROM 'LogGroup1', 'LogGroup2', ...'LogGroupn'
WHERE Column1 = 'ABC'
```

Restricciones

Cuando utilice comandos SQL o PPL, encierre determinados campos entre comillas invertidas para consultarlos. Las comillas invertidas son obligatorias para los campos con caracteres especiales (no alfabéticos ni numéricos). Por ejemplo, encierre `@message`, `Operation.Export,` y `Test::Field` entre comillas invertidas. No es necesario incluir las columnas con nombres exclusivamente alfabéticos entre comillas invertidas.

Ejemplo de consulta con campos sencillos:

```
SELECT SessionToken, Operation, StartTime  FROM `LogGroup-A`
LIMIT 1000;
```

La misma consulta con comillas invertidas agregadas:

```
SELECT `SessionToken`, `Operation`, `StartTime`  FROM `LogGroup-A`
LIMIT 1000;
```

Para ver otras restricciones generales que no son específicas de los CloudWatch registros, consultethe section called “Restricciones SQL generales”.

Consultas y cuotas de ejemplo

> **Note**
>
> Lo siguiente se aplica tanto a los usuarios de CloudWatch Logs Insights como a los OpenSearch usuarios que consultan CloudWatch datos.

Para ver ejemplos de consultas SQL que puede usar en CloudWatch Logs, consulte Consultas guardadas y de ejemplo en la consola de Amazon CloudWatch Logs Insights para ver ejemplos.

Para obtener información sobre los límites que se aplican al consultar CloudWatch los registros del OpenSearch servicio, consulte las cuotas de CloudWatch registros en la Guía del usuario de Amazon CloudWatch Logs. Los límites se refieren al número de grupos de CloudWatch registros que se pueden consultar, el número máximo de consultas simultáneas que se pueden ejecutar, el tiempo máximo de ejecución de las consultas y el número máximo de filas devueltas en los resultados. Los límites son los mismos independientemente del idioma que utilice para consultar los CloudWatch registros (es decir, OpenSearch PPL, SQL y Logs Insights).

Comandos SQL

Temas

- Funciones de cadena
- Funciones de fecha y hora
- Funciones de agregación

- Funciones condicionales
- Funciones JSON
- Funciones de matriz
- Funciones de ventana
- Funciones de conversión
- Funciones de predicados
- Funciones de mapa
- Funciones matemáticas
- Funciones generadoras
- Cláusula SELECT
- Cláusula WHERE
- Cláusula GROUP BY
- Cláusula HAVING
- Cláusula ORDER BY
- Cláusula JOIN
- Cláusula LIMIT
- Cláusula CASE
- Expresiones de tabla comunes
- EXPLAIN
- Cláusula LATERAL SUBQUERY
- Cláusula LATERAL VIEW
- Predicado LIKE
- OFFSET
- Cláusula PIVOT
- Operadores de establecimiento
- Cláusula SORT BY
- UNPIVOT

Funciones de cadena

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
ascii(str)	Devuelve el valor numérico del primer carácter de `str`.
base64(bin)	Convierte el argumento de binario `bin` a una cadena de base 64.
bit_length(expr)	Devuelve la longitud en bits de los datos de la cadena o el número de bits de los datos binarios.
btrim(str)	Elimina los caracteres de espacio inicial y final de `str`.
btrim(str, trimStr)	Elimine los caracteres `trimStr` iniciales y finales de `str`.
char(expr)	Devuelve el carácter ASCII que tiene el equivalente binario a `expr`. Si n es mayor que 256, el resultado equivale a chr (n % 256)
char_length(expr)	Devuelve la longitud en caracteres de la cadena de datos o el número de bytes de los datos binarios. La longitud de los datos de cadena incluye los espacios finales. La longitud de los datos binarios incluye los ceros binarios.
character_length(expr)	Devuelve la longitud en caracteres de la cadena de datos o el número de bytes de

Función	Description (Descripción)
	los datos binarios. La longitud de los datos de cadena incluye los espacios finales. La longitud de los datos binarios incluye los ceros binarios.
chr(expr)	Devuelve el carácter ASCII que tiene el equivalente binario a `expr`. Si n es mayor que 256, el resultado equivale a chr (n % 256)
concat_ws(sep[, str | array(str)]+)	Devuelve la concatenación de las cadenas separadas por `sep`, omitiendo los valores nulos.
contains(left, right)	Devuelve un BOOLEANO. El valor es verdadero si right se encuentra dentro de left. Devuelve NULL si alguna de las expresiones de entrada es NULL. De lo contrario, devuelve que el valor es falso. Tanto left como right deben ser de tipo STRING o BINARY.
decode(bin, charset)	Decodifica el primer argumento utilizando el juego de caracteres del segundo argumento.
decode(expr, search, result [, search, result] ... [, default])	Compara expr con cada valor de búsqueda en orden. Si expr es igual a un valor de búsqueda, decode devuelve el resultado correspondiente. Si no se encuentra coincidencia, devuelve el valor predeterminado. Si se omite el valor predeterminado, devuelve un valor nulo.
elt(n, input1, input2, ...)	Devuelve la n-ésima entrada, por ejemplo, devuelve `input2` cuando n es 2.
encode(str, charset)	Codifica el primer argumento utilizando el conjunto de caracteres del segundo argumento .

Función	Description (Descripción)
endswith(left, right)	Devuelve un BOOLEANO. El valor es verdadero si left termina con right. Devuelve NULL si alguna de las expresiones de entrada es NULL. De lo contrario, devuelve que el valor es falso. Tanto left como right deben ser de tipo STRING o BINARY.
find_in_set(str, str_array)	Devuelve el índice (basado en 1) de la cadena dada (`str`) en la lista delimitada por comas (`str_array`). Devuelve 0 si no se encontró la cadena o si la cadena dada (`str`) contiene una coma.
format_number(expr1, expr2)	Formatea el número `expr1` como '#, ###, ##.##', redondeándolo a `expr2` decimales. Si `expr2` es 0, el resultado no tiene separador decimal ni parte fraccionaria. `expr2` también acepta un formato especificado por el usuario. Se supone que funciona como el FORMAT de MySQL.
format_string(strfmt, obj, ...)	Devuelve una cadena formateada a partir de cadenas de formato de estilo printf.
initcap(str)	Devuelve `str` con la primera letra de cada palabra en mayúscula. Todas las demás letras están en minúscula. Las palabras se delimitan mediante espacios en blanco.
instr(str, substr)	Devuelve el índice (basado en 1) de la primera aparición de `substr` en `str`.
lcase(str)	Devuelve `str` con todos los caracteres cambiados a minúsculas.

Función	Description (Descripción)
left(str, len)	Devuelve los caracteres más a la izquierda `len` (`len` pueden ser de tipo cadena) de la cadena `str`; si `len` es menor o igual que 0, el resultado es una cadena vacía.
len(expr)	Devuelve la longitud en caracteres de la cadena de datos o el número de bytes de los datos binarios. La longitud de los datos de cadena incluye los espacios finales. La longitud de los datos binarios incluye los ceros binarios.
length(expr)	Devuelve la longitud en caracteres de la cadena de datos o el número de bytes de los datos binarios. La longitud de los datos de cadena incluye los espacios finales. La longitud de los datos binarios incluye los ceros binarios.
levenshtein(str1, str2[, threshold])	Devuelve la distancia de Levenshtein entre las dos cadenas dadas. Si se establece un umbral y la distancia es superior a este, devuelve -1.
locate(substr, str[, pos])	Devuelve la posición de la primera aparición de `substr` en `str` después de la posición de `pos`. El valor `pos` dado y el valor devuelto se basan en 1.
lower(str)	Devuelve `str` con todos los caracteres cambiados a minúsculas.

Función	Description (Descripción)
lpad(str, len[, pad])	Devuelve `str`, rellenado a la izquierda con `pad` una longitud de `len`. Si `str` es más larga que `len`, el valor devuelto se acorta al `len` de caracteres de ancho. Si no se especifica `pad`, `str` se rellenará a la izquierda con espacios si es una cadena de caracteres y con ceros si es una secuencia de bytes.
ltrim(str)	Elimina los caracteres de espacio iniciales de `str`.
luhn_check(str)	Comprueba que una cadena de dígitos es válida según el algoritmo de Luhn. Esta función de suma de verificación se aplica en general a los números de tarjetas de crédito y a los números de identificación gubername ntales para distinguir los números válidos de los números mal escritos o incorrectos.
mask(input[, upperChar, lowerChar, digitChar, otherChar])	Enmascara el valor de cadena determinado. La función reemplaza los caracteres por “X” o “x” y los números por “n”. Esto puede resultar útil para crear copias de tablas que tiene información confidencial que ha sido eliminada .
octet_length(expr)	Devuelve la longitud en bytes de los datos de cadena o el número de bytes de los datos binarios.
overlay(input, replace, pos[, len])	Reemplaza `input` por `replace` que comience en `pos` y sea `len` de largo.

Función	Description (Descripción)
position(substr, str[, pos])	Devuelve la posición de la primera aparición de `substr` en `str` después de la posición de `pos`. El valor `pos` dado y el valor devuelto se basan en 1.
printf(strfmt, obj, ...)	Devuelve una cadena formateada a partir de cadenas de formato de estilo printf.
regexp_count(str, regexp)	Devuelve un recuento del número de veces que el patrón `regexp` de expresión regular coincide en la cadena `str`.
regexp_extract(str, regexp[, idx])	Extraiga la primera cadena de la `str` que coincida con la `regexp` expresión y que corresponda al índice del grupo de expresion es regulares.
regexp_extract_all(str, regexp[, idx])	Extraiga todas las cadenas que coincidan con la expresión `str` y que correspondan al índice del grupo de expresiones regulares `regexp`.
regexp_instr(str, regexp)	Busca una cadena para un patrón de expresión regular y devuelve un valor entero que indica la posición de inicio o de finalizac ión de la subcadena coincidente. Las posicione s se basan en 1, no en 0. Devuelve 0 si no se encuentran coincidencias.
regexp_replace(str, regexp, rep[, position])	Sustituye todas las subcadenas `str` de esa coincidencia `regexp` por `rep`.
regexp_substr(str, regexp)	Devuelve la subcadena que coincide con la expresión `regexp` regular de la cadena `str`. Si no se encuentra la expresión regular, el resultado también será nulo.

Función	Description (Descripción)
repeat(str, n)	Devuelve la cadena que repite el valor de cadena dado n veces.
replace(str, search[, replace])	Sustituye todas las apariciones de `search` por `replace`.
right(str, len)	Devuelve los caracteres más a la derecha `len` (`len` pueden ser de tipo cadena) de la cadena `str`; si `len` es menor o igual que 0, el resultado es una cadena vacía.
rpad(str, len[, pad])	Devuelve `str`, rellenado a la derecha con `pad` una longitud de `len`. Si `str` es más larga que `len`, el valor devuelto se acorta al número de caracteres de `len`. Si `pad` no se especifica, `str` se rellenará a la derecha con espacios si es una cadena de caracteres y con ceros si es una cadena binaria.
rtrim(str)	Elimina los caracteres de espacio del final de `str`.
sentences(str[, lang, country])	`str` se divide en un conjunto de palabras.
soundex(str)	Devuelve el código Soundex de la cadena.
space(n)	Devuelve una cadena que consta de n espacios.
split(str, regex, limit)	Se divide `str` en función de las ocurrencias que coinciden `regex` y devuelve una matriz con una longitud máxima de `limit`

Función	Description (Descripción)
split_part(str, delimiter, partNum)	Divide `str` por delimitador y devuelve la parte solicitada de la división (basada en 1). Si alguna entrada es nula, devuelve null. Si `partNum` está fuera del rango de partes divididas, devuelve una cadena vacía. Devuelve error si `partNum` es 0. Si `partNum` es negativo, las partes se cuentan hacia atrás desde el final de la cadena. Si `delimiter` es una cadena vacía, `str` no está dividida.
startswith(left, right)	Devuelve un BOOLEANO. El valor es verdadero si left comienza con right. Devuelve NULL si alguna de las expresiones de entrada es NULL. De lo contrario, devuelve que el valor es falso. Tanto left como right deben ser de tipo STRING o BINARY.
substr(str, pos[, len])	Devuelve la subcadena `str` que comienza en `pos` y es de longitud `len`, o el segmento de la matriz de bytes que comienza en `pos` y es de longitud `len`.
substr(str FROM pos[FOR len]])	Devuelve la subcadena `str` que comienza en `pos` y es de longitud `len`, o el segmento de la matriz de bytes que comienza en `pos` y es de longitud `len`.
substring(str, pos[, len])	Devuelve la subcadena `str` que comienza en `pos` y es de longitud `len`, o el segmento de la matriz de bytes que comienza en `pos` y es de longitud `len`.

Función	Description (Descripción)
substring(str FROM pos[FOR len]])	Devuelve la subcadena `str` que comienza en `pos` y es de longitud `len`, o el segmento de la matriz de bytes que comienza en `pos` y es de longitud `len`.
substring_index(str, delim, count)	Devuelve la subcadena `str` antes de la aparición `count` del delimitador `delim`. Si `count` es positivo, devuelve todo lo que esté a la izquierda del delimitador final (contando desde la izquierda). Si `count` es negativo, devuelve todo lo que esté a la derecha del delimitador final (contando desde la derecha). La función substring_index realiza una búsqueda de `delim` que distingue entre mayúsculas y minúsculas.
to_binary(str[, fmt])	Convierte la entrada `str` en un valor binario en función del `fmt` proporcionado. `fmt` puede ser un literal de cadena “hex”, “utf-8”, “utf8” o “base64” que no distinga mayúsculas de minúsculas. De forma predeterminada, el formato binario para la conversión es “hex” si se omite el `fmt`. La función devuelve NULL si al menos uno de los parámetros de entrada es NULL.

Función	Description (Descripción)
to_char(numberExpr, formatExpr)	Convierte `numberExpr` en una cadena basada en `formatExpr`. Genera una excepción si no se cumple ninguna de las condiciones. El formato puede constar de los siguientes caracteres, sin distinguir entre mayúsculas y minúsculas: “0” o “9”: especifica un dígito esperado entre 0 y 9. Una secuencia de 0 o 9 en la cadena de formato coincide con una secuencia de dígitos del valor de entrada, lo que genera una cadena de resultado s de la misma longitud que la secuencia correspondiente de la cadena de formato. La cadena resultante se rellena con ceros a la izquierda si la secuencia 0/9 incluye más dígitos que la parte correspondiente del valor decimal, comienza por 0 y está antes de la coma decimal. De lo contrario, se rellena con espacios. “.” o “D”: especifica la posición de la coma decimal (opcional, solo se permite una vez). “,” o “G”: especifica la posición del separador de agrupamiento (miles) (,). Debe haber un 0 o un 9 a la izquierda y a la derecha de cada separador de agrupamiento. '

Función	Description (Descripción)
tto_number(expr, fmt)	Convierte la cadena 'expr' en un número según el formato de cadena 'fmt'. Genera una excepción si no se cumple ninguna de las condiciones. El formato puede constar de los siguientes caracteres, sin distinguir entre mayúsculas y minúsculas: “0” o “9”: especifica un dígito esperado entre 0 y 9. Una secuencia de 0 o 9 en la cadena de formato coincide con una secuencia de dígitos de la cadena de entrada. Si la secuencia 0/9 comienza por 0 y está antes de la coma decimal, solo puede coincidir con una secuencia de dígitos del mismo tamaño. De lo contrario, si la secuencia comienza con 9 o está después de la coma decimal, puede coincidir con una secuencia de dígitos que tenga el mismo tamaño o menor. “.” o “D”: especifica la posición de la coma decimal (opcional, solo se permite una vez). “,” o “G”: especifica la posición del separador de agrupamiento (miles) (,). Debe haber un 0 o un 9 a la izquierda y a la derecha de cada separador de agrupamiento. “expr” debe coincidir con el separador de agrupamiento correspondiente al tamaño del número. '

Función	Description (Descripción)
to_varchar(numberExpr, formatExpr)	Convierte `numberExpr` en una cadena basada en `formatExpr`. Genera una excepción si no se cumple ninguna de las condiciones. El formato puede constar de los siguientes caracteres, sin distinguir entre mayúsculas y minúsculas: “0” o “9”: especifica un dígito esperado entre 0 y 9. Una secuencia de 0 o 9 en la cadena de formato coincide con una secuencia de dígitos del valor de entrada, lo que genera una cadena de resultado s de la misma longitud que la secuencia correspondiente de la cadena de formato. La cadena resultante se rellena con ceros a la izquierda si la secuencia 0/9 incluye más dígitos que la parte correspondiente del valor decimal, comienza por 0 y está antes de la coma decimal. De lo contrario, se rellena con espacios. “.” o “D”: especifica la posición de la coma decimal (opcional, solo se permite una vez). “,” o “G”: especifica la posición del separador de agrupamiento (miles) (,). Debe haber un 0 o un 9 a la izquierda y a la derecha de cada separador de agrupamiento. '
translate(input, from, to)	Traduce la cadena `input` y los caracteres encontrados en la cadena `from` reemplaza dos por los caracteres correspondientes en la cadena `to`.
trim(str)	Elimina los caracteres de espacio inicial y final de `str`.
trim(BOTH FROM str)	Elimina los caracteres de espacio inicial y final de `str`.

Función	Description (Descripción)
trim(LEADING FROM str)	Elimina los caracteres de espacio iniciales de `str`.
trim(TRAILING FROM str)	Elimina los caracteres de espacio del final de `str`.
trim(trimStr FROM str)	Elimine los caracteres `trimStr` iniciales y finales de `str`.
trim(BOTH trimStr FROM str)	Elimine los caracteres `trimStr` iniciales y finales de `str`.
trim(LEADING trimStr FROM str)	Elimine los caracteres `trimStr` iniciales de `str`.
trim(TRAILING trimStr FROM str)	Elimine los caracteres `trimStr` finales de `str`.
try_to_binary(str[, fmt])	Se trata de una versión de `to_binary` especial que realiza la misma operación, pero devuelve un valor NULL en lugar de generar un error si no se puede realizar la conversión.
try_to_number(expr, fmt)	Convierte la cadena 'expr' en un número según el formato de la cadena `fmt`. Devuelve NULL si la cadena 'expr' no coincide con el formato esperado. El formato sigue la misma semántica que la función to_number.
ucase(str)	Devuelve `str` con todos los caracteres cambiados a mayúsculas.
unbase64(str)	Convierte el argumento de una cadena `str` de base 64 en un binario.
upper(str)	Devuelve `str` con todos los caracteres cambiados a mayúsculas.

Ejemplos

```
-- 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|
+---------------+
```

Funciones de fecha y hora

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
add_months(start_date, num_months)	Devuelve la fecha que es `num_months` después de `start_date` .
convert_timezone([sourceTz,]targetTz, sourceTs)	Convierte la marca de tiempo `sourceTs` sin zona horaria `sourceTz` de a la zona horaria `targetTz`.
curdate()	Devuelve la fecha actual al inicio de la evaluación de la consulta. Todas las llamadas a curdate dentro de la misma consulta devuelven el mismo valor.
current_date()	Devuelve la fecha actual al inicio de la evaluación de la consulta. Todas las llamadas a current_date dentro de la misma consulta devuelven el mismo valor.
current_date	Devuelve la fecha actual al inicio de la evaluación de la consulta.
current_timestamp()	Devuelve la marca de tiempo actual al inicio de la evaluación de la consulta. Todas las

Función	Description (Descripción)
	llamadas a current_timestamp dentro de la misma consulta devuelven el mismo valor.
current_timestamp	Devuelve la marca de tiempo actual al inicio de la evaluación de la consulta.
current_timezone()	Devuelve la zona horaria local de la sesión actual.
date_add(start_date, num_days)	Devuelve la fecha que es `num_days` después de `start_date` .
date_diff(endDate, startDate)	Devuelve el número de días comprendido entre `startDate` y `endDate`.
date_format(timestamp, fmt)	Se convierte `timestamp` en un valor de cadena en el formato especificado por el formato de fecha `fmt`.
date_from_unix_date(days)	Crea una fecha a partir del número de días transcurridos desde 1970-01-01.
date_part(field, source)	Extrae una parte de la fuente date/timestamp o del intervalo.
date_sub(start_date, num_days)	Devuelve la fecha que es `num_days` antes de `start_date` .
date_trunc(fmt, ts)	Devuelve la marca de tiempo `ts` truncada a la unidad especificada por el modelo de formato `fmt`.
dateadd(start_date, num_days)	Devuelve la fecha que es `num_days` después de `start_date` .
datediff(endDate, startDate)	Devuelve el número de días comprendido entre `startDate` y `endDate`.

Función	Description (Descripción)
datepart(field, source)	Extrae una parte de la fuente del intervalo date/timestamp o.
day(date)	Devuelve el día del mes de la fecha/marca horaria.
dayofmonth(date)	Devuelve el día del mes de la fecha/marca horaria.
dayofweek(date)	Devuelve el día de la semana de date/time stamp (1 = domingo, 2 = lunes,..., 7 = sábado).
dayofyear(date)	Devuelve el día del año de la fecha/marca horaria.
extract(field FROM source)	Extrae una parte de la fuente del intervalo date/timestamp o.
from_unixtime(unix_time[, fmt])	Devuelve el valor `unix_time` especificado en `fmt`.
from_utc_timestamp(timestamp, timezone)	Dada una marca de tiempo como “2017-07-14 02:40:00.0”, la interpreta como una hora en UTC y representa esa hora como una marca de tiempo en la zona horaria determinada. Por ejemplo, “GMT+1” daría como resultado “2017-07-14 03:40:00.0”.
hour(timestamp)	Devuelve el componente horario de la cadena/marca de tiempo.
last_day(date)	Devuelve el último día del mes al que pertenece la fecha.

Función	Description (Descripción)
localtimestamp()	Devuelve la marca de tiempo actual sin zona horaria al inicio de la evaluación de la consulta. Todas las llamadas a localtimestamp dentro de la misma consulta devuelven el mismo valor.
localtimestamp	Devuelve la fecha y hora local actual en la zona horaria de la sesión al inicio de la evaluación de la consulta.
make_date(year, month, day)	Crea una fecha a partir de los campos de año, mes y día.
make_dt_interval([days[, hours[, mins[, secs]]]])	Calcula DayTimeIntervalType la duración a partir de días, horas, minutos y segundos.
make_interval([years[, months[, weeks[, days[, hours[, mins[, secs]]]]]]])	Hace intervalos entre años, meses, semanas, días, horas, minutos y segundos.
make_timestamp(year, month, day, hour, min, sec[, timezone])	Crea una marca de tiempo a partir de los campos de año, mes, día, hora, minuto, segundo y zona horaria.
make_timestamp_ltz(year, month, day, hour, min, sec[, timezone])	Crea la marca de tiempo actual con la zona horaria local a partir de los campos de año, mes, día, hora, minutos, segundos y zona horaria.
make_timestamp_ntz(year, month, day, hour, min, sec)	Crea una fecha y hora local a partir de los campos de año, mes, día, hora, minutos y segundos.
make_ym_interval([years[, months]])	Hace un intervalo año-mes a partir de años, meses.
minute(timestamp)	Devuelve el componente de minutos de la cadena/marca de tiempo.

Función	Description (Descripción)
month(date)	Devuelve el componente mensual de la fecha/marca horaria.
months_between(timestamp1, timestamp2[, roundOff])	Si `timestamp1` es posterior a `timestamp2`, entonces el resultado es positivo. Si `timestamp1` y `timestamp2` son el mismo día del mes, o ambos son el último día del mes, se ignorará la hora del día. De lo contrario, la diferencia se calcula en función de 31 días al mes y se redondea a 8 dígitos, a menos que exista roundOff=false.
next_day(start_date, day_of_week)	Devuelve la primera fecha posterior a `start_date` y cuyo nombre es el indicado. La función devuelve NULL si al menos uno de los parámetros de entrada es NULL.
now()	Devuelve la marca de tiempo actual al inicio de la evaluación de la consulta.
quarter(date)	Devuelve el trimestre del año correspondiente a la fecha, en un intervalo de 1 a 4.
second(timestamp)	Devuelve el segundo componente de la cadena/marca de tiempo.
session_window(time_column, gap_duration)	Genera una ventana de sesión con una marca de tiempo que especifica la duración de la columna y el intervalo. Consulte “Tipos de intervalos” en el documento de la guía de transmisión estructurada para obtener una explicación detallada y ejemplos.
timestamp_micros(microseconds)	Crea una marca de tiempo a partir del número de microsegundos transcurridos desde la época UTC.

Función	Description (Descripción)
timestamp_millis(milliseconds)	Crea una marca de tiempo a partir del número de milisegundos transcurridos desde la época UTC.
timestamp_seconds(seconds)	Crea una marca de tiempo a partir del número de segundos (puede ser fraccionario) desde la época UTC.
to_date(date_str[, fmt])	Analiza la expresión `date_str` con la expresión `fmt` hasta una fecha. Devuelve un valor nulo con una entrada no válida. De forma predeterminada, sigue las reglas de selección hasta una fecha si `fmt` se omite.
to_timestamp(timestamp_str[, fmt])	Analiza la expresión `timestamp_str` junto con la expresión `fmt` en una marca de tiempo. Devuelve un valor nulo con una entrada no válida. De forma predeterminada, sigue las reglas de conversión a una marca de tiempo si `fmt` se omite.
to_timestamp_ltz(timestamp_str[, fmt])	Analiza la expresión `timestamp_str` con la expresión `fmt` para convertirla en una marca de tiempo con la zona horaria local. Devuelve un valor nulo con una entrada no válida. De forma predeterminada, sigue las reglas de conversión a una marca de tiempo si `fmt` se omite.
to_timestamp_ntz(timestamp_str[, fmt])	Analiza la expresión `timestamp_str` con la expresión `fmt` para convertirla en una marca de tiempo sin zona horaria. Devuelve un valor nulo con una entrada no válida. De forma predeterminada, sigue las reglas de conversió n a una marca de tiempo si `fmt` se omite.

Función	Description (Descripción)
to_unix_timestamp(timeExp[, fmt])	Devuelve la marca de tiempo de UNIX de la hora dada.
to_utc_timestamp(timestamp, timezone)	Dada una marca de tiempo como “2017-07-14 02:40:00.0”, la interpreta como una hora de la zona horaria determinada y representa esa hora como una marca de tiempo en UTC. Por ejemplo, “GMT+1” daría como resultado “2017-07-14 01:40:00.0”.
trunc(date, fmt)	Devuelve `date` con la parte horaria del día truncada a la unidad especificada por el modelo de formato `fmt`.
try_to_timestamp(timestamp_str[, fmt])	Analiza la expresión `timestamp_str` junto con la expresión `fmt` en una marca de tiempo.
unix_date(date)	Devuelve el número de días transcurridos desde 1970-01-01.
unix_micros(timestamp)	Devuelve el número de microsegundos transcurridos desde 1970-01-01 00:00:00 UTC.
unix_millis(timestamp)	Devuelve el número de milisegundos transcurridos desde 1970-01-01 00:00:00 UTC. Trunca los niveles de precisión más altos.
unix_seconds(timestamp)	Devuelve el número de segundos transcurridos desde 1970-01-01 00:00:00 UTC. Trunca los niveles de precisión más altos.
unix_timestamp([timeExp[, fmt]])	Devuelve la marca de tiempo de UNIX de la hora actual o especificada.
weekday(date)	Devuelve el día de la semana de date/time stamp (0 = lunes, 1 = martes,..., 6 = domingo).

Función	Description (Descripción)
weekofyear(date)	Devuelve la semana del año de la fecha indicada. Se considera que una semana comienza un lunes y la semana 1 es la primera semana con más de 3 días.
window(time_column, window_duration[, slide_duration[, start_time]])	Clasifica las filas en una o más ventanas de tiempo con una columna que especifique la marca de tiempo. Los inicios de las ventanas son inclusivos, pero los finales son exclusivos, por ejemplo, las 12:05 estarán en la ventana [12:05,12:10) pero no en [12:00,12:05). Windows admite una precisión de microsegundos. No se admite Windows para el orden de los meses. Consulte “Operaciones de intervalo en tiempos de eventos” en el documento de la guía de transmisión estructurada para obtener una explicación detallada y ejemplos.
window_time(window_column)	Extraiga el valor de tiempo de time/session la columna de la ventana, que se puede usar como valor de tiempo del evento de la ventana. El tiempo extraído es (window.end - 1), lo que refleja el hecho de que los intervalos de agregación tienen un límite superior exclusivo (inicio, final). Consulte “Operaciones de intervalo en tiempos de eventos” en el documento de la guía de transmisión estructurada para obtener una explicación detallada y ejemplos.
year(date)	Devuelve el componente anual de la fecha/marca horaria.

Ejemplos

```
-- 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|
+----------------+
```

Funciones de agregación

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consultethe section called “Comandos SQL compatibles”.

Las funciones de agregado funcionan con los valores de las filas para realizar cálculos matemáticos como la suma, el promedio, el recuento, minimum/maximum los valores, la desviación estándar y la estimación, así como algunas operaciones no matemáticas.

Sintaxis

```
aggregate_function(input1 [, input2, ...]) FILTER (WHERE boolean_expression)
```

Parámetros

- `boolean_expression`: especifica cualquier expresión que se evalúe como un resultado de tipo booleano. Se pueden combinar dos o más expresiones mediante los operadores lógicos (AND, OR).

Funciones agregadas de conjuntos ordenados

Estas funciones de agregación utilizan una sintaxis diferente a la de las demás funciones de agregación para especificar una expresión (normalmente el nombre de una columna) con la que ordenar los valores.

Sintaxis

```
{ PERCENTILE_CONT | PERCENTILE_DISC }(percentile) WITHIN GROUP (ORDER BY
 { order_by_expression [ ASC | DESC ] [ NULLS { FIRST | LAST } ] [ , ... ] }) FILTER
 (WHERE boolean_expression)
```

Parámetros

- `percentile`: el percentil del valor que se desea buscar. El percentil debe ser una constante entre 0,0 y 1,0.
- `order_by_expression`: la expresión (normalmente el nombre de una columna) con la que se ordenan los valores antes de agregarlos.
- `boolean_expression`: especifica cualquier expresión que se evalúe como un resultado de tipo booleano. Se pueden combinar dos o más expresiones mediante los operadores lógicos (AND, OR).

Ejemplos

```
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|
+----------+-------+--------+-------+--------+-----+-----+-----+-----+
```

Funciones condicionales

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consultethe section called “Comandos SQL compatibles”.

Función	Description (Descripción)
coalesce(expr1, expr2, ...)	Devuelve el primer argumento no nulo si existe. De lo contrario, es nulo.
if(expr1, expr2, expr3)	Si se evalúa `expr1` como verdadero, devuelve `expr2`; de lo contrario, devuelve `expr3`.
ifnull(expr1, expr2)	Devuelve `expr2` si `expr1` es nulo, o `expr1`.
nanvl(expr1, expr2)	Devuelve `expr1` si no es NaN, o `expr2`.
nullif(expr1, expr2)	Devuelve null si `expr1` es igual a `expr2`, si no, devuelve `expr1`.
nvl(expr1, expr2)	Devuelve `expr2` si `expr1` es nulo, o `expr1`.
nvl2(expr1, expr2, expr3)	Devuelve `expr2` si `expr1` no es nulo, o `expr3`.
CASE WHEN expr1 THEN expr2 [WHEN expr3 THEN expr4]* [ELSE expr5] END	Cuando `expr1` es true, devuelve `expr2`; si no, cuando `expr3` es true, devuelve `expr4`; si no, devuelve `expr5`.

Ejemplos

```
-- 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|
+-----------------------------------------------+
```

Funciones JSON

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
from_json (jsonStr, schema[, options])	Devuelve un valor de estructura con los valores `jsonStr` y `schema` dados.
get_json_ object(js on_txt, path)	Extrae un objeto json de `path`.

Función	Description (Descripción)
json_array_length(jsonArray)	Devuelve la cantidad de elementos en la matriz exterior de una cadena JSON.
json_object_keys(json_object)	Devuelve todas las claves del objeto JSON más externo como una matriz.
json_tuple(jsonStr, p1, p2, ..., pn)	Devuelve una tupla como la función get_json_object, pero toma varios nombres. Todos los parámetros de entrada y los tipos de columnas de salida son cadenas.
schema_of_json(json[, options])	Devuelve el esquema en el formato DDL de la cadena JSON.
to_json(expr[, options])	Devuelve una cadena JSON con un valor de estructura determinado

Ejemplos

```
-- 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}]               |
+-------------------------+
```

Funciones de matriz

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
array(expr, ...)	Devuelve una matriz con los elementos dados.
array_append(array, element)	Agrega el elemento al final de la matriz pasada como primer argumento. El tipo de elemento debe ser similar al tipo de los elementos de la matriz. El elemento nulo también se agrega a la matriz. Pero si la matriz se pasa, es NULL, la salida es NULL
array_compact(array)	Elimina los valores nulos de la matriz.
array_contains(array, value)	Devuelve verdadero si la matriz contiene el valor.
array_distinct(array)	Elimina los valores duplicados de la matriz.
array_except(array1, array2)	Devuelve una matriz de los elementos de la array1 pero no de la array2, sin duplicados.
array_insert(x, pos, val)	Coloca val en el punto de índice de la matriz x. Los índices comienzan en 1. El índice negativo máximo es -1 para el que la función inserta un

Función	Description (Descripción)
	nuevo elemento después del último elemento actual. El índice por encima del tamaño de la matriz se agrega a la matriz, o antepone la matriz si el índice es negativo, con elementos “nulos”.
array_intersect(array1, array2)	Devuelve una matriz de los elementos en la intersección de array1 y array2, sin duplicados.
array_join(array, delimiter[, nullReplacement])	Concatena los elementos de la matriz dada mediante el delimitador y una cadena opcional para reemplazar los valores nulos. Si no se establece ningún valor para nullReplacement, se filtra cualquier valor nulo.
array_max(array)	Devuelve el valor máximo de la matriz. NaN es mayor que cualquier elemento que no sea NaN por double/float tipo. Se omiten los elementos NULL.
array_min(array)	Devuelve el valor mínimo de la matriz. NaN es mayor que cualquier elemento que no sea NaN por double/float tipo. Se omiten los elementos NULL.
array_position(array, element)	Devuelve el índice (basado en 1) del primer elemento coincidente de la matriz siempre que sea largo, o 0 si no se encuentra ninguna coincidencia.

Función	Description (Descripción)
array_prepend(array, element)	Agrega el elemento al principio de la matriz pasada como primer argumento. El tipo de elemento debe ser el mismo que el tipo de los elementos de la matriz. El elemento nulo también se antepone a la matriz. Pero si la matriz pasada era NULL, la salida va a ser NULL
array_remove(array, element)	Elimine todos los elementos que sean iguales al elemento de la matriz.
array_repeat(element, count)	Devuelve la matriz que contiene los tiempos de recuento de elementos.
array_union(array1, array2)	Devuelve una matriz de los elementos de la unión de array1 y array2, sin duplicados.
arrays_overlap(a1, a2)	Devuelve verdadero si a1 contiene al menos un elemento no nulo presente también en a2. Si las matrices no tienen ningún elemento común, ninguna está vacía y alguna de ellas contiene un elemento nulo, se devuelve null y falso en caso contrario.
arrays_zip(a1, a2, ...)	Devuelve una matriz combinada de estructur as en la que la estructura n-ésima contiene todos los valores n-ésimos de las matrices de entrada.
aplanar () arrayOfArrays	Transforma una matriz de matrices en una sola matriz.
get(array, index)	Devuelve el elemento de la matriz en un índice dado (basado en 0). Si el índice apunta fuera de los límites de la matriz, esta función devuelve NULL.

Función	Description (Descripción)
sequence(start, stop, step)	Genera una matriz de elementos desde el principio hasta la parada (ambos incluidos), incrementándose paso a paso. El tipo de valor devuelto es el mismo que el tipo de la expresión. Los tipos compatibles son byte, short, integer, long, date, timestamp. Las expresiones de inicio y parada deben tener el mismo tipo de resolución. Si las expresiones de inicio y finalización se resuelven en el tipo “date” o “timestamp”, la expresión escalonada debe resolverse en el tipo “interval”, “year-mon th interval” o “day-time interval”; de lo contrario , debe ser del mismo tipo que las expresiones de inicio y finalización.
shuffle(array)	Devuelve una permutación aleatoria de la matriz dada.
slice(x, start, length)	Subestablece la matriz x a partir del inicio del índice (los índices de la matriz comienzan en 1 o comienzan desde el final si el inicio es negativo) con la longitud especificada.
sort_array(array[, ascendingOrder])	Ordena la matriz de entrada en orden ascendente o descendente de acuerdo con el orden natural de los elementos de la matriz. NaN es mayor que cualquier elemento que no sea NaN por double/float tipo. Los elementos nulos se colocarán al principio de la matriz devuelta en orden ascendente o al final de la matriz devuelta en orden descendente.

Ejemplos

```
-- 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]|
+---------------------------------------+
```

Funciones de ventana

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Las funciones de ventana funcionan en un grupo de filas, denominado ventana y calculan un valor de devolución para cada fila en función del grupo de filas. Las funciones de ventana son útiles para procesar tareas como calcular un promedio móvil, computar una estadística acumulada o acceder al valor de las filas dada la posición relativa de la fila actual.

Sintaxis

```
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 ] )
```

Parámetros

- Funciones de clasificación

  Sintaxis: `RANK | DENSE_RANK | PERCENT_RANK | NTILE | ROW_NUMBER`

  Funciones analíticas

  Sintaxis: `CUME_DIST | LAG | LEAD | NTH_VALUE | FIRST_VALUE | LAST_VALUE`

Funciones de agregación

Sintaxis: `MAX | MIN | COUNT | SUM | AVG | ...`

- `nulls_option`: Especifica si se deben omitir o no los valores nulos al evaluar la función de ventana. RESPECT NULLS significa no omitir los valores nulos, mientras que IGNORE NULLS significa omitirlos. Si no se especifica, el valor predeterminado es RESPECT NULLS.

  Sintaxis: `{ IGNORE | RESPECT } NULLS`

  Nota: `Only LAG | LEAD | NTH_VALUE | FIRST_VALUE | LAST_VALUE` se pueden utilizar con `IGNORE NULLS`.
- `window_frame`: Especifica en qué fila se debe iniciar la ventana y dónde se debe terminar.

  Sintaxis: `{ RANGE | ROWS } { frame_start | BETWEEN frame_start AND frame_end }`

  frame_start y frame_end tienen la siguiente sintaxis:

  Sintaxis: `UNBOUNDED PRECEDING | offset PRECEDING | CURRENT ROW | offset FOLLOWING | UNBOUNDED FOLLOWING`

  offset: especifica el desplazamiento desde la posición de la fila actual.

  Nota: Si se omite frame_end, el valor predeterminado es CURRENT ROW.

Ejemplos

```
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|
+--+----+----+----+---------+-----------+----------+
```

Funciones de conversión

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
bigint(expr)	Transfiere el valor `expr` al tipo de datos de destino `bigint`.

Función	Description (Descripción)
binary(expr)	Transfiere el valor `expr` al tipo de datos de destino `binario`.
boolean(expr)	Transfiere el valor `expr` al tipo de datos de destino `boolean`.
cast(expr AS type)	Convierte el valor `expr` al tipo de datos de destino `type`.
date(expr)	Transfiere el valor `expr` al tipo de datos de destino `date`.
decimal(expr)	Convierte el valor `expr` al tipo de datos de destino `decimal`.
double(expr)	Convierte el valor `expr` al tipo de datos de destino `double`.
float(expr)	Convierte el valor `expr` al tipo de datos de destino `float`.
int(expr)	Transfiere el valor `expr` al tipo de datos de destino `int`.
smallint(expr)	Transfiere el valor `expr` al tipo de datos de destino `smallint`.
string(expr)	Transfiere el valor `expr` al tipo de datos de destino `string`.
timestamp(expr)	Transfiere el valor `expr` al tipo de datos de destino `timestamp`.
tinyint(expr)	Transfiere el valor `expr` al tipo de datos de destino `tinyint`.

Ejemplos

```
-- cast
SELECT cast(field as int);
+---------------+
|CAST(field AS INT)|
+---------------+
|             10|
+---------------+
```

Funciones de predicados

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
! expr	No lógico.
expr1 < expr2	Devuelve el valor verdadero si `expr1` es menor que `expr2`.
expr1 <= expr2	Devuelve verdadero si `expr1` es menor o igual que `expr2`.
expr1 <=> expr2	Devuelve el mismo resultado que el operador EQUAL (=) para los operandos no nulos, pero devuelve verdadero si ambos son nulos y falso si uno de ellos es nulo.
expr1 = expr2	Devuelve verdadero si `expr1` es igual a `expr2`, o falso en caso contrario.
expr1 == expr2	Devuelve verdadero si `expr1` es igual a `expr2`, o falso en caso contrario.
expr1 > expr2	Devuelve verdadero si `expr1` es mayor que `expr2`.
expr1 >= expr2	Devuelve verdadero si `expr1` es mayor o igual que `expr2`.
expr1 y expr2	AND lógico.
str ilike pattern[ESCAPE escape]	Devuelve verdadero si str hace coincidir `pattern` con `escape` sin distinguir mayúsculas de minúsculas, null si algún argumento es nulo, falso en caso contrario.
expr1 in(expr2, expr3, ...)	Devuelve verdadero si `expr` es igual a cualquier valN.
isnan(expr)	Devuelve verdadero si `expr` es NaN, o falso en caso contrario.

Función	Description (Descripción)
isnotnull(expr)	Devuelve verdadero si `expr` no es nulo, o falso en caso contrario.
isnull(expr)	Devuelve verdadero si `expr` es nulo o falso en caso contrario.
str like pattern[ESCAPE escape]	Devuelve verdadero si str coincide con `pattern` con `escape`, null si algún argumento es nulo, falso en caso contrario.
not expr	No lógico.
expr1 o expr2	OR lógico.
regexp(str, regexp)	Devuelve verdadero si `str` coincide con `regexp`, o falso en caso contrario.
regexp_like(str, regexp)	Devuelve verdadero si `str` coincide con `regexp`, o falso en caso contrario.
rlike(str, regexp)	Devuelve verdadero si `str` coincide con `regexp`, o falso en caso contrario.

Ejemplos

```
-- !
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|
+---------------+
```

Funciones de mapa

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
element_at(array, index)	Devuelve el elemento de la matriz en un índice dado (basado en 1).
element_at(map, key)	Devuelve el valor de una clave dada. La función devuelve NULL si la clave no se encuentra en el mapa.
map(key0, value0, key1, value1, ...)	Crea un mapa con los key/value pares dados.
map_concat(map, ...)	Devuelve la unión de todos los mapas dados
map_contains_key(map, key)	Devuelve true si el mapa contiene la clave especificada.
map_entries(map)	Devuelve una matriz desordenada de todas las entradas del mapa dado.
map_from_arrays(keys, values)	Crea un mapa con un par de las key/value matrices dadas. Todos los elementos de las claves no deben ser un valor nulo
map_from_entries () arrayOfEntries	Devuelve un mapa creado a partir de la matriz de entradas dada.
map_keys(map)	Devuelve una matriz desordenada que contiene las claves del mapa.
map_values(map)	Devuelve una matriz desordenada que contiene los valores del mapa.
str_to_map (texto [, PairDelim [,]]) keyValueD elim	Crea un mapa después de dividir el texto en pares clave/valor mediante delimitadores. Los delimitadores predeterminados son ',' para `pairDelim` y ':' para ``. keyValueDelim Tanto `PairDelim` como `` se tratan como expresion es regulares. keyValueDelim

Función	Description (Descripción)
try_element_at(array, index)	Devuelve el elemento de la matriz en un índice dado (basado en 1). Si el índice es 0, el sistema generará un error. Si el índice es inferior a 0, accede a los elementos desde el último hasta el primero. La función siempre devuelve NULL si el índice supera la longitud de la matriz.
try_element_at(map, key)	Devuelve el valor de una clave dada. La función siempre devuelve NULL si la clave no está incluida en el mapa.

Ejemplos

```
-- 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|
+---------------------------------+
```

Funciones matemáticas

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
expr1 % expr2	Devuelve el resto después de `expr1`/`expr2`.
expr1 * expr2	Devuelve `expr1` por `expr2`.
expr1 + expr2	Devuelve `expr1` más `expr2`.
expr1 - expr2	Devuelve `expr1` menos `expr2`.
expr1 / expr2	Devuelve `expr1` dividido `expr2`. Siempre realiza una división de punto flotante.
abs(expr)	Devuelve el valor absoluto del valor numérico o del intervalo.
acos(expr)	Devuelve el coseno inverso (también conocido como arcoseno) de `expr`, como si se hubiera calculado con `java.lang.math.acos`.
acosh(expr)	Devuelve el coseno hiperbólico inverso de `expr`.
asin(expr)	Devuelve el seno inverso (también conocido como arco seno), el arco seno de `expr`, como si se hubiera calculado con `java.lang.math.asin`.
asinh(expr)	Devuelve el seno hiperbólico inverso de `expr`.
atan(expr)	Devuelve la tangente inversa (también conocida como arco tangente) de `expr`, como si la hubiera calculado `java.lang.math.atan`
atan2(exprY, exprX)	Devuelve el ángulo en radianes entre el eje x positivo de un plano y el punto dado por las coordenadas (`exprX`, `exprY`), como si se hubiera calculado con `java.lang.Math.atan2`.

Función	Description (Descripción)
atanh(expr)	Devuelve la tangente hiperbólica inversa de `expr`.
bin(expr)	Devuelve la representación en cadena del valor largo `expr` representado en binario.
bround(expr, d)	Devuelve `expr` redondeado a `d` decimales utilizando el modo de redondeo HALF_EVEN.
cbrt(expr)	Devuelve la raíz cúbica de `expr`.
ceil(expr[, scale])	Devuelve el número más pequeño después de redondear hacia arriba que no sea menor que `expr`. Se puede especificar un parámetro `scale` opcional para controlar el comportam iento de redondeo.
ceiling(expr[, scale])	Devuelve el número más pequeño después de redondear hacia arriba que no sea menor que `expr`. Se puede especificar un parámetro `scale` opcional para controlar el comportam iento de redondeo.
conv(num, from_base, to_base)	Convierte `num` de `from_base` a `to_base`.
cos(expr)	Devuelve el coseno de `expr`, como si lo hubiera calculado `java.lang.Math.cos`.
cosh(expr)	Devuelve el coseno hiperbólico de `expr`, como si se hubiera calculado con `java.lan g.Math.cosh`.
cot(expr)	Devuelve la cotangente de `expr`, como si se hubiera calculado con `1/java.lang.Math.tan`.
csc(expr)	Devuelve la cosecante de `expr`, como si se hubiera calculado con `1/java.lang.Math.sin`.

Función	Description (Descripción)
degrees(expr)	Convierte radianes en grados.
expr1 div expr2	Divide `expr1` por `expr2`. Devuelve NULL si un operando es NULL o `expr2` es 0. El resultado se convierte en largo.
e()	Devuelve el número de Euler, e.
expr(expr)	Devuelve e a la potencia de `expr`.
expm1(expr) - Devuelve exp(`expr`)	1
factorial(expr)	Devuelve la factorial de `expr`. `expr` es [0.. 20]. De lo contrario, es nulo.
floor(expr[, scale])	Devuelve el número mayor después de redondear para abajo que no sea mayor que `expr`. Se puede especificar un parámetro `scale` opcional para controlar el comportam iento de redondeo.
greatest(expr, ...)	Devuelve el valor máximo de todos los parámetros, omitiendo los valores nulos.
hex(expr)	Convierte `expr` en hexadecimal.
hypot(expr1, expr2)	Devuelve sqrt(`expr1`**2 + `expr2`**2).
least(expr, ...)	Devuelve el valor mínimo de todos los parámetros, omitiendo los valores nulos.
ln(expr)	Devuelve el logaritmo natural (de base e) de `expr`.
log(base, expr)	Devuelve el logaritmo de `expr` con `base`.
log10(expr)	Devuelve el logaritmo en base 10 de `expr`.

Función	Description (Descripción)
log1p(expr)	Devuelve log(1 + `expr`).
log2(expr)	Devuelve el logaritmo de `expr` con base 2.
expr1 mod expr2	Devuelve el resto después de `expr1`/`expr2`.
negative(expr)	Devuelve el valor negado de `expr`.
pi()	Devuelve pi.
pmod(expr1, expr2)	Devuelve el valor positivo del resto de `expr1`/`expr2`.
positive(expr)	Devuelve el valor de `expr`.
pow(expr1, expr2)	Eleva `expr1` a la potencia de `expr2`.
power(expr1, expr2)	Eleva `expr1` a la potencia de `expr2`.
radians(expr)	Convierte grados en radianes.
rand([seed])	Devuelve un valor aleatorio con valores independientes e idénticos (i.i.d.) distribuidos uniformemente en [0, 1).
randn([seed])	Devuelve un valor aleatorio con valores independientes y de distribución idéntica (i.i.d.) extraídos de la distribución normal estándar.
random([seed])	Devuelve un valor aleatorio con valores independientes e idénticos (i.i.d.) distribuidos uniformemente en [0, 1).
rint(expr)	Devuelve el valor doble que tenga el valor más cercano al argumento y que sea igual a un entero matemático.

Función	Description (Descripción)
round(expr, d)	Devuelve `expr` redondeado a `d` decimales utilizando el modo de redondeo HALF_UP.
sec(expr)	Devuelve la secante de `expr`, como si se hubiera calculado con `1/java.lang.Math.cos`.
shiftleft(base, expr)	Cambio bit a bit hacia la izquierda.
sign(expr)	Devuelve -1.0, 0.0 o 1.0 si `expr` es negativo, 0 o positivo.
signum(expr)	Devuelve -1.0, 0.0 o 1.0 si `expr` es negativo, 0 o positivo.
sin(expr)	Devuelve el seno de `expr`, como si lo hubiera calculado `java.lang.Math.sin`.
sinh(expr)	Devuelve el seno hiperbólico de `expr`, como si se hubiera calculado mediante `java.lang.Math.sinh`.
sqrt(expr)	Devuelve la raíz cuadrada de `expr`.
tan(expr)	Devuelve la tangente de `expr`, como si se hubiera calculado con `java.lang.Math.tan`.
tanh(expr)	Devuelve la tangente hiperbólica de `expr`, como si se hubiera calculado con `java.lang.Math.tanh`.
try_add(expr1, expr2)	Devuelve la suma de `expr1`y `expr2` y el resultado es nulo en caso de desbordamiento. Los tipos de entrada aceptables son los mismos que con el operador `+`.

Función	Description (Descripción)
try_divide(dividend, divisor)	Devuelve `dividend`/`divisor`. Siempre realiza una división de punto flotante. Su resultado siempre es nulo si `expr2` es 0. El `dividend` debe ser un número o un intervalo. El `divisor` debe ser numérico.
try_multiply(expr1, expr2)	Devuelve `expr1`*`expr2` y el resultado es nulo en caso de desbordamiento. Los tipos de entrada aceptables son los mismos que con el operador `*`.
try_subtract(expr1, expr2)	Devuelve `expr1`-`expr2` y el resultado es nulo en caso de desbordamiento. Los tipos de entrada aceptables son los mismos para el operador `-`.
unhex(expr)	Convierte `expr` hexadecimal en binario.
width_bucket(value, min_value, max_value, num_bucket)	Devuelve el número de bucket al que se asignaría `value` en un histograma de ancho equivalente con buckets de `num_bucket`, en el rango de `min_value` a `max_value`.

Ejemplos

```
-- %
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|
+-----------------------------------------------------------------------+
```

Funciones generadoras

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten estas funciones de SQL, consulte. the section called “Comandos SQL compatibles”

Función	Description (Descripción)
explode(expr)	Separa los elementos de la matriz `expr` en varias filas o los elementos del mapa `expr` en varias filas y columnas. A menos que se especifique lo contrario, utiliza el nombre de columna predeterminado `col` para los elementos de la matriz o `key` y `value` para los elementos del mapa.
explode_outer(expr)	Separa los elementos de la matriz `expr` en varias filas o los elementos del mapa `expr` en varias filas y columnas. A menos que se especifique lo contrario, utiliza el nombre de columna predeterminado `col` para los elementos de la matriz o `key` y `value` para los elementos del mapa.
inline(expr)	Descompone un conjunto de estructuras en una tabla. Utiliza los nombres de columna col1, col2, etc. de forma predeterminada, a menos que se especifique lo contrario.
inline_outer(expr)	Descompone un conjunto de estructuras en una tabla. Utiliza los nombres de columna col1, col2, etc. de forma predeterminada, a menos que se especifique lo contrario.
posexplode(expr)	Separa los elementos de la matriz `expr` en varias filas con posicione s, o los elementos del mapa `expr` en varias filas y columnas con posiciones. A menos que se especifique lo contrario, utiliza el nombre de columna `pos` para la posición, `col` para los elementos de la matriz o `key` y `value` para los elementos del mapa.
posexplode_outer(e xpr)	Separa los elementos de la matriz `expr` en varias filas con posicione s, o los elementos del mapa `expr` en varias filas y columnas con posiciones. A menos que se especifique lo contrario, utiliza el nombre

Función	Description (Descripción)
	de columna `pos` para la posición, `col` para los elementos de la matriz o `key` y `value` para los elementos del mapa.
stack(n, expr1, ..., exprk)	Separa `expr1`,..., `expr` en `n` filas. Utiliza los nombres de columna col0, col1, etc. de forma predeterminada a menos que se especifique lo contrario.

Ejemplos

```
-- 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|
+----+----+
```

Cláusula SELECT

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

OpenSearch SQL admite una `SELECT` sentencia que se utiliza para recuperar conjuntos de resultados de una o más tablas. En la siguiente sección se describe la sintaxis general de la consulta y las diferentes estructuras de una consulta.

Sintaxis

```
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 } ]
```

Mientras que `select_statement` se define como:

```
SELECT [ ALL | DISTINCT ] { [ [ named_expression ] [ , ... ] ] }
FROM { from_item [ , ... ] }
[ PIVOT clause ]
[ UNPIVOT clause ]
[ LATERAL VIEW clause ] [ ... ]
[ WHERE boolean_expression ]
[ GROUP BY expression [ , ... ] ]
[ HAVING boolean_expression ]
```

Parámetros

- TODAS

  Selecciona todas las filas coincidentes de la relación y está habilitada de forma predeterminada.
- DISTINTO

  Selecciona todas las filas coincidentes de la relación después de eliminar los duplicados de los resultados.

- named_expression

  Una expresión con un nombre asignado. En general, denota una expresión de columna.

  Sintaxis: `expression [[AS] alias]`

- from_item

  Relación de tabla

  Relación de unión

  Relación de tablas dinámicas

  Relación no dinámica

  Funciones con valores de tabla

  Tabla en línea

  `[ LATERAL ] ( Subquery )`

- PIVOTE

  La cláusula `PIVOT` se utiliza desde la perspectiva de los datos. Puede obtener los valores agregados en función de un valor de columna específico.

- DESPIVOTAR

  La cláusula `UNPIVOT` transforma las columnas en filas. Es lo contrario a `PIVOT`, excepto en lo que respecta a la agregación de valores.

- LATERAL VIEW

  La cláusula `LATERAL VIEW` se usa junto con funciones generadoras, por ejemplo `EXPLODE`, que generarán una tabla virtual que contiene una o más filas.

  `LATERAL VIEW` aplicará las filas a cada fila de salida original.

- WHERE

  Filtra el resultado de la cláusula `FROM` en función de los predicados proporcionados.

- AGRUPAR POR

  Especifica las expresiones que se utilizan para agrupar las filas.

Se utiliza junto con las funciones de agregado (`MIN`, `MAX`, `COUNT`, `SUM`, `AVG`, etc.) para agrupar las filas en función de las expresiones de agrupación y los valores agregados de cada grupo.

Cuando se adjunta una cláusula `FILTER` a una función agregada, solo las filas coincidentes se pasan a esa función.

- TENIENDO

  Especifica los predicados mediante los que se filtran las filas producidas por `GROUP BY`.

  La cláusula `HAVING` se utiliza para filtrar las filas una vez realizada la agrupación.

  Si `HAVING` se especifica sin `GROUP BY`, indica que `GROUP BY` no tiene expresiones de agrupamiento (agregado global).

- ORDER BY

  Especifica el orden de las filas del conjunto completo de resultados de la consulta.

  Las filas de salida se ordenan en las particiones.

  Este parámetro se excluye mutuamente con `SORT BY` y `DISTRIBUTE BY` no se puede especificar de forma conjunta.

- SORT BY

  Especifica el orden por el que se ordenan las filas dentro de cada partición.

  Este parámetro se excluye mutuamente con `ORDER BY` y no se puede especificar de forma conjunta.

- LIMIT

  Especifica el número máximo de filas que puede devolver una sentencia o subconsulta.

  Esta cláusula se usa principalmente junto con `ORDER BY` para producir un resultado determinista.

- boolean_expression

  Especifica cualquier expresión que se evalúe como un valor booleano.

  Se pueden combinar dos o más expresiones mediante los operadores lógicos (`AND`, `OR`).

- expresión

Combinación de uno o varios valores, operadores o funciones SQL que dan como resultado un valor.

- named_window

  Especifica los alias de una o más especificaciones de la ventana de origen.

  Se puede hacer referencia a las especificaciones de la ventana de origen en las definiciones de ventana de la consulta.

Cláusula WHERE

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consultethe section called “Comandos SQL compatibles”.

La cláusula `WHERE` se utiliza para limitar los resultados de la cláusula `FROM` de una consulta o subconsulta en función de la condición especificada.

Sintaxis

```
WHERE boolean_expression
```

Parámetros

- boolean_expression

  Especifica cualquier expresión que se evalúe como un valor booleano.

  Se pueden combinar dos o más expresiones mediante los operadores lógicos (`AND`, `OR`).

Ejemplos

```
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|
+---+----+----+
```

Cláusula GROUP BY

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

La cláusula `GROUP BY` se utiliza para agrupar las filas en función de un conjunto de expresiones de agrupamiento especificadas y calcular las agregaciones en el grupo de filas en función de una o más funciones de agregación especificadas.

El sistema también realiza múltiples agregaciones para el mismo conjunto de registros de entrada mediante las cláusulas `GROUPING SETS`, `CUBE`, `ROLLUP`. Las expresiones de agrupamiento y las agregaciones avanzadas se pueden mezclar en la cláusula `GROUP BY` y anidar en una cláusula `GROUPING SETS` . Consulte la sección `Mixed/Nested Grouping Analytics` para obtener más información.

Cuando se adjunta una cláusula `FILTER` a una función agregada, solo las filas coincidentes se pasan a esa función.

Sintaxis

```
GROUP BY group_expression [ , group_expression [ , ... ] ] [ WITH { ROLLUP | CUBE } ]
GROUP BY { group_expression | { ROLLUP | CUBE | GROUPING SETS } (grouping_set
 [ , ...]) } [ , ... ]
```

Mientras que las funciones agregadas se definen de la siguiente manera:

```
aggregate_name ( [ DISTINCT ] expression [ , ... ] ) [ FILTER ( WHERE
 boolean_expression ) ]
```

Parámetros

- group_expression

  Especifica los criterios en función de los cuales se agrupan las filas. La agrupación de filas se realiza en función de los valores de resultado de las expresiones de agrupación.

  Una expresión de agrupamiento puede ser un nombre de columna como `GROUP BY a`, una posición de columna como `GROUP BY 0` o una expresión similar como `GROUP BY a + b`.
- grouping_set

  Un conjunto de agrupamiento se especifica mediante cero o más expresiones separadas por comas entre paréntesis. Cuando el conjunto de agrupamiento solo tiene un elemento, se pueden omitir los paréntesis.

  Por ejemplo, `GROUPING SETS ((a), (b))` es igual que `GROUPING SETS (a, b)`.

  Sintaxis: `{ ( [ expression [ , ... ] ] ) | expression }`
- GROUPING SETS

  Agrupa las filas de cada conjunto de agrupamiento que se especifique después de `GROUPING SETS`.

  Por ejemplo, `GROUP BY GROUPING SETS ((warehouse), (product))` es semánticamente equivalente a la unión de los resultados de `GROUP BY warehouse` y `GROUP BY product`. Esta cláusula es una forma abreviada de UNION ALL, en la que cada parte del operador `UNION ALL` realiza la agregación de cada conjunto de agrupamiento especificado en la cláusula `GROUPING SETS`.

Del mismo modo, `GROUP BY GROUPING SETS ((warehouse, product), (product), ())` es semánticamente equivalente a la unión de los resultados de `GROUP BY warehouse, product`, `GROUP BY product` y un agregado global.

- ROLLUP

  Especifica varios niveles de agregación en una sentencia. Esta cláusula se utiliza para calcular las agregaciones en función de varios conjuntos de agrupamiento. `ROLLUP` es una forma abreviada de `GROUPING SETS`.

  Por ejemplo, `GROUP BY warehouse, product WITH ROLLUP` or `GROUP BY ROLLUP(warehouse, product)` equivale a `GROUP BY GROUPING SETS((warehouse, product), (warehouse), ())`.

  `GROUP BY ROLLUP(warehouse, product, (warehouse, location))` es equivalente a `GROUP BY GROUPING SETS((warehouse, product, location), (warehouse, product), (warehouse), ())`.

  Los elementos N de una especificación ROLLUP dan como resultado N+1 GROUPING SETS.

- CUBE

  La cláusula CUBE se utiliza para realizar agregaciones basadas en la combinación de columnas de agrupamiento especificadas en la cláusula GROUP BY. CUBE es una forma abreviada de GROUPING SETS.

  Por ejemplo, `GROUP BY warehouse, product WITH CUBE` or `GROUP BY CUBE(warehouse, product)` equivale a `GROUP BY GROUPING SETS((warehouse, product), (warehouse), (product), ())`.

  `GROUP BY CUBE(warehouse, product, (warehouse, location))` es equivalente a `GROUP BY GROUPING SETS((warehouse, product, location), (warehouse, product), (warehouse, location), (product, warehouse, location), (warehouse), (product), (warehouse, product), ())`. Los elementos N de una especificación `CUBE` dan como resultado 2^N `GROUPING SETS`.

- Análisis de agrupamiento mixto/anidado

  Una cláusula `GROUP BY` puede incluir múltiples group_expressions y múltiples `CUBE|ROLLUP|GROUPING SETS`. `GROUPING SETS` también puede tener cláusulas `CUBE|ROLLUP|GROUPING SETS` anidadas, como `GROUPING SETS(ROLLUP(warehouse,`

location), CUBE(warehouse, location)), GROUPING SETS(warehouse, GROUPING SETS(location, GROUPING SETS(ROLLUP(warehouse, location),, CUBE(warehouse, location)))).

CUBE | ROLLUP es solo azúcar de sintáctico para GROUPING SETS. Consulte las secciones anteriores para saber cómo traducir CUBE | ROLLUP a GROUPING SETS. group_expression se puede tratar como un solo grupo GROUPING SETS en este contexto.

En el caso de los múltiplos de GROUPING SETS en la cláusula GROUP BY, generamos un único GROUPING SETS haciendo un producto cruzado del original GROUPING SETS. Para anidar GROUPING SETS en la cláusula GROUPING SETS, simplemente cogemos sus conjuntos de agrupamiento y los eliminamos.

Por ejemplo, GROUP BY warehouse, GROUPING SETS((product), ()), GROUPING SETS((location, size), (location), (size), ()) and GROUP BY warehouse, ROLLUP(product), CUBE(location, size) equivale 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))) es equivalente a GROUP BY GROUPING SETS((warehouse), (warehouse, product)).

- aggregate_name

  Especifica el nombre de una función agregada (MIN, MAX, COUNT, SUM, AVG, etc.).

- DISTINTO

  Elimina los duplicados de las filas de entrada antes de pasarlos a las funciones de agregado.

- FILTRAR

  Filtra las filas de entrada para las que el valor boolean_expression de la cláusula WHERE se evalúe como verdadero se pasan a la función de agregado; las demás filas se descartan.

Ejemplos

```
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     |
+-------------------+------------------+----------+
```

Cláusula HAVING

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consultethe section called “Comandos SQL compatibles”.

La cláusula HAVING se utiliza para filtrar los resultados generados por GROUP BY en función de la condición especificada. Suele utilizarse junto con una cláusula GROUP BY.

Sintaxis

```
HAVING boolean_expression
```

Parámetros

- boolean_expression

  Especifica cualquier expresión que se evalúe como un valor booleano. Se pueden combinar dos o más expresiones mediante los operadores lógicos (AND, OR).

  Nota: Las expresiones especificadas en la cláusula HAVING solo pueden hacer referencia a:

  1. Constantes
  2. Expresiones que aparecen en GROUP BY
  3. Funciones de agregación

Ejemplos

```
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|
+---+
```

Cláusula ORDER BY

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

La cláusula `ORDER BY` se utiliza para devolver las filas de resultados ordenadas en el orden especificado por el usuario. A diferencia de la cláusula SORT BY, esta cláusula garantiza un orden total en la salida.

Sintaxis

```
ORDER BY { expression [ sort_direction | nulls_sort_order ] [ , ... ] }
```

Parámetros

- ORDER BY

  Especifica una lista de expresiones separadas por comas junto con parámetros `sort_direction` y `nulls_sort_order` opcionales que se utilizan para ordenar las filas.

- sort_direction

  De manera opcional, especifica si se deben ordenar en orden ascendente o descendente.

  Los valores válidos para la dirección de ordenación son `ASC` ascendentes y `DESC` descendentes.

  Si la dirección de ordenación no se especifica de forma explícita, las filas se ordenan por defecto en orden ascendente.

  Sintaxis: `[ ASC | DESC ]`

- nulls_sort_order

  Si lo desea, especifica si `NULL` los valores se devuelven como valores before/after no nulos.

  Si no se especifica null_sort_order, `NULLs` ordena primero si el orden es `ASC` y NULLS ordenará en último lugar si el orden de clasificación es `DESC`.

1. Si `NULLS FIRST` se especifica, los valores NULL se devuelven primero, independientemente del orden de clasificación.

2. Si `NULLS LAST` se especifica, los valores NULL se devuelven en último lugar, independientemente del orden de clasificación.

Sintaxis: `[ NULLS { FIRST | LAST } ]`

Ejemplos

```
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|
+---+-----+----+
```

Cláusula JOIN

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Una unión SQL se utiliza para combinar filas de dos relaciones en función de los criterios de unión. En la siguiente sección se describe la sintaxis general de las uniones y los distintos tipos de uniones, junto con ejemplos.

Sintaxis

```
relation INNER JOIN relation [ join_criteria ]
```

Parámetros

- relation

  Especifica la relación que se va a unir.
- join_type

  Especifica el tipo de unión.

  Sintaxis: `INNER | CROSS | LEFT OUTER`
- join_criteria

  Especifica cómo se combinarán las filas de una relación con las filas de otra relación.

  Sintaxis: `ON boolean_expression | USING ( column_name [ , ... ] )`
- boolean_expression

  Especifica una expresión con un tipo de devolución booleano.

Tipos de unión

- Unión interna

La combinación interna debe especificarse de forma explícita. Selecciona filas que tienen valores coincidentes en ambas relaciones.

Sintaxis: `relation INNER JOIN relation [ join_criteria ]`

- Combinación izquierda

  Una combinación por la izquierda devuelve todos los valores de la relación izquierda y los valores coincidentes de la relación derecha, o agrega NULL si no hay ninguna coincidencia. También se conoce como unión externa izquierda.

  Sintaxis: `relation LEFT OUTER JOIN relation [ join_criteria ]`

- Unión cruzada

  Una unión cruzada devuelve el producto cartesiano de dos relaciones.

  Sintaxis: `relation CROSS JOIN relation [ join_criteria ]`

Ejemplos

```
-- 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|
+---+-----+------+-----------+
```

Cláusula LIMIT

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consultethe section called “Comandos SQL compatibles”.

La cláusula `LIMIT` se utiliza para restringir el número de filas devueltas por la sentencia `SELECT`. En general, esta cláusula se usa junto con `ORDER BY` para garantizar que los resultados sean deterministas.

Sintaxis

```
LIMIT { ALL | integer_expression }
```

Parámetros

- TODAS

  Si se especifica, la consulta devuelve todas las filas. En otras palabras, no se aplica ningún límite si se especifica esta opción.
- integer_expression

  Especifica una expresión plegable que devuelve un entero.

Ejemplos

```
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|
+-------+---+
```

Cláusula CASE

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consultethe section called “Comandos SQL compatibles”.

La cláusula CASE utiliza una regla para devolver un resultado específico en función de la condición especificada, de forma similar a las sentencias if/else de otros lenguajes de programación.

Sintaxis

```
CASE [ expression ] { WHEN boolean_expression THEN then_expression } [ ... ]
[ ELSE else_expression ]
END
```

Parámetros

- boolean_expression

  Especifica cualquier expresión que se evalúe como un valor booleano.

  Se pueden combinar dos o más expresiones mediante los operadores lógicos (AND, OR).
- then_expression

  Especifica la expresión then en función de la condición boolean_expression.

  `then_expression` y `else_expression` todas deben ser del mismo tipo o estar sujetas a un tipo común.
- else_expression

  Especifica la expresión predeterminada.

  `then_expression` y `else_expression` todas deben ser del mismo tipo o estar sujetas a un tipo común.

Ejemplos

```
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
                |
+------
+-------------------------------------------------------------------------------------------------------
+
```

Expresiones de tabla comunes

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Una expresión de tabla común (CTE) define un conjunto de resultados temporal al que un usuario puede hacer referencia posiblemente varias veces dentro del ámbito de una sentencia SQL. Una CTE se utiliza principalmente en una declaración `SELECT`.

Sintaxis

```
WITH common_table_expression [ , ... ]
```

Mientras que `common_table_expression` se define como:

```
Syntexpression_name [ ( column_name [ , ... ] ) ] [ AS ] ( query )
```

Parámetros

- expression_name

  Especifica un nombre para la expresión de tabla común.
- consulta

  Una sentencia `SELECT`.

Ejemplos

```
-- 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**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

La sentencia EXPLAIN se utiliza para proporcionar planes lógicos/físicos para una sentencia de entrada. De forma predeterminada, esta cláusula proporciona información únicamente sobre un plan físico.

Sintaxis

```
EXPLAIN [ EXTENDED | CODEGEN | COST | FORMATTED ] statement
```

Parámetros

- EXTENDED

  Genera un plan lógico diseccionado, un plan lógico analizado, un plan lógico optimizado y un plan físico.

  El plan lógico diseccionado es un plan sin resolver que se extrajo de la consulta.

  Los planes lógicos analizados se transforman y traducen `unresolvedAttribute` y `unresolvedRelation` a objetos completamente tipificados.

  El plan lógico optimizado se transforma mediante un conjunto de reglas de optimización, lo que da como resultado el plan físico.
- CODEGEN

  Genera el código para la declaración, si la hubiera y un plan físico.
- COST

  Si las estadísticas de los nodos del plan están disponibles, genera un plan lógico y las estadísticas.
- FORMATTED

  Genera dos secciones: un esquema físico del plan y detalles de los nodos.
- statement

  Especifica una sentencia SQL que se va a explicar.

Ejemplos

```
-- 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]
|
+-----------------------------------------------------+
```

Cláusula LATERAL SUBQUERY

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

`LATERAL SUBQUERY` es una subconsulta precedida de la palabra clave `LATERAL`. Proporciona una forma de hacer referencia a las columnas de la cláusula `FROM` anterior. Sin la palabra clave `LATERAL`, las subconsultas solo pueden hacer referencia a las columnas de la consulta externa, pero no a las de la cláusula `FROM`. `LATERAL SUBQUERY` simplifica y hace que las consultas complicadas sean más sencillas y eficaces.

Sintaxis

```
[ LATERAL ] primary_relation [ join_relation ]
```

Parámetros

- primary_relation

  Especifica la relación principal. Puede ser una de las siguientes opciones:

  1. Relación de tabla
  2. Consulta con alias

     Sintaxis: `( query ) [ [ AS ] alias ]`
  3. Relación con alias

     `Syntax: ( relation ) [ [ AS ] alias ]`

Ejemplos

```
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   |
+-------+--------+--------+
```

Cláusula LATERAL VIEW

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

La cláusula `LATERAL VIEW` se usa junto con funciones generadoras, por ejemplo `EXPLODE`, que generarán una tabla virtual que contiene una o más filas. `LATERAL VIEW` aplicará las filas a cada fila de salida original.

Sintaxis

```
LATERAL VIEW [ OUTER ] generator_function ( expression [ , ... ] ) [ table_alias ] AS
 column_alias [ , ... ]
```

Parámetros

- EXTERIOR

  Si `OUTER` se especifica, devuelve nulo si una entrada array/map está vacía o nula.
- generator_function

  Especifica una función generadora (`EXPLODE`, `INLINE`, etc.).
- table_alias

  El alias de `generator_function`, que es opcional.
- column_alias

  Muestra los alias de las columnas de `generator_function`, que se pueden usar en las filas de salida.

  Puede tener varios alias si `generator_function` tiene varias columnas de salida.

Ejemplos

```
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   |
+------+-------+-------+--------+-----------+--------+
```

Predicado LIKE

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consultethe section called “Comandos SQL compatibles”.

Se utiliza un predicado `LIKE` para buscar un patrón específico. Este predicado también admite varios patrones con cuantificadores como `ANY`, `SOME` y `ALL`.

Sintaxis

```
[ NOT ] { LIKE search_pattern [ ESCAPE esc_char ] | [ RLIKE | REGEXP ] regex_pattern }
[ NOT ] { LIKE quantifiers ( search_pattern [ , ... ]) }
```

Parámetros

- search_pattern

  Especifica un patrón de cadenas que se va a buscar mediante la cláusula LIKE. Puede contener caracteres especiales que coincidan con el patrón:

  - `%` coincide con 0 o más caracteres.
  - `_` coincide exactamente con un carácter.
- esc_char

  Especifica el carácter de escape. El carácter de escape predeterminado es `\`.
- regex_pattern

  Especifica un patrón de búsqueda de expresiones regulares que se va a buscar mediante la cláusula `RLIKE` o `REGEXP`.

- quantifiers

  Especifica los cuantificadores de predicados que incluyen ANY, SOME y ALL.

  ANY o SOME significan que si uno de los patrones coincide con la entrada, devuelve verdadero.

  ALL significa que si todos los patrones coinciden con la entrada, entonces devuelve verdadero.

Ejemplos

```
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**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

La cláusula `OFFSET` se utiliza para especificar el número de filas que se van a omitir antes de empezar a devolver las filas devueltas por la sentencia `SELECT`. En general, esta cláusula se utiliza en conjunto con `ORDER BY` para garantizar que los resultados sean deterministas.

Sintaxis

```
OFFSET integer_expression
```

Parámetros

- integer_expression

  Especifica una expresión plegable que devuelve un entero.

Ejemplos

```
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|
```

```
+-------+---+
```

Cláusula PIVOT

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

La cláusula `PIVOT` se utiliza desde la perspectiva de los datos. Podemos obtener los valores agregados en función de valores de columna específicos, que se convertirán en varias columnas utilizadas en la cláusula `SELECT`. La cláusula `PIVOT` se puede especificar después del nombre de la tabla o la subconsulta.

Sintaxis

```
PIVOT ( { aggregate_expression [ AS aggregate_expression_alias ] } [ , ... ] FOR
 column_list IN ( expression_list ) )
```

Parámetros

- expresión_de_agregación

  Especifica una expresión agregada (`SUM(a)`, `COUNT(DISTINCT b)`, etc.

- aggregate_expression_alias

  Especifica un alias para la expresión agregada.

- column_list

  Contiene columnas en la cláusula `FROM`, que especifica las columnas que desea reemplazar por columnas nuevas. Puede utilizar corchetes para rodear las columnas, por ejemplo `(c1, c2)`.

- expression_list

  Especifica las columnas nuevas, que se utilizan para hacer coincidir los valores en `column_list` como la condición de agregación. También puede usar alias para los bots.

Ejemplos

```
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  |
+------+-----------+-------+-------+-------+-------+
```

Operadores de establecimiento

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

Los operadores de conjuntos se utilizan para combinar dos relaciones de entrada en una sola. OpenSearch SQL admite tres tipos de operadores de conjuntos:

- `EXCEPT` o `MINUS`
- `INTERSECT`
- `UNION`

Las relaciones de entrada deben tener el mismo número de columnas y tipos de datos compatibles para las columnas respectivas.

EXCEPTO

`EXCEPT` y `EXCEPT ALL` devuelven las filas que se encuentran en una relación pero no en la otra. `EXCEPT` (alternativamente, `EXCEPT DISTINCT`) solo toma filas distintas, pero `EXCEPT ALL` no elimina los duplicados de las filas de resultados. Tenga en cuenta que `MINUS` es un alias de `EXCEPT`.

Sintaxis

```
[ ( ] relation [ ) ] EXCEPT | MINUS [ ALL | DISTINCT ] [ ( ] relation [ ) ]
```

Ejemplos

```
-- 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|
+---+
```

INTERSECTAR

`INTERSECT` y `INTERSECT ALL` devuelven las filas que se encuentran en ambas relaciones. `INTERSECT` (alternativamente, `INTERSECT DISTINCT`) solo toma filas distintas, pero `INTERSECT ALL` no elimina los duplicados de las filas de resultados.

Sintaxis

```
[ ( ] relation [ ) ] INTERSECT [ ALL | DISTINCT ] [ ( ] relation [ ) ]
```

Ejemplos

```
(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|
+---+
```

UNIÓN

`UNION` y `UNION ALL` devuelven las filas que se encuentran en cualquiera de las dos relaciones. `UNION` (alternativamente, `UNION DISTINCT`) solo toma filas distintas, pero `UNION ALL` no elimina los duplicados de las filas de resultados.

Sintaxis

```
 [ ( ] relation [ ) ] UNION [ ALL | DISTINCT ] [ ( ] relation [ ) ]
```

Ejemplos

```
(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|
+---+
```

Cláusula SORT BY

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consultethe section called “Comandos SQL compatibles”.

La cláusula `SORT BY` se utiliza para devolver las filas de resultados ordenadas dentro de cada partición en el orden especificado por el usuario. Cuando hay más de una partición, `SORT BY` puede devolver un resultado parcialmente ordenado. Esto es diferente a la cláusula `ORDER BY` que garantiza un orden total de la salida.

Sintaxis

```
SORT BY { expression [ sort_direction | nulls_sort_order ] [ , ... ] }
```

Parámetros

- SORT BY

  Especifica una lista de expresiones separadas por comas junto con los parámetros opcionales sort_direction y nulls_sort_order que se utilizan para ordenar las filas de cada partición.
- sort_direction

  De manera opcional, especifica si se deben ordenar en orden ascendente o descendente.

  Los valores válidos para la dirección de ordenación son `ASC` ascendentes y `DESC` descendentes.

  Si la dirección de ordenación no se especifica de forma explícita, las filas se ordenan por defecto en orden ascendente.

  Sintaxis: `[ ASC | DESC ]`
- nulls_sort_order

  De manera opcional, especifica si los valores NULL se devuelven antes o después de los valores no NULL.

  Si no `null_sort_order` se especifica, NULLs ordene primero si el orden de clasificación es `ASC` y NULL ordene en último lugar si el orden de clasificación sí lo es. `DESC`

  1. Si `NULLS FIRST` se especifica, los valores NULL se devuelven primero, independientemente del orden de clasificación.

  2. Si `NULLS LAST` se especifica, los valores NULL se devuelven en último lugar, independientemente del orden de clasificación.

  Sintaxis: `[ NULLS { FIRST | LAST } ]`

Ejemplos

```
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**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando SQL, consulte. the section called “Comandos SQL compatibles”

La cláusula `UNPIVOT` transforma varias columnas en varias filas utilizadas en la cláusula `SELECT`. La cláusula `UNPIVOT` se puede especificar después del nombre de la tabla o la subconsulta.

Sintaxis

```
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] [, ...])
```

Parámetros

- unpivot_column

  Contiene las columnas de la `FROM` cláusula, que especifican las columnas que queremos dejar de pivotar.
- name_column

El nombre de la columna que contiene los nombres de las columnas no pivotantes.

- values_column

  El nombre de la columna que contiene los valores de las columnas no pivotantes.

Ejemplos

```
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           |
+------+------------------+---------------+----------------+
```

Comandos compatibles PPL

Las siguientes tablas muestran qué comandos de PPL admite OpenSearch Dashboards para consultar CloudWatch Logs, Amazon S3 o Security Lake, y qué comandos admite CloudWatch Logs Insights. CloudWatch Logs Insights utiliza la misma sintaxis de PPL que los OpenSearch paneles cuando consulta los CloudWatch registros, y en las tablas se hace referencia a ambos como registros. CloudWatch

> **Note**
>
> Al analizar datos fuera del OpenSearch Servicio, es posible que los comandos se ejecuten de forma diferente a como lo hacen en los índices. OpenSearch

Temas

- Comandos
- Funciones
- Información adicional para CloudWatch los usuarios de Logs Insights que utilizan OpenSearch PPL

Comandos

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “campos”	Muestra un conjunto de campos que necesitan ser proyectados.	compatible	compatible	compatible	`fields field1, field2`
the section called “donde”	Filtra los datos en función de las condiciones que especifiq ue.	compatible	compatible	compatible	`where field1="success" | where field2 != "i-023fe0a90929d8822" | fields field3, col4, col5, col6 | head 1000`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “stats”	Realiza agregaciones y cálculos.	compatible	compatible	compatible	`stats count(), count(`field1`), min(`field1`), max(`field1`), avg(`field1`) by field2 | head 1000`
the section called “parse”	Extrae un patrón de expresión regular (regex) de una cadena y muestra el patrón extraído. El patrón extraído se puede utilizar además para crear nuevos campos o filtrar datos.	compatible	compatible	compatible	`parse `field1` ".*/(?<field2>[^/]+$)" | where field2 = "requestId" | fields field2, `field2` | head 1000`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “patrones”	Extrae los patrones de registro de un campo de texto y agrega los resultados al resultado de la búsqueda. Al agrupar los registros por sus patrones, resulta más fácil agregar estadísticas de grandes volúmenes de datos de registro para analizarlos y solucionar problemas.	no compatible	compatible	compatible	`patterns new_field='no_numbers' pattern='[0-9]' message | fields message, no_numbers`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “sort”	Ordena los resultados mostrados por un nombre de campo. Use ordenar: FieldName para ordenar en orden descendente.	compatible	compatible	compatible	`stats count(), count(`field1`), min(`field1`) as field1Alias, max(`field1`), avg(`field1`) by field2 | sort - field1Alias | head 1000`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “eval”	Modifica o procesa el valor de un campo y lo almacena en un campo diferente. Esto resulta útil para modificar matemáticamente una columna, aplicar funciones de cadena a una columna o aplicar funciones de fecha a una columna.	compatible	compatible	compatible	`eval field2 = `field1` * 2 | fields field1, field2 | head 20`
the section called “rename”	Cambia el nombre de uno o más campos del resultado de la búsqueda.	compatible	compatible	compatible	`rename field2 as field1 | fields field1`
the section called “head”	Limita los resultados de la consulta a las N primeras filas.	compatible	compatible	compatible	`fields `@message` | head 20`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “grok”	Analiza un campo de texto con un patrón grok basado en una expresión regular y agrega los resultados al resultado de la búsqueda.	compatible	compatible	compatible	`grok email '.+@%{HOSTNAME:host}' | fields email`
the section called “top”	Busca los valores más frecuentes de un campo.	compatible	compatible	compatible	`top 2 Field1 by Field2`
the section called “dedup”	Elimina entradas duplicadas en función de los campos que especifique.	compatible	compatible	compatible	`dedup field1 | fields field1, field2, field3`
the section called “unirse”	Une dos conjuntos de datos.	compatible	compatible	compatible	`source=customer | join ON c_custkey = o_custkey orders | head 10`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “lookup”	Enriquece los datos de búsqueda mediante la adición o sustitución de datos de un índice de búsqueda (tabla de dimensiones). Puede ampliar los campos de un índice con valores de una tabla de dimensiones, agregar o reemplazar valores cuando se cumpla la condición de búsqueda.	no compatible	compatible	compatible	`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 (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “subquery”	Realiza consultas anidadas y complejas en sus instrucciones de lenguaje de procesamiento de canalizaciones (PPL).	compatible	compatible	compatible	`where id in [ subquery source=users | where user in [ subquery source=actions | where action="login" | fields user ] | fields uid ]`
the section called “rare”	Busca los valores menos frecuentes de todos los campos de la lista de campos.	compatible	compatible	compatible	`rare Field1 by Field2`
the section called “trendline”	Calcula los promedios móviles de los campos.	compatible	compatible	compatible	`trendline sma(2, field1) as field1Alias`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “eventstats”	Enriquece los datos del evento con estadísticas resumidas calculadas. Analiza los campos específicos de sus eventos, calcula varias medidas estadísticas y, a continuación, agrega estos resultados a cada evento original como campos nuevos.	compatible (excepto `count()`)	compatible	compatible	`eventstats sum(field1) by field2`
the section called “flatten”	Aplana un campo. El campo debe ser de este tipo: `struct<?,?> or array<struct<?,?>>`	compatible	compatible	compatible	`source=table | flatten field1`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “fieldsummary”	Calcula las estadísticas básicas de cada campo (recuento, recuento distinto, mínimo, máximo, promedio, stddev y media).	Compatible (un campo por consulta)	compatible	compatible	`where field1 != 200 | fieldsummary includefields=field1 nulls=true`
the section called “fillnull”	Rellena los campos nulos con el valor que proporcione. Se puede usar en uno o más campos.	compatible	compatible	compatible	`fields field1 | eval field2=field1 | fillnull value=0 field1`
the section called “Expandir”	Divide un campo que contiene varios valores en filas independientes, creando una nueva fila para cada valor del campo especificado.	compatible	compatible	compatible	`expand employee | stats max(salary) as max by state, company`

Comando PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “describe”	Obtiene información detallada sobre la estructura y los metadatos de las tablas, los esquemas y los catálogos	no compatible	compatible	compatible	`describe schema.table`

Funciones

Función PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Cadena” (`CONCAT`, `CONCAT_WS` , `LENGTH`, `LOWER`, `LTRIM`, `POSITION`, `REVERSE`, `RIGHT`, `RTRIM`, `SUBSTRING` , `TRIM`, `UPPER`)	Funciones integradas en PPL que pueden manipular y transformar datos de texto y cadenas de texto dentro de las consultas de PPL. Por	compatible	compatible	compatible	`eval col1Len = LENGTH(col1) | fields col1Len`

Función PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
	ejemplo, convertir mayúsculas y minúsculas, combinar cadenas, extraer partes y limpiar el texto.				

Función PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called "Fecha y hora" (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, WEEKDAY,	Funciones integradas para administrar y transformar los datos de fecha y hora en las consultas PPL. Por ejemplo, date_add, date_format, datediff y current_date.	compatible	compatible	compatible	`eval newDate = ADDDATE(DATE('2020-08-26'), 1) | fields newDate`

Función PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
`WEEK_OF_YEAR` , `DATE_ADD`, `DATE_SUB`, `TIMESTAMPADD` , `TIMESTAMPDIFF` , `UTC_TIMESTAMP` , `CURRENT_TIMEZONE`)					
the section called “Condición” (`EXISTS`, `IF`, `IFNULL`, `ISNOTNULL` , `ISNULL`, `NULLIF`)	Funciones integradas que realizan cálculos en varias filas para generar un único valor resumido. Por ejemplo, sum, count, avg, max y min.	compatible	compatible	compatible	`eval field2 = isnull(col1) | fields field2, col1, field3`

Función PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Matemático” (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)	Funciones integradas para realizar cálculos matemáticos y transformaciones en consultas PPL. Por ejemplo: abs (valor absoluto), round (redondea números), sqrt (raíz cuadrada), pow (cálculo de potencia) y ceil (redondea al entero más cercano).	compatible	compatible	compatible	`eval field2 = ACOS(col1) | fields col1`

Función PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Expressions” (Operadores aritméticos (+, -, *), operadores de predicados (>. <, IN))	Las funciones integradas para las expresiones, en particular las expresiones de valores, devuelven un valor escalar. Las expresiones tienen distintos tipos y formas.	compatible	compatible	compatible	`where age > (25 + 5) | fields age`
the section called “Dirección IP” (CIDRMATCH)	Funciones integradas para administrar direcciones IP, como el CIDR.	compatible	compatible	compatible	`where cidrmatch (ip, '********** ***/24') | fields ip`

Función PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called "JSON" (ARRAY_LENGTH , ARRAY_LENGTH , JSON, JSON_ARRAY , JSON_EXTRACT , JSON_KEYS , JSON_OBJECT , JSON_VALID , TO_JSON_STRING)	Funciones integradas para el manejo de JSON, incluidas las matrices, la extracción y la validación.	compatible	compatible	compatible	`eval `json_extract('{"a":"b"}', '$.a')` = json_extract('{"a":"b"}', '$a')`
the section called "Lambda" (EXISTS, FILTER, REDUCE, TRANSFORM)	Funciones integradas para el manejo de JSON, incluidas las matrices, la extracción y la validación.	no compatible	compatible	compatible	`eval array = json_array(1, -1, 2), result = filter(array, x -> x > 0) | fields result`

Función PPL	Description (Descripción)	CloudWatch Registros	Amazon S3	Security Lake	Comando de ejemplo:
the section called “Criptográfica” (MD5, SHA1, SHA2)	Funciones integradas que le permiten generar huellas dactilares únicas de los datos, que se pueden utilizar para la verificación, la comparación o como parte de protocolos de seguridad más complejos.	compatible	compatible	compatible	`eval` `MD5('hello')` = `MD5('hello')` | `fields` `MD5('hello')`

Información adicional para CloudWatch los usuarios de Logs Insights que utilizan OpenSearch PPL

Aunque CloudWatch Logs Insights admite la mayoría de los comandos y funciones de OpenSearch PPL, algunos comandos y funciones no lo son actualmente. Por ejemplo, actualmente no admite

los comandos de búsqueda en PPL. A partir del 2 de junio de 2025, CloudWatch Logs Insights ahora admite las funciones JOIN, subqueries, Flatten, Fillnull, Expand, Cidrmatch y JSON en PPL. Para obtener una lista completa de los comandos y funciones de consulta compatibles, consulta las columnas de Amazon CloudWatch Logs en las tablas anteriores.

Consultas y cuotas de ejemplo

Lo siguiente se aplica tanto a los usuarios de CloudWatch Logs Insights como a los OpenSearch usuarios que consultan CloudWatch datos.

Para obtener información sobre los límites que se aplican al consultar CloudWatch los registros del OpenSearch servicio, consulte las cuotas de CloudWatch registros en la Guía del usuario de Amazon CloudWatch Logs. Los límites se refieren al número de grupos de CloudWatch registros que se pueden consultar, el número máximo de consultas simultáneas que se pueden ejecutar, el tiempo máximo de ejecución de las consultas y el número máximo de filas devueltas en los resultados. Los límites son los mismos independientemente del idioma que utilice para consultar CloudWatch los registros (es decir, OpenSearch PPL, SQL y Logs Insights QL).

Comandos PPL

Temas

- comentario
- Comando de correlación
- dedup command
- 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 rare
- Comando rename
- Comando search
- Comando sort
- Comandos stats
- Comando subquery
- Comando top
- Comando trendline
- Comando where
- Resumen de campo
- Comando expand
- Funciones PPL

comentario

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

PPL admite tanto los comentarios de línea como los comentarios bloqueados. El sistema no evalúa el texto de los comentarios.

Comentarios de línea

Los comentarios de línea comienzan con dos barras diagonales // y terminan con una línea nueva.

Ejemplo:

```
os> source=accounts | top gender // finds most common gender of all the accounts
fetched rows / total rows = 2/2
```

```
+----------+
| gender   |
|----------|
| M        |
| F        |
+----------+
```

Comentarios bloqueados

Los comentarios bloqueados comienzan con una barra seguida de un asterisco * y terminan con un asterisco seguido de una barra */.

Ejemplo:

```
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 de correlación

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Puede correlacionar diferentes orígenes de datos según dimensiones y plazos comunes.

Esta correlación es crucial cuando se trata de grandes cantidades de datos de varios mercados verticales que comparten los mismos periodos de tiempo pero que no están sincronizados formalmente.

Al correlacionar estos diferentes orígenes de datos en función de plazos y dimensiones similares, puede enriquecer sus datos y descubrir información valiosa.

Ejemplo

El dominio de la observabilidad tiene tres orígenes de datos distintos:

- Registros
- Métricas
- Rastros

Estas orígenes de datos pueden compartir dimensiones comunes. Para realizar la transición de un origen de datos a otro, debe correlacionarlos correctamente. Mediante las convenciones de nomenclatura semántica, puede identificar los elementos compartidos en los registros, los rastros y las métricas.

Ejemplo:

```
{
  "@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"
}
```

En este ejemplo, se muestra un registro de AWS ELB que proviene de un servicio que reside en. AWS Muestra una respuesta backend de HTTP con un código de estado 500, lo que indica un error. Esto podría activar una alerta o formar parte de su proceso de supervisión habitual. El siguiente paso es recopilar datos relevantes sobre este suceso para llevar a cabo una investigación exhaustiva.

Si bien es posible que se sienta tentado a consultar todos los datos relacionados con el periodo de tiempo, este enfoque puede resultar abrumador. Podría terminar con demasiada información y dedicar más tiempo a filtrar los datos irrelevantes que a identificar la raíz del problema.

En su lugar, puede utilizar un enfoque más específico mediante la correlación de los datos de diferentes orígenes. Puede utilizar estas dimensiones para la correlación:

- IP: `"ip": "10.0.0.1" | "ip": "**************"`
- Puerto: `"port": 2817 | "target_port": "10.0.0.1:80"`

Si tiene acceso a índices de métricas y trazas adicionales y está familiarizado con la estructura de su esquema, puede crear una consulta de correlación más precisa.

Este es un ejemplo de un documento de índice de rastreo que contiene información HTTP que quizás quiera correlacionar:

```
{
  "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": {
         ...
        }
      }
    }
  }
}
```

En este enfoque, puede ver los puntos `traceId` y los http client/server `ip` que se pueden correlacionar con los registros del ELB para comprender mejor el comportamiento y el estado del sistema.

Nuevo comando de consulta de correlación

Este es el nuevo comando que permitiría este tipo de investigación:

```
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 )
```

A continuación, se muestra lo que hace cada parte del comando:

1. `source alb_logs, traces`: esto selecciona los orígenes de datos que desea correlacionar.
2. `where ip="10.0.0.1" AND cloud.provider="aws"`: esto reduce el alcance de la búsqueda.
3. `correlate exact fields(traceId, ip)`: esto le indica al sistema que correlacione los datos en función de las coincidencias exactas de los siguientes campos:
   - El campo `ip` tiene una condición de filtro explícita, por lo que se utilizará en la correlación de todas los orígenes de datos.
   - El campo `traceId` no tiene un filtro explícito, por lo que coincidirá con los mismos traceId en todas las orígenes de datos.

Los nombres de los campos indican el significado lógico de la función dentro del comando de correlación. La condición de unión real depende de la sentencia de asignación que proporcione.

El término `exact` significa que las declaraciones de correlación requerirán que todos los campos coincidan para cumplir con la declaración de consulta.

El término `approximate` intentará coincidir en el mejor de los casos y no rechazará las filas con coincidencias parciales.

Tratamiento de diferentes asignaciones de campos

En los casos en que el mismo campo lógico (por ejemplo `ip`) tenga nombres diferentes en los orígenes de datos, debe proporcionar la asignación explícita de los campos de ruta. Para solucionar este problema, puede ampliar las condiciones de correlación para que coincidan con distintos nombres de campo con significados lógicos similares. A continuación le explicamos cómo puede configurar esto:

```
alb_logs.ip = traces.attributes.http.server.address, alb_logs.traceId = traces.traceId
```

Para cada campo que participe en la unión de correlación, debe proporcionar una sentencia de asignación relevante que incluya todas las tablas que se van a unir mediante este comando de correlación.

Ejemplo

En este ejemplo, hay dos orígenes: `alb_logs, traces`

Hay 2 campos: `traceId, ip`

Hay dos instrucciones de asignación: `alb_logs.ip = traces.attributes.http.server.address, alb_logs.traceId = traces.traceId`

Delimitación de los plazos de correlación

Para simplificar el trabajo realizado por el motor de ejecución (controlador), puede agregar la sentencia de alcance. Esto indica explícitamente a la consulta de unión el momento en el que debe realizarse esta búsqueda.

`scope(@timestamp, 1D)` i

En este ejemplo, el ámbito de búsqueda se centra en el día a día, por lo que las correlaciones que aparecen el mismo día se agrupan. Este mecanismo de determinación del alcance simplifica y permite un mejor control de los resultados, lo que permite una resolución de búsqueda incremental en función de sus necesidades.

Controladores de apoyo

El nuevo comando de correlación es en realidad un comando de unión “oculto”. Por lo tanto, solo los siguientes controladores PPL admiten este comando. En estos controladores, el comando de correlación se traducirá directamente al plan lógico Catalyst Join correspondiente.

Ejemplo

```
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)
```

Plan lógico:

```
'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]
```

El motor catalizador optimiza esta consulta según el orden de unión más eficiente.

dedup command

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `dedup` para eliminar documentos idénticos de los resultados de la búsqueda en función de los campos especificados.

Sintaxis

Utilice la siguiente sintaxis:

```
dedup [int] <field-list> [keepempty=<bool>] [consecutive=<bool>]
```

`int`

- Opcional.
- El comando `dedup` conserva varios eventos para cada combinación cuando se especifica <int>. El número para <int> debe ser mayor que 0. Si no especifica un número, solo se conservara el primer evento que se produzca. Todos los demás duplicados se eliminan de los resultados.
- Valor predeterminado: 1

`keepempty`

- Opcional.
- Si es verdadero, mantiene los documentos en los que algún campo de la lista de campos tenga un valor NULL o MISSING.

- Predeterminado: false

`consecutive`

- Opcional.
- Si es verdadero, elimina solo los eventos con combinaciones de valores duplicados consecutivas.
- Predeterminado: false

`field-list`

- Obligatorio.
- Una lista de campos delimitada por comas. Se requiere como mínimo un campo de metadatos.

Ejemplo 1: Deduplicación por un campo

En este ejemplo se muestra cómo desdoblar documentos mediante el campo de género.

Consulta PPL:

```
os> source=accounts | dedup gender | fields account_number, gender;
fetched rows / total rows = 2/2
+------------------+----------+
| account_number   | gender   |
|------------------+----------|
| 1                | M        |
| 13               | F        |
+------------------+----------+
```

Ejemplo 2: Conserve 2 documentos duplicados

En el ejemplo se muestra cómo deduplicar los documentos con el campo de género, conservando dos duplicados.

Consulta 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        |
+------------------+----------+
```

Ejemplo 3: Mantener o ignorar el campo vacío de forma predeterminada

El ejemplo muestra cómo deduplicar el documento manteniendo el campo de valor nulo.

Consulta 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   |
+------------------+-----------------------+
```

El ejemplo muestra cómo deduplicar el documento ignorando el campo de valor vacío.

Consulta 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   |
+------------------+-----------------------+
```

Ejemplo 4: Deduplicación en documentos consecutivos

En el ejemplo se muestra cómo deduplicar en documentos consecutivos.

Consulta 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        |
+------------------+----------+
```

Ejemplos adicionales

- `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` (no se admite la deduplicación consecutiva)

Limitación

- En `| 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
```

- En `| 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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `describe` para obtener información detallada sobre la estructura y los metadatos de las tablas, los esquemas y los catálogos. A continuación, se muestran varios ejemplos y casos de uso del comando `describe`.

Describir

- `describe table` Este comando es igual al comando `DESCRIBE EXTENDED table` de SQL
- `describe schema.table`
- `describe schema.\`table\``
- `describe catalog.schema.table`
- `describe catalog.schema.\`table\``
- `describe \`catalog\`.\`schema\`.\`table\``

Comando eval

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

El comando `eval` evalúa la expresión y agrega el resultado al resultado de la búsqueda.

Sintaxis

Utilice la siguiente sintaxis:

```
eval <field>=<expression> ["," <field>=<expression> ]...
```

- `field`: obligatorio. Si el nombre de campo no existe, se agrega un campo nuevo. Si el nombre de campo ya existe, se sobrescribirá.
- `expression`: obligatorio. Cualquier expresión admitida por el sistema.

Ejemplo 1: creación de un campo nuevo

En este ejemplo se muestra cómo crear un nuevo campo `doubleAge` para cada documento. El nuevo `doubleAge` es el resultado de la evaluación de la edad multiplicado por 2.

Consulta 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          |
+-------+-------------+
```

Ejemplo 2: anular el campo existente

En este ejemplo, se muestra cómo anular el campo de edad existente por edad más 1.

Consulta PPL:

```
os> source=accounts | eval age = age + 1 | fields age ;
fetched rows / total rows = 4/4
+-------+
| age   |
|-------|
| 33    |
| 37    |
| 29    |
| 34    |
+-------+
```

Ejemplo 3: cree el nuevo campo con el campo definido en eval

En este ejemplo se muestra cómo crear un nuevo campo `ddAge` con un campo definido en el comando eval. El nuevo campo `ddAge` es el resultado de la evaluación de `doubleAge` multiplicado por 2, donde `doubleAge` se define en el comando eval.

Consulta 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     |
+-------+-------------+---------+
```

Supuestos: `a`, `b` y `c` son campos existentes en `table`

Ejemplos adicionales

- `source = table | eval f = 1 | fields a,b,c,f`
- `source = table | eval f = 1` (campos de salida 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`

Evaluación con un ejemplo de 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')
```

Evaluación con otro ejemplo de caso:

Supuestos: `a`, `b` y `c` son campos existentes en `table`

Ejemplos adicionales

- `source = table | eval f = 1 | fields a,b,c,f`

- `source = table | eval f = 1` (campos de salida 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`

Evaluación con un ejemplo de 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')
```

Evaluación con otro ejemplo de 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
```

Limitaciones

- No se admite el anular los campos existentes. Las consultas que lo intenten generarán excepciones con el mensaje “La referencia 'a' es ambigua”.

```
- `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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `eventstats` para enriquecer los datos del evento con estadísticas resumidas calculadas. Funciona mediante el análisis de campos específicos de los eventos, el cálculo de diversas medidas estadísticas y, a continuación, la adición de estos resultados como campos nuevos a cada evento original.

Aspectos clave de eventstats

1. Realiza cálculos en todo el conjunto de resultados o dentro de grupos definidos.
2. Los eventos originales permanecen intactos y se agregan nuevos campos para contener los resultados estadísticos.
3. El comando es particularmente útil para realizar análisis comparativos, identificar valores atípicos o proporcionar contexto adicional a eventos individuales.

Diferencia entre stats y eventstats

Los comandos `stats` y `eventstats` se utilizan para calcular estadísticas, pero tienen algunas diferencias clave en su funcionamiento y en lo que producen.

Formato de salida

- `stats`: produce una tabla de resumen con solo las estadísticas calculadas.
- `eventstats`: agrega las estadísticas calculadas como campos nuevos a los eventos existentes y conserva los datos originales.

Retención de eventos

- `stats`: reduce el conjunto de resultados a solo un resumen estadístico y descarta los eventos individuales.
- `eventstats`: retiene todos los eventos originales y agrega nuevos campos con las estadísticas calculadas.

Casos de uso

- `stats`: ideal para crear paneles o informes resumidos. Suele utilizarse como comando final para resumir los resultados.
- `eventstats`: útil cuando se necesita enriquecer los eventos con un contexto estadístico para su posterior análisis o filtrado. Se puede utilizar en mitad de una búsqueda para agregar estadísticas que se pueden utilizar en comandos posteriores.

Sintaxis

Utilice la siguiente sintaxis:

```
eventstats <aggregation>... [by-clause]
```

agregación

- Obligatorio.
- Una función de agregación.
- El argumento de agregación debe ser un campo.

by-clause

- Opcional.
- Sintaxis: `by [span-expression,] [field,]...`
- La cláusula by puede incluir campos y expresiones, como funciones escalares y funciones de agregación. También puede usar la cláusula span para dividir un campo específico en grupos de intervalos iguales. A continuación, el comando eventstats realiza la agregación en función de estos buckets de intervalo.
- Predeterminado: si no especificas una cláusula by, el comando eventstats agrega todo el conjunto de resultados.

span-expression

- Opcional, como máximo uno.
- Sintaxis: `span(field_expr, interval_expr)`
- La unidad de la expresión de intervalo es la unidad natural por defecto. Sin embargo, para los campos de tipo fecha y hora, debe especificar la unidad en la expresión de intervalo cuando utilice unidades de fecha y hora.

  Por ejemplo, para dividir el campo `age` en buckets por 10 años, utilice `span(age, 10)`. Para los campos basados en el tiempo, puede dividir un campo `timestamp` en intervalos de una hora si utiliza `span(timestamp, 1h)`.

Unidades de tiempo disponibles

Unidades de intervalo
milisegundo (ms)
segundo (s)
minuto (m, distingue entre mayúsculas y minúsculas)
hora (h)
día (d)
semana (s)

Unidades de intervalo
mes (M, distingue entre mayúsculas y minúsculas)
trimestre (q)
año (y)

Funciones de agregación

COUNT

COUNT devuelve un recuento del número de expresiones en las filas recuperadas por una sentencia SELECT.

En el CloudWatch caso de los registros, no COUNT se admiten consultas.

Ejemplo:

```
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) devuelve la suma de expr.

Ejemplo:

```
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)` devuelve el valor medio de expr.

Ejemplo:

```
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)` devuelve el valor máximo de expr.

Ejemplo

```
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)` devuelve el valor mínimo de expr.

Ejemplo

```
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)` devuelve la desviación estándar muestral de expr.

Ejemplo

```
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)` devuelve la desviación estándar poblacional de expr.

Ejemplo

```
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)` devuelve el valor percentil aproximado de expr en el porcentaje especificado.

porcentaje

- El número debe ser una constante entre 0 y 100.

Ejemplo

```
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
       |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+--------------------------------+
```

Ejemplo 1: calcular el promedio, la suma y el recuento de un campo por grupo

En el ejemplo se muestra cómo calcular la edad media, la edad sumada y el recuento de eventos de todas las cuentas agrupadas por género.

```
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     |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----------+-----------+--------+
```

Ejemplo 2: Calcule el recuento en un intervalo

El ejemplo obtiene el recuento de la edad en un intervalo de 10 años.

```
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        |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+----------+
```

Ejemplo 3: Calcule el recuento por género y rango

El ejemplo obtiene el recuento de la edad por un intervalo de 5 años y el grupo por género.

```
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   |
+----------------+----------+-----------+----------+-----+--------
+----------------------+-------------+--------------------------+--------+-------
+-----+
```

De uso

- `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)`

Agregaciones 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`

Agregaciones con intervalo de ventana de tiempo (función de ventana invertida)

- `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`

Las agregaciones se agrupan en varios niveles

- `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**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `expand` para aplanar un campo del tipo:

- `Array<Any>`
- `Map<Any>`

Sintaxis

Utilice la siguiente sintaxis:

```
expand <field> [As alias]
```

campo

- El campo que se va a expandir (explotar). Debe ser de un tipo compatible.

alias

- Opcional. El nombre que se va a utilizar en lugar del nombre del campo original.

De uso

El comando `expand` genera una fila para cada elemento de la matriz o campo de mapa especificado, donde:

- Los elementos de la matriz se convierten en filas individuales.
- Los pares clave-valor de la asignación se dividen en filas independientes y cada valor clave se representa como una fila.

- Cuando se proporciona un alias, los valores desglosados se representan con el alias en lugar del nombre de campo original.

- Esto se puede usar en combinación con otros comandos, como `stats`, `eval` y `parse` para manipular o extraer datos después de la expansión.

Ejemplos

- `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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

El comando `explain` lo ayuda a entender los planes de ejecución de las consultas, lo que le permite analizarlas y optimizarlas para obtener un mejor rendimiento. Esta introducción proporciona una descripción general y concisa del propósito del comando explain y de su importancia en la optimización de las consultas.

Comment

- `source=accounts | top gender // finds most common gender of all the accounts` (comentario de línea)

- `source=accounts | dedup 2 gender /* dedup the document with gender field keep 2 duplication */ | fields account_number, gender` (comentario bloqueado)

Describir

- `describe table` Este comando es igual al comando `DESCRIBE EXTENDED table` de SQL
- `describe schema.table`
- `describe schema.\`table\``
- `describe catalog.schema.table`
- `describe catalog.schema.\`table\``
- `describe \`catalog\`.\`schema\`.\`table\``

Explain

- `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`

Fields

- `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`

Resumen de 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 anidado

- `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`

Filtros

- `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: Esto devuelve un >= 1 y un <= 4, es decir, [1, 4]
- `source = table | where b not between '2024-09-10' and '2025-09-10'`, Nota: Esto devuelve b >= '**********' y 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`

Consultas relacionadas con la 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`

Filtros complejos

```
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
```

Filtros con condiciones lógicas

- `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`

Eval

Supuestos: a, b, c son campos existentes en `table`

- `source = table | eval f = 1 | fields a,b,c,f`
- `source = table | eval f = 1` (campos de salida 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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Description (Descripción)

Use el comando `fillnull` para reemplazar los valores nulos por un valor especificado en uno o más campos de los resultados de la búsqueda.

Sintaxis

Utilice la siguiente sintaxis:

```
fillnull [with <null-replacement> in <nullable-field>["," <nullable-field>]] | [using
 <source-field> = <null-replacement> ["<source-field> = <null-replacement>]]
```

- null-replacement: obligatorio. El valor utilizado para reemplazar los valores nulos.
- nullable-field: obligatorio. Referencia de campos. Los valores nulos de este campo se sustituirán por el valor especificado en la sustitución de valores nulos.

Ejemplo 1: Fillnull en un campo

En el ejemplo se muestra cómo utilizar fillnull en un solo 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         |
```

Ejemplo 2: Fillnull aplicado a varios campos

El ejemplo muestra la aplicación fillnull a varios campos.

```
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   |
```

Ejemplo 3: Se aplica fillnull a varios campos con varios valores de reemplazo nulos.

El ejemplo muestra fillnull con varios valores que se utilizan para reemplazar los valores nulos.

- `/error` en el campo `request_path`
- `1970-01-01 00:00:00` en el campo `timestamp`

```
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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `fields` para conservar o eliminar campos del resultado de la búsqueda.

Sintaxis

Utilice la siguiente sintaxis:

```
field [+|-] <field-list>
```

- `index`: opcional.

  Si se utiliza el signo más (+), solo se conservarán los campos especificados en la lista de campos.

  Si se utiliza el signo menos (-), se eliminarán todos los campos especificados en la lista de campos.

Predeterminado: +

- `field list`: obligatorio. Una lista delimitada por comas de campos para conservar o eliminar.

Ejemplo 1: seleccionar campos específicos del resultado

En este ejemplo se muestra cómo obtener campos `account_number`, `firstname` y `lastname` de los resultados de la búsqueda.

Consulta 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         |
+------------------+-------------+------------+
```

Ejemplo 2: Eliminar los campos especificados del resultado

En este ejemplo se muestra cómo quitar el campo `account_number` de los resultados de búsqueda.

Consulta 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         |
+-------------+------------+
```

Ejemplos adicionales

- `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`

Ejemplo de campos anidados:

```
`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**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando flatten para expandir los campos de los siguientes tipos:

- `struct<?,?>`
- `array<struct<?,?>>`

Sintaxis

Utilice la siguiente sintaxis:

```
flatten <field>
```

- campo: el campo que se va a aplanar. El campo debe ser de un tipo compatible.

Esquema

col_name	data_type
_time	cadena
puentes	array<struct<length:bigint,name:string>>
ciudad	cadena
coor	struct<alt:bigint,lat:double,long:double>
país	cadena

Datos

_time	puentes	ciudad	coor	país
2024-09-13T12:00:00	[{801, Puente de la Torre}, {928, Puente de Londres}]	Londres	{35, 51.5074, -0.1278}	Inglaterra
2024-09-13T12:00:00	[{232, Puente Nuevo}, {160, Puente Alejandro III}]	Paris	{35, 48.8566, 2.3522}	Francia
2024-09-13T12:00:00	[{48, Puente de Rialto}, {11, Puente	Venecia	{2, 45.4408, 12.3155}	Italia

_time	puentes	ciudad	coor	país
	de los Suspiros}]			
2024-09-13T12:00:00	[{***, Puente de Carlos}, {343, Puente de la Legión}]	Praga	{200, 50.0755, 14.4378}	República Checa
2024-09-13T12:00:00	[{375, Puente de las Cadenas}, {333, Puente de la Libertad}]	Budapest	{96, 47.4979, 19.0402}	Hungría
1990-09-13T12:00:00	NULL	Varsovia	NULL	Polonia

Ejemplo 1: estructura aplanada

En este ejemplo se muestra cómo aplanar un campo de estructura.

Consulta PPL:

```
source=table | flatten coor
```

_time	puentes	ciudad	país	alt	lat	long
2024-09-13T12:00:00	[{801, Puente de la Torre}, {928,	Londres	Inglaterra	35	51.5074	-0,1278

_time	puentes	ciudad	país	alt	lat	long
	Puente de Londres}]					
2024-09-13T12:00:00	[{232, Puente Nuevo}, {160, Puente Alejandro III}]	Paris	Francia	35	48,8566	2.3522
2024-09-13T12:00:00	[{48, Puente de Rialto}, {11, Puente de los Suspiros}]	Venecia	Italia	2	45,4408	12,3155
2024-09-13T12:00:00	[{516, Puente de Carlos}, {343, Puente de la Legión}]	Praga	República Checa	200	50,0755	14,4378

_time	puentes	ciudad	país	alt	lat	long
2024-09-13T12:00:00	[{375, Puente de las Cadenas} {333, Puente de la Libertad}]	Budapest	Hungría	96	47,4979	19,0402
1990-09-13T 12:00:00	NULL	Varsovia	Polonia	NULL	NULL	NULL

Ejemplo 2: aplanar una matriz

En el ejemplo se muestra cómo aplanar una matriz de campos de estructura.

Consulta PPL:

```
source=table | flatten bridges
```

_time	ciudad	coor	país	length	name
2024-09-13T12:00:00	Londres	{35, 51.5074, -0.1278}	Inglaterra	801	Puente de la Torre
2024-09-13T12:00:00	Londres	{35, 51.5074, -0.1278}	Inglaterra	928	Puente de Londres
2024-09-13T12:00:00	Paris	{35, 48.8566, 2.3522}	Francia	232	Puente Nuevo
2024-09-13T12:00:00	Paris	{35, 48.8566, 2.3522}	Francia	160	Puente Alejandro III

_time	ciudad	coor	país	length	name
2024-09-13T12:00:00	Venecia	{2, 45.4408, 12.3155}	Italia	48	Puente de Rialto
2024-09-13T12:00:00	Venecia	{2, 45.4408, 12.3155}	Italia	11	Puente de los Suspiros
2024-09-13T12:00:00	Praga	{200, 50.0755, 14.4378}	República Checa	516	Puente de Carlos
2024-09-13T12:00:00	Praga	{200, 50.0755, 14.4378}	República Checa	343	Puente de la Legión
2024-09-13T12:00:00	Budapest	{96, 47.4979, 19.0402}	Hungría	375	Puente de las Cadenas
2024-09-13T12:00:00	Budapest	{96, 47.4979, 19.0402}	Hungría	333	Puente de la Libertad
1990-09-13T12:00:00	Varsovia	NULL	Polonia	NULL	NULL

Ejemplo 3: aplanar una matriz y una estructura

En este ejemplo, se muestra cómo aplanar varios campos.

Consulta PPL:

```
source=table | flatten bridges | flatten coor
```

_time	ciudad	país	length	name	alt	lat	long
2024-09-13T12:00:00	Londres	Inglaterra	801	Puente de la Torre	35	51,5074	-0,1278

_time	ciudad	país	length	name	alt	lat	long
2024-09-13T12:00:00	Londres	Inglaterra	928	Puente de Londres	35	51,5074	-0,1278
2024-09-13T12:00:00	Paris	Francia	232	Puente Nuevo	35	48,8566	2.3522
2024-09-13T12:00:00	Paris	Francia	160	Puente Alejandro III	35	48,8566	2.3522
2024-09-13T12:00:00	Venecia	Italia	48	Puente de Rialto	2	45,4408	12,3155
2024-09-13T12:00:00	Venecia	Italia	11	Puente de los Suspiros	2	45,4408	12,3155
2024-09-13T12:00:00	Praga	República Checa	516	Puente de Carlos	200	50,0755	14,4378
2024-09-13T12:00:00	Praga	República Checa	343	Puente de la Legión	200	50,0755	14,4378
2024-09-13T12:00:00	Budape	Hungría	375	Puente de las Cadenas	96	47,4979	19,0402
2024-09-13T12:00:00	Budape	Hungría	333	Puente de la Libertad	96	47,4979	19,0402

_time	ciudad	país	length	name	alt	lat	long
1990-09-13T 12:00:00	Varsovi	Polonia	NULL	NULL	NULL	NULL	NULL

Comando grok

> **Note**
>
> Para ver qué integraciones de fuentes AWS de datos admiten este comando PPL, consulte. the section called “Comandos”

El comando `grok` analiza un campo de texto con un patrón grok y agrega los resultados al resultado de la búsqueda.

Sintaxis

Utilice la siguiente sintaxis:

```
grok <field> <pattern>
```

campo

- Obligatorio.
- El campo debe ser un campo de texto.

pattern

- Obligatorio.
- El patrón grok utilizado para extraer nuevos campos del campo de texto dado.
- Si ya existe un nombre de campo nuevo, reemplazará al campo original.

Patrón de Grok

El patrón grok se utiliza para hacer coincidir el campo de texto de cada documento y extraer nuevos campos.

Ejemplo 1: creación de un campo nuevo

En este ejemplo se muestra cómo crear un nuevo campo `host` para cada documento. `host` será el nombre de host que aparece después `@` en el campo `email`. Si se analiza un campo nulo, se devuelve una cadena vacía.

```
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 |
+-------------------------+-------------+
```

Ejemplo 2: anular el campo existente

En este ejemplo se muestra cómo anular el campo `address` existente quitando el número de calle.

```
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    |
+------------------+
```

Ejemplo 3: Uso de grok para analizar registros

En este ejemplo, se muestra cómo utilizar grok para analizar registros sin procesar.

```
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
    |
+-----------------------------------------------------------------------------------------------------------
+----------------------------+-----------+---------+
```

Limitaciones

El comando grok tiene las mismas limitaciones que el comando parse.

Comando head

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `head` para devolver el primer número N de resultados especificados después de un desplazamiento opcional en el orden de búsqueda.

Sintaxis

Utilice la siguiente sintaxis:

```
head [<size>] [from <offset>]
```

<size>

- Número entero opcional.
- El número de resultados que se devolverá.
- Valor predeterminado: 10

<offset>

- Entero después de opcional `from`.
- Número de resultados que se va a saltear.
- Predeterminado: 0

Ejemplo 1: Obtenga los primeros 10 resultados

En este ejemplo se muestra cómo recuperar un máximo de 10 resultados del índice de cuentas.

Consulta PPL:

```
os> source=accounts | fields firstname, age | head;
fetched rows / total rows = 4/4
+-------------+-------+
| firstname   | age   |
|-------------+-------|
| Jane        | 32    |
| John        | 36    |
| Jorge       | 28    |
| Juan        | 33    |
+-------------+-------+
```

Ejemplo 2: Obtenga los primeros N resultados

El ejemplo muestra los primeros N resultados del índice de cuentas.

Consulta PPL:

```
os> source=accounts | fields firstname, age | head 3;
fetched rows / total rows = 3/3
+-------------+-------+
| firstname   | age   |
```

```
|-------------+-------|
| Jane        | 32    |
| John        | 36    |
| Jorge       | 28    |
+-------------+-------+
```

Ejemplo 3: Obtenga los primeros N resultados después de la compensación M

En este ejemplo se muestra cómo recuperar los primeros N resultados después de omitir M resultados del índice de cuentas.

Consulta 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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

El comando join le permite combinar datos de varios orígenes en función de campos comunes, lo que le permite realizar análisis complejos y obtener información más detallada a partir de sus conjuntos de datos distribuidos.

Esquema

Hay al menos dos índices: `otel-v1-apm-span-*` (grande) y `otel-v1-apm-service-map` (pequeño).

Campos relevantes de los índices:

`otel-v1-apm-span-*`

- traceId: un identificador único para un rastreo. Todos los tramos de la misma traza comparten el mismo traceId.
- spanId: identificador único para un tramo dentro de una traza, que se asigna cuando se crea el tramo.
- parentSpanId - El spAnid del tramo principal de este tramo. Si se trata de un intervalo raíz, este campo debe estar vacío.
- durationInNanos - La diferencia en nanosegundos entre StartTime y EndTime. (esto está `latency` en la interfaz de usuario)
- serviceName: el recurso desde el que se origina el tramo.
- traceGroup: nombre del tramo raíz de la traza.

`otel-v1-apm-service-map`

- serviceName: nombre del servicio que registró el intervalo.
- destination.domain: el serviceName del servicio que llama este cliente.
- destination.resource: el nombre del intervalo (API, operación, etc.) al que llama este cliente.
- target.domain: el serviceName del servicio que llama un cliente.
- target.resource: el nombre del intervalo (API, operación, etc.) al que llama un cliente.
- traceGroupName - El nombre del intervalo de nivel superior que inició la cadena de solicitudes.

Requisito

Admita join para calcular lo siguiente:

Para cada servicio, combine el índice span con el índice del mapa de servicios para calcular las métricas con distintos tipos de filtros.

Esta consulta de ejemplo calcula la latencia al filtrarla por grupo de rastreo `client_cancel_order` para el servicio `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`)
```

Migración a PPL

Sintaxis del comando join

```
SEARCH source=<left-table>
| <other piped command>
| [joinType] JOIN
    [leftAlias]
    ON joinCriteria
    <right-table>
| <other piped command>
```

Reescrituras

```
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
```

joinType

- Sintaxis: `INNER | LEFT OUTER | CROSS`
- Opcional
- El tipo de unión que se va a realizar. Si no se especifica, el valor predeterminado es `INNER`.

leftAlias

- Sintaxis: `left = <leftAlias>`
- Opcional
- El alias de la subconsulta que se debe usar con el lado izquierdo de la combinación, para evitar nombres ambiguos.

joinCriteria

- Sintaxis: `<expression>`
- Obligatorio
- La sintaxis comienza con `ON`. Puede ser cualquier expresión de comparación. En general, los criterios de unión son los siguientes `<leftAlias>.<leftField>=<rightAlias>.<rightField>`.

  Por ejemplo: `l.id = r.id`. Si los criterios de unión contienen varias condiciones, puede especificar operador `AND` y `OR` entre cada expresión de comparación. Por ejemplo, `l.id = r.id AND l.email = r.email AND (r.age > 65 OR r.age < 18)`.

Más ejemplos

Migración desde una consulta 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;
```

Reescrito por PPL Join Query:

```
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
```

Limitación: no se admiten subbúsquedas en el lado derecho de la unión.

Si se admiten las subbúsquedas, puedes reescribir la consulta PPL anterior de la siguiente manera:

```
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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `lookup` para enriquecer los datos de búsqueda mediante la adición o sustitución de datos de un índice de búsqueda (tabla de dimensiones). Este comando le permite extender los

campos de un índice con valores de una tabla de dimensiones. También se puede utilizar para agregar o reemplazar valores cuando se cumplen las condiciones de búsqueda. El comando `lookup` es más adecuado que el comando `Join` para enriquecer los datos de origen con un conjunto de datos estático.

Sintaxis

Utilice la siguiente sintaxis:

```
SEARCH source=<sourceIndex>
| <other piped command>
| LOOKUP <lookupIndex> (<lookupMappingField> [AS <sourceMappingField>])...
    [(REPLACE | APPEND) (<inputField> [AS <outputField>])...]
| <other piped command>
```

lookupIndex

- Obligatorio.
- El nombre del índice de búsqueda (tabla de dimensiones).

lookupMappingField

- Obligatorio.
- Una clave de asignación en el índice de búsqueda, análoga a una clave de unión de la tabla de la derecha. Puede especificar múltiples nombres de esquema separados por comas.

sourceMappingField

- Opcional.
- Predeterminado: **< >**lookupMappingField.
- Una clave de asignación de la consulta de origen, análoga a una clave de unión del lado izquierdo.

inputField

- Opcional.
- Predeterminado: todos los campos del índice de búsqueda en los que se encuentran valores coincidentes.

- Campo del índice de búsqueda en el que los valores coincidentes se aplican a la salida del resultado. Puede especificar múltiples nombres de esquema separados por comas.

outputField

- Opcional.
- Predeterminado: `<inputField>`.
- Un campo de salida. Puede especificar varios campos de salida. Si especifica un nombre de campo existente en la consulta de origen, sus valores se sustituirán o se agregarán por valores coincidentes de inputField. Si especifica un nombre de campo nuevo, se agregará a los resultados.

REPLACE | APPEND

- Opcional.
- Predeterminado: REPLACE
- Especifica cómo administrar los valores coincidentes. Si especifica REPLACE, los valores coincidentes del campo **<lookupIndex>** sobrescribirán los valores del resultado. Si especifica `APPEND`, los valores coincidentes del campo **<lookupIndex>** solo se incorporarán a los valores faltantes en el resultado.

De uso

- LOOKUP **<lookupIndex>** id AS cid REPLACE mail AS email
- LOOKUP **<lookupIndex>** name REPLACE mail AS email
- LOOKUP **<lookupIndex>** id AS cid, name APPEND address, mail AS email
- LOOKUP **<lookupIndex>** id

Ejemplo

Vea los siguientes ejemplos de .

```
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

El comando `parse` analiza un campo de texto con una expresión regular y agrega el resultado al resultado de la búsqueda.

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando PPL, consulte. the section called “Comandos”

Sintaxis

Utilice la siguiente sintaxis:

```
parse <field> <pattern>
```

`field`

- Obligatorio.
- El campo debe ser un campo de texto.

`pattern`

- Cadena obligatoria.
- Este es el patrón de expresión regular que se utiliza para extraer nuevos campos del campo de texto dado.
- Si ya existe un nombre de campo nuevo, reemplazará al campo original.

Expresión regular

El patrón de expresiones regulares se utiliza para hacer coincidir todo el campo de texto de cada documento con el motor de expresiones regulares de Java. Cada grupo de captura nombrado en la expresión se convertirá en un campo nuevo `STRING`.

Ejemplo 1: creación de un campo nuevo

En el ejemplo se muestra cómo crear un nuevo campo `host` para cada documento. `host` será el nombre del host que aparece después `@` en el campo `email`. Si se analiza un campo nulo, se devuelve una cadena vacía.

Consulta 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 |
+-----------------------+-------------+
```

Ejemplo 2: Anular un campo existente

El ejemplo muestra cómo anular el campo `address` existente quitando el número de calle.

Consulta 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    |
+------------------+
```

Ejemplo 3: Filtrar y ordenar por campo analizado

En el ejemplo se muestra cómo ordenar los números de calle superiores a 500 en el campo `address`.

Consulta 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   |
+----------------+----------------+
```

Limitaciones

El comando parse presenta algunas limitaciones:

- Los campos definidos por parse no se pueden volver a analizar.

  Este comando no funcionara:

  ```
  source=accounts | parse address '\d+ (?<street>.+)' | parse street '\w+ (?<road>\w+)'
  ```

- Los campos definidos por parse no se pueden anular con otros comandos.

  `where` no coincidirá con ningún documento ya que `street` no se puede anular:

  ```
  source=accounts | parse address '\d+ (?<street>.+)' | eval street='1' | where
   street='1' ;
  ```

- El campo de texto utilizado por parse no se puede anular.

  `street` no se analizará correctamente ya que `address` está anulado:

  ```
  source=accounts | parse address '\d+ (?<street>.+)' | eval address='1' ;
  ```

- Los campos definidos por parse no se pueden filtrar ni ordenar después de usarlos en el comando `stats`.

`where` en el siguiente comando no funcionará:

```
source=accounts | parse email '.+@(?<host>.+)' | stats avg(age) by host | where
 host=pyrami.com ;
```

Comando patterns

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

El comando `patterns` extrae los patrones de registro de un campo de texto y agrega los resultados al resultado de la búsqueda. Al agrupar los registros por sus patrones, resulta más fácil agregar estadísticas de grandes volúmenes de datos de registro para analizarlos y solucionar problemas.

Sintaxis

Utilice la siguiente sintaxis:

```
patterns [new_field=<new-field-name>] [pattern=<pattern>] <field>
```

new-field-name

- Cadena opcional.
- Este es el nombre del nuevo campo para los patrones extraídos.
- El valor predeterminado es `patterns_field`.
- Si el nombre ya existe, sustituirá al campo original.

pattern

- Cadena opcional.
- Este es el patrón de caracteres regex que deben filtrarse del campo de texto.
- Si no está presente, el patrón predeterminado son los caracteres alfanuméricos (`[a-zA-Z\d]`).

campo

- Obligatorio.
- El campo debe ser un campo de texto.

Ejemplo 1: creación de un campo nuevo

En el ejemplo se muestra cómo utilizar los signos de puntuación extraídos en `email` cada documento. Si se analiza un campo nulo, se devuelve una cadena vacía.

Consulta 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   | @.               |
+-----------------------+------------------+
```

Ejemplo 2: Extraer patrones de registro

El ejemplo muestra cómo extraer signos de puntuación de un campo de registro sin procesar utilizando los patrones predeterminados.

Consulta 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
                                       | ... - - [//::: -] " / /."       |
+-------------------------------------------------------------------------------------------
+--------------------------------+
```

Ejemplo 3: Extraiga patrones de registro con un patrón de expresiones regulares personalizado

El ejemplo muestra cómo extraer signos de puntuación de un campo de registro sin procesar utilizando patrones definidos por el usuario.

Consulta 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/."
                                       |
+-------------------------------------------------------------------------------------------
+-----------------------------------------------------------------------------
+
```

Limitación

El comando patterns tiene las mismas limitaciones que el comando parse.

Comando rare

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `rare` para buscar la tupla de valores menos común de todos los campos de la lista de campos.

> **Note**
>
> Se devuelve un máximo de 10 resultados por cada tupla distinta de valores de los campos agrupados.

Sintaxis

Utilice la siguiente sintaxis:

```
rare [N] <field-list> [by-clause] rare_approx [N] <field-list> [by-clause]
```

field-list

- Obligatorio.
- Una lista delimitada por comas de nombres de campo.

by-clause

- Opcional.
- Uno o más campos por los que agrupar los resultados.

N

- El número de resultados que se devolverá.

- Valor predeterminado: 10

rare_approx

- El recuento aproximado de los campos raros (n) mediante el uso de la cardinalidad estimada mediante el algoritmo ++. HyperLogLog

Ejemplo 1: Encuentre los valores menos comunes en un campo

El ejemplo busca el género menos común de todas las cuentas.

Consulta 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        |
+----------+
```

Ejemplo 2: encuentra los valores menos comunes organizados por género

En el ejemplo, se busca la edad menos común de todas las cuentas agrupadas por género.

Consulta 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 rename

Utilice el comando `rename` para cambiar los nombres de uno o más campos del resultado de la búsqueda.

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Sintaxis

Utilice la siguiente sintaxis:

```
rename <source-field> AS <target-field>["," <source-field> AS <target-field>]...
```

source-field

- Obligatorio.
- Este es el nombre del campo que desea cambiar de nombre.

target-field

- Obligatorio.
- Este es el nombre al que quiere cambiarle el nombre.

Ejemplo 1: cambiar el nombre de un campo

En este ejemplo, se muestra cómo renombrar un único campo.

Consulta PPL:

```
os> source=accounts | rename account_number as an | fields an;
fetched rows / total rows = 4/4
+------+
| an   |
|------|
| 1    |
| 6    |
```

```
| 13   |
| 18   |
+------+
```

Ejemplo 2: cambiar el nombre de varios campos

En este ejemplo, se muestra cómo renombrar varios campos.

Consulta 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    |
+------+---------+
```

Limitaciones

- No se admite la sustitución de un campo existente:

```
source=accounts | grok address '%{NUMBER} %{GREEDYDATA:address}' | fields address
```

Comando search

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `search` para recuperar documentos de un índice. El comando `search` solo se puede usar como primer comando en una consulta PPL.

Sintaxis

Utilice la siguiente sintaxis:

```
search source=[<remote-cluster>:]<index> [boolean-expression]
```

buscar

- Opcional.
- Busca palabras clave, que se pueden omitir.

índice

- Obligatorio.
- El comando de búsqueda debe especificar el índice desde el que realizar la consulta.
- El nombre del índice puede tener un prefijo `<cluster name>:` para las búsquedas entre clústeres.

bool-expression

- Opcional.
- Una cadena de expresión que se evalúa como un valor booleano.

Ejemplo 1: obtener todos los datos

El ejemplo muestra cómo recuperar todo el documento del índice de cuentas.

Consulta 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         |
+------------------+-------------+----------------------+-----------+----------
+--------+----------------+---------+-------+----------------------+------------+
```

Ejemplo 2: Obtener datos con una condición

El ejemplo muestra cómo recuperar todo el documento del índice de cuentas con .

Consulta 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

Utilice el comando `sort` para ordenar los resultados de la búsqueda por campos específicos.

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Sintaxis

Utilice la siguiente sintaxis:

```
sort <[+|-] sort-field>...
```

[+|-]

- Opcional.
- El signo más [+] significa orden ascendente con NULL/MISSING los valores en primer lugar.
- El signo menos [-] indica el orden descendente con los NULL/MISSING valores al final.
- Predeterminado: orden ascendente con NULL/MISSING los valores primero.

sort-field

- Obligatorio.
- El campo que se utiliza para ordenar.

Ejemplo 1: Ordenar por un campo

El ejemplo muestra cómo ordenar el documento con el campo de edad en orden ascendente.

Consulta 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    |
+------------------+-------+
```

Ejemplo 2: ordenar por un campo y devolver todos los resultados

El ejemplo muestra cómo ordenar el documento con el campo de edad en orden ascendente.

Consulta 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    |
+------------------+-------+
```

Ejemplo 3: ordenar por un campo en orden descendente

El ejemplo muestra cómo ordenar el documento con el campo de edad en orden descendente.

Consulta 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    |
+------------------+-------+
```

Ejemplo 4: Ordenar por varios campos

El ejemplo muestra cómo ordenar el documento con el campo de género en orden ascendente y el campo de edad en orden descendente.

Consulta 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    |
+------------------+----------+-------+
```

Ejemplo 5: Ordenar por campo e incluir un valor nulo

En el ejemplo se muestra cómo ordenar el campo del empleador por la opción predeterminada (en orden ascendente y con los valores nulo primero). El resultado muestra que el valor nulo está en la primera fila.

Consulta PPL:

```
os> source=accounts | sort employer | fields employer;
fetched rows / total rows = 4/4
+------------+
| employer   |
|------------|
| null       |
| AnyCompany |
| AnyCorp    |
| AnyOrgty   |
+------------+
```

Comandos stats

Utilice el comando `stats` para calcular la agregación a partir del resultado de la búsqueda.

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Manejo de valores NULL o MISSING

Manejo de valores NULL o MISSING

Función	NULL	AUSENTE
COUNT	No contado	No contado
SUM	Ignore	Ignore
AVG	Ignore	Ignore

Función	NULL	AUSENTE
MAX	Ignore	Ignore
MIN	Ignore	Ignore

Sintaxis

Utilice la siguiente sintaxis:

```
stats <aggregation>... [by-clause]
```

agregación

- Obligatorio.
- Función de agregación aplicada a un campo.

by-clause

- Opcional.
- Sintaxis: `by [span-expression,] [field,]...`
- Especifica los campos y expresiones para agrupar los resultados de la agregación. La cláusula secundaria le permite agrupar los resultados de agregación mediante campos y expresiones. Puede utilizar funciones escalares, funciones de agregación e incluso expresiones de extensión para dividir campos específicos en grupos de intervalos iguales.
- Predeterminado: si `<by-clause>` no se especifica, el comando stats devuelve una sola fila que representa la agregación de todo el conjunto de resultados.

span-expression

- Opcional, como máximo uno.
- Sintaxis: `span(field_expr, interval_expr)`
- La unidad de la expresión de intervalo es la unidad natural por defecto. Si el campo es de tipo fecha y hora y el intervalo está en date/time unidades, especifique la unidad en la expresión del intervalo.

- Por ejemplo, parece que se divide el campo `age` en buckets de 10 años `span(age, 10)`. Para dividir un campo de fecha y hora en intervalos de una hora, utilice `span(timestamp, 1h)`.

Unidades de tiempo disponibles

Unidades de intervalo
milisegundo (ms)
segundo (s)
minuto (m, distingue entre mayúsculas y minúsculas)
hora (h)
día (d)
semana (s)
mes (M, distingue entre mayúsculas y minúsculas)
trimestre (q)
año (y)

Funciones de agregación

COUNT

Devuelve un recuento del número de expr en las filas recuperadas por una sentencia SELECT.

Ejemplo:

```
os> source=accounts | stats count();
fetched rows / total rows = 1/1
+-----------+
| count()   |
|-----------|
| 4         |
+-----------+
```

SUM

Se utiliza `SUM(expr)` para devolver la suma de expr.

Ejemplo

```
os> source=accounts | stats sum(age) by gender;
fetched rows / total rows = 2/2
+------------+----------+
| sum(age)   | gender   |
|------------+----------|
| 28         | F        |
| 101        | M        |
+------------+----------+
```

AVG

Se utiliza `AVG(expr)` para devolver el valor medio de la expr.

Ejemplo

```
os> source=accounts | stats avg(age) by gender;
fetched rows / total rows = 2/2
+--------------------+----------+
| avg(age)           | gender   |
|--------------------+----------|
| 28.0               | F        |
| 33.666666666666664 | M        |
+--------------------+----------+
```

MAX

Se utiliza `MAX(expr)` para devolver el valor máximo de expr.

Ejemplo

```
os> source=accounts | stats max(age);
fetched rows / total rows = 1/1
+------------+
| max(age)   |
|------------|
| 36         |
```

```
+------------+
```

MIN

Se utiliza `MIN(expr)` para devolver el valor mínimo de expr.

Ejemplo

```
os> source=accounts | stats min(age);
fetched rows / total rows = 1/1
+------------+
| min(age)   |
|------------|
| 28         |
+------------+
```

STDDEV_SAMP

Se utiliza `STDDEV_SAMP(expr)` para devolver la desviación estándar muestral de expr.

Ejemplo:

```
os> source=accounts | stats stddev_samp(age);
fetched rows / total rows = 1/1
+--------------------+
| stddev_samp(age)   |
|--------------------|
| 3.304037933599835  |
+--------------------+
```

STDDEV_POP

Se utiliza `STDDEV_POP(expr)` para devolver la desviación estándar poblacional de expr.

Ejemplo:

```
os> source=accounts | stats stddev_pop(age);
fetched rows / total rows = 1/1
+--------------------+
| stddev_pop(age)    |
|--------------------|
| 2.**************** |
```

```
+--------------------+
```

TAKE

Se utiliza `TAKE(field [, size])` para devolver los valores originales de un campo. No garantiza el orden de los valores.

campo

- Obligatorio.
- El campo debe ser un campo de texto.

tamaño

- Número entero opcional.
- Debe devolverse el número de valores.
- El valor predeterminado es 10.

Ejemplo

```
os> source=accounts | stats take(firstname);
fetched rows / total rows = 1/1
+-----------------------------+
| take(firstname)             |
|-----------------------------|
| [Jane, Mary, Nikki, Juan    |
+-----------------------------+
```

PERCENTILE o PERCENTILE_APPROX

Utilice `PERCENTILE(expr, percent)` o `PERCENTILE_APPROX(expr, percent)` para devolver el valor percentil aproximado de expr en el porcentaje especificado.

porcentaje

- El número debe ser una constante entre 0 y 100.

Ejemplo

```
os> source=accounts | stats percentile(age, 90) by gender;
fetched rows / total rows = 2/2
+----------------------+----------+
| percentile(age, 90)   | gender   |
|----------------------+----------|
| 28                    | F        |
| 36                    | M        |
+----------------------+----------+
```

Ejemplo 1: Calcular el recuento de eventos

El ejemplo muestra cómo calcular el recuento de eventos en las cuentas.

```
os> source=accounts | stats count();
fetched rows / total rows = 1/1
+-----------+
| count()   |
|-----------|
| 4         |
+-----------+
```

Ejemplo 2: Calcular el promedio de un campo

En el ejemplo se muestra cómo calcular la edad media de todas las cuentas.

```
os> source=accounts | stats avg(age);
fetched rows / total rows = 1/1
+------------+
| avg(age)   |
|------------|
| 32.25      |
+------------+
```

Ejemplo 3: Calcular el promedio de un campo por grupo

El ejemplo muestra cómo calcular la edad media de todas las cuentas, agrupadas por género.

```
os> source=accounts | stats avg(age) by gender;
fetched rows / total rows = 2/2
+--------------------+----------+
| avg(age)           | gender   |
|--------------------+----------|
```

```
| 28.0               | F        |
| 33.666666666666664 | M        |
+--------------------+----------+
```

Ejemplo 4: Calcular el promedio, la suma y el recuento de un campo por grupo

En el ejemplo se muestra cómo calcular la edad media, la edad sumada y el recuento de eventos de todas las cuentas, agrupadas por género.

```
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        |
+--------------------+------------+-----------+----------+
```

Ejemplo 5: Calcular el máximo de un campo

El ejemplo calcula la edad máxima de todas las cuentas.

```
os> source=accounts | stats max(age);
fetched rows / total rows = 1/1
+------------+
| max(age)   |
|------------|
| 36         |
+------------+
```

Ejemplo 6: Calcular el máximo y el mínimo de un campo por grupo

El ejemplo calcula los valores de edad máxima y mínima para todas las cuentas, agrupados por género.

```
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        |
```

```
+------------+-----------+----------+
```

Ejemplo 7: Calcular el recuento distinto de un campo

Para obtener el recuento de valores distintos de un campo, puede utilizar la función `DISTINCT_COUNT` (o `DC`) en lugar de `COUNT`. El ejemplo calcula tanto el recuento como el recuento distinto del campo de género de todas las cuentas.

```
os> source=accounts | stats count(gender), distinct_count(gender);
fetched rows / total rows = 1/1
+-----------------+--------------------------+
| count(gender)   | distinct_count(gender)   |
|-----------------+--------------------------|
| 4               | 2                        |
+-----------------+--------------------------+
```

Ejemplo 8: Calcular el recuento mediante un intervalo

El ejemplo obtiene el recuento de la edad en un intervalo de 10 años.

```
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         |
+--------------+------------+
```

Ejemplo 9: Calcule el recuento por género y rango

Este ejemplo cuenta los registros agrupados por género y edad en intervalos de 5 años.

```
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        |
+-------+------------+----------+
```

La expresión span siempre aparece como la primera clave de agrupamiento, independientemente del orden especificado en el 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        |
+-------+------------+----------+
```

Ejemplo 10: Calcule el recuento y obtenga la lista de correo electrónico por género y rango

El ejemplo obtiene el recuento de edad por un intervalo de 10 años y el grupo por género. Además, para cada fila, obtiene una lista de 5 correos electrónicos como máximo.

```
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        |
+-------+-----------------------------------------------------+------------+----------+
```

Ejemplo 11: Calcular el percentil de un campo

El ejemplo muestra cómo calcular el percentil de 90 de edad de todas las cuentas.

```
os> source=accounts | stats percentile(age, 90);
fetched rows / total rows = 1/1
+----------------------+
| percentile(age, 90)   |
|----------------------|
| 36                   |
+----------------------+
```

Ejemplo 12: Calcular el percentil de un campo por grupo

En el ejemplo se muestra cómo calcular el percentil de 90 de edad de todas las cuentas agrupadas por género.

```
os> source=accounts | stats percentile(age, 90) by gender;
fetched rows / total rows = 2/2
+----------------------+----------+
| percentile(age, 90)   | gender   |
|----------------------+----------|
| 28                    | F        |
| 36                    | M        |
+----------------------+----------+
```

Ejemplo 13: Calcule el percentil por género y rango

El ejemplo obtiene el percentil de 90 años de edad con un intervalo de 10 años y el grupo por género.

```
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)`
```

Agregaciones 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`
```

Agregaciones con intervalo de tiempo (función de ventana invertida)

```
- `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`
```

Las agregaciones se agrupan en varios niveles

```
- `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**
>
> Para ver qué integraciones de fuentes de AWS datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `subquery` para realizar consultas complejas y anidadas en sus sentencias del lenguaje de procesamiento canalizado (PPL).

```
source=logs | where field in [ subquery source=events | where condition | fields
 field ]
```

En este ejemplo, la búsqueda principal (`source=logs`) se filtra según los resultados de la subconsulta (`source=events`).

El comando subquery admite varios niveles de anidación para el análisis de datos complejos.

Ejemplo de subconsultas anidadas

```
source=logs | where id in [ subquery source=users | where user in [ subquery
 source=actions | where action="login" | fields user] | fields uid ]
```

InSubquery Uso

- `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 ]` (filtrado de búsqueda con subconsulta)
- `source = outer a not in [ source = inner | fields b ]` (filtrado de búsqueda con subconsulta)
- `source = outer | where a in [ source = inner1 | where b not in [ source = inner2 | fields c ] | fields b ]` (anidado)
- `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` (como filtro de unión)

Ejemplos de migración de SQL con PPL en la subconsulta

TPC-H Q4 (subconsulta interna con agregación)

```
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
```

Reescrito por una consulta PPL: InSubquery

```
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 (anidado en la subconsulta)

```
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
```

Reescrito por una consulta 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 uso

Supuestos :a, b son campos del exterior de la tabla c, d son campos del interior de la tabla e, f son campos del interior de la tabla2.

- `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 ]` (filtrado de búsqueda con subconsulta)
- `source = outer not exists [ source = inner | where a = c ]` (filtrado de búsqueda con subconsulta)

- `source = table as t1 exists [ source = table as t2 | where t1.a = t2.a ]` (el alias de la tabla es útil en una subconsulta existente)
- `source = outer | where exists [ source = inner1 | where a = c and exists [ source = inner2 | where c = e ] ]` (anidado)
- `source = outer | where exists [ source = inner1 | where a = c | where exists [ source = inner2 | where c = e ] ]` (anidado)
- `source = outer | where exists [ source = inner | where c > 10 ]` (existe una correlación no correlacionada)
- `source = outer | where not exists [ source = inner | where c > 10 ]` (existe una relación no correlacionada)
- `source = outer | where exists [ source = inner ] | eval l = "nonEmpty" | fields l` (existe una relación especial no correlacionada)

ScalarSubquery uso

Supuestos: `a` `b` son campos del exterior de la tabla `c`, `d` son campos del interior de la tabla `e`, `f` son campos de la tabla anidados

Subconsulta escalar no correlacionada

Seleccione en:

- `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`

En dónde:

- `source = outer | where a > [ source = inner | stats min(c) ] | fields a`

Búsqueda filtrada en:

- `source = outer a > [ source = inner | stats min(c) ] | fields a`

Subconsulta escalar correlacionada

Seleccione en:

- `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`

En dónde:

- `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`

Búsqueda filtrada en:

- `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`

Subconsulta escalar anidada

- `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 ]`

(Relación) Subconsulta

`InSubquery`, `ExistsSubquery` y `ScalarSubquery` son todas expresiones de subconsulta. Pero `RelationSubquery` no es una expresión de subconsulta, es un plan de subconsulta que se usa comúnmente en las cláusulas Join o From.

- `source = table1 | join left = l right = r [ source = table2 | where d > 10 | head 5 ]` (subconsulta en el lado derecho de la unión)
- `source = [ source = table1 | join left = l right = r [ source = table2 | where d > 10 | head 5 ] | stats count(a) by b ] as outer | head 1`

Contexto adicional

`InSubquery`, `ExistsSubquery` y `ScalarSubquery` son expresiones de subconsulta que se utilizan habitualmente en cláusulas `where` y filtros de búsqueda.

Comando Where:

```
| where <boolean expression> | ...
```

Filtro de búsqueda:

```
search source=* <boolean expression> | ...
```

Se puede usar una expresión de subconsulta en una expresión booleana:

```
| where orders.order_id in [ source=returns | where return_reason="damaged" | field
 order_id ]
```

El `orders.order_id in [ source=... ]` es un `<boolean expression>`.

En general, denominamos expresión `InSubquery` a este tipo de cláusula de subconsulta. Es una `<boolean expression>`.

Subconsulta con diferentes tipos de unión

En el ejemplo se utiliza 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 diferencia de InSubquery ExistsSubquery, y ScalarSubquery, a no RelationSubquery es una expresión de subconsulta. En cambio, es un plan de subconsultas.

```
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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `top` para buscar la tupla de valores más común de todos los campos de la lista de campos.

Sintaxis

Utilice la siguiente sintaxis:

```
top [N] <field-list> [by-clause] top_approx [N] <field-list> [by-clause]
```

N

- El número de resultados que se devolverá.
- Valor predeterminado: 10

field-list

- Obligatorio.
- Una lista delimitada por comas de nombres de campo.

by-clause

- Opcional.
- Uno o más campos por los que agrupar los resultados.

top_approx

- Un recuento aproximado de los (n) campos principales mediante el algoritmo de cardinalidad estimada mediante HyperLogLog ++.

Ejemplo 1: Encuentre los valores más comunes en un campo

En el ejemplo, se busca el género más común en todas las cuentas.

Consulta PPL:

```
os> source=accounts | top gender;
os> source=accounts | top_approx gender;
fetched rows / total rows = 2/2
+----------+
| gender   |
|----------|
| M        |
| F        |
+----------+
```

Ejemplo 2: busca los valores más comunes en un campo (limitado a 1)

En el ejemplo, se busca el género más común para todas las cuentas.

Consulta PPL:

```
os> source=accounts | top_approx 1 gender;
fetched rows / total rows = 1/1
+----------+
| gender   |
|----------|
| M        |
+----------+
```

Ejemplo 3: busca los valores más comunes agrupados por género

En el ejemplo, se busca la edad más común para todas las cuentas, agrupada por género.

Consulta 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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Utilice el comando `trendline` para calcular los promedios móviles de los campos.

Sintaxis

Utilice la siguiente sintaxis.

```
TRENDLINE [sort <[+|-] sort-field>] SMA(number-of-datapoints, field) [AS alias]
 [SMA(number-of-datapoints, field) [AS alias]]...
```

[+|-]

- Opcional.
- El signo más [+] significa orden ascendente con NULL/MISSING los valores en primer lugar.
- El signo menos [-] indica el orden descendente con los NULL/MISSING valores al final.
- Predeterminado: orden ascendente con NULL/MISSING los valores primero.

sort-field

- Obligatorio cuando se utiliza la clasificación.
- El campo que se utiliza para ordenar.

number-of-datapoints

- Obligatorio.
- El número de puntos de datos que calculan el promedio móvil.
- Debe ser mayor que cero.

campo

- Obligatorio.
- El nombre del campo para el que se debe calcular el promedio móvil.

alias

- Opcional.
- El nombre de la columna resultante que contiene el promedio móvil.

Solo se admite el tipo de promedio móvil simple (Simple Moving Average, SMA). Se calcula de la siguiente manera:

```
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
```

Ejemplo 1: Calcular el promedio móvil simple para una serie temporal de temperaturas

El ejemplo calcula el promedio móvil simple sobre las temperaturas utilizando dos puntos de datos.

Consulta 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|
+-----------+---------+--------------------+----------+
```

Ejemplo 2: Calcular promedios móviles simples para una serie temporal de temperaturas mediante la clasificación

El ejemplo calcula dos promedios móviles simples sobre las temperaturas utilizando dos y tres puntos de datos ordenados de forma descendente por el identificador del dispositivo.

Consulta 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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

El comando `where` utiliza una expresión booleana para filtrar el resultado de la búsqueda. Solo devuelve el resultado cuando la expresión booleana se evalúa como verdadera.

Sintaxis

Utilice la siguiente sintaxis:

```
where <boolean-expression>
```

bool-expression

- Opcional.
- Cualquier expresión que pueda evaluarse en un valor booleano.

Ejemplo 1: filtrado de conjunto de resultados con condición

El ejemplo muestra cómo obtener documentos del índice de cuentas que cumplan condiciones específicas.

Consulta 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        |
+------------------+----------+
```

Ejemplos adicionales

Filtros con condiciones lógicas

- `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: Esto devuelve un **>=** 1 y un **<=** 4, es decir, [1, 4]
- `source = table | where b not between '2024-09-10' and '2025-09-10'`, Nota: Esto devuelve b **>=** '**********' y 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
```

Resumen de campo

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten este comando PPL, consulte. the section called “Comandos”

Use el comando `fieldsummary` para calcular las estadísticas básicas de cada campo (count, distinct count, min, max, avg, stddev, mean) y determine el tipo de datos de cada campo. Este comando se puede utilizar con cualquier barra vertical anterior y las tendrá en cuenta.

Sintaxis

Utilice la siguiente sintaxis. Para CloudWatch los casos de uso de registros, solo se admite un campo en una consulta.

```
... | fieldsummary <field-list> (nulls=true/false)
```

includefields

- Lista de todas las columnas que se van a recopilar con las estadísticas en un conjunto de resultados unificado.

Nulos

- Opcional.
- Si se establece como verdadero, incluye valores nulos en los cálculos de agregación (sustituye el valor nulo por cero para los valores numéricos).

Ejemplo 1

Consulta 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"  |
+-----------------+-------------+-----------+-----------+-----------+-----------
+-----------+-----------+----------------|
```

Ejemplo 2

Consulta 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**
>
> Para ver qué integraciones AWS de fuentes de datos admiten esta función de PPL, consulte. the section called “Funciones”

Utilice el comando `expand` para aplanar un campo de tipo Array<Any> o Map<Any>, lo que genera filas individuales para cada elemento o par clave-valor.

Sintaxis

Utilice la siguiente sintaxis:

```
expand <field> [As alias]
```

campo

- El campo que se va a expandir (explotar).
- El campo debe ser de un tipo compatible.

alias

- Opcional.
- El nombre que se va a utilizar en lugar del nombre del campo original.

Directrices de uso

El comando expand genera una fila para cada elemento de la matriz o campo de mapa especificado, donde:

- Los elementos de la matriz se convierten en filas individuales.
- Los pares clave-valor de la asignación se dividen en filas independientes y cada valor clave se representa como una fila.
- Cuando se proporciona un alias, los valores desglosados se representan con el alias en lugar del nombre de campo original.

Puede usar este comando en combinación con otros comandos, como stats, eval y parse, para manipular o extraer datos después de la expansión.

Ejemplos

- `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`

Puede usar el comando expand en combinación con otros comandos, como eval, stats y más. El uso de varios comandos de expansión creará un producto cartesiano de todos los elementos internos de cada matriz o mapa compuesto.

Consulta desplegable de SQL eficaz

El comando expand se traduce en una operación SQL equivalente mediante LATERAL VIEW explode, lo que permite la fragmentación eficiente de matrices o mapas a nivel de consulta SQL.

```
SELECT customer exploded_productId
FROM table
LATERAL VIEW explode(productId) AS exploded_productId
```

El comando explode ofrece las siguientes funciones:

- Es una operación de columna que devuelve una columna nueva.
- Crea una nueva fila para cada elemento de la columna expandida.
- Los valores nulos internos se ignoran como parte del campo expandido (no hay filas created/exploded para valores nulos).

Funciones PPL

Temas

- Funciones de condiciones PPL
- Funciones hash criptográficas PPL
- Funciones de fecha y hora PPL
- Expresiones PPL

- Funciones PPL de dirección IP
- Nuevas funciones PPL JSON
- Funciones de Lambda PPL
- Funciones matemáticas PPL
- Funciones de cadena PPL
- Función de conversión de tipos PPL

Funciones de condiciones PPL

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten esta función PPL, consulte. the section called “Funciones”

ISNULL

Descripción: `isnull(field)` devuelve el valor verdadero si el campo es nulo.

Tipo de argumento:

- Cualquier tipo de datos compatible.

Tipo de devolución:

- BOOLEANO

Ejemplo:

```
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        |
+----------+-------------+-------------+
```

ISNOTNULL

Descripción: `isnotnull(field)` devuelve el valor verdadero si el campo no es nulo.

Tipo de argumento:

- Cualquier tipo de datos compatible.

Tipo de devolución:

- BOOLEANO

Ejemplo:

```
os> source=accounts | where not isnotnull(employer) | fields account_number, employer
fetched rows / total rows = 1/1
+------------------+------------+
| account_number   | employer   |
|------------------+------------|
| 18               | null       |
+------------------+------------+
```

EXISTS

Ejemplo:

```
os> source=accounts | where exists(email) | fields account_number, email
fetched rows / total rows = 1/1
```

IFNULL

Descripción: `ifnull(field1, field2)` devuelve `field2` si `field1` es nulo.

Tipo de argumento:

- Cualquier tipo de datos compatible.

- Si los dos parámetros son de tipos diferentes, la función no pasará la comprobación semántica.

Tipo de devolución:

- Cualquiera

Ejemplo:

```
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

Descripción: `nullif(field1, field2)` devuelve nulo si dos parámetros son iguales; de lo contrario, devuelve field1.

Tipo de argumento:

- Cualquier tipo de datos compatible.
- Si los dos parámetros son de tipos diferentes, la función no pasará la comprobación semántica.

Tipo de devolución:

- Cualquiera

Ejemplo:

```
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

Descripción: `if(condition, expr1, expr2)` devuelve `expr1` si la condición es verdadera; de lo contrario, devuelve `expr2`.

Tipo de argumento:

- Cualquier tipo de datos compatible.
- Si los dos parámetros son de tipos diferentes, la función no pasará la comprobación semántica.

Tipo de devolución:

- Cualquiera

Ejemplo:

```
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      |
+----------+-------------+------------+
```

Funciones hash criptográficas PPL

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten esta función PPL, consulte. the section called “Funciones”

MD5

MD5 calcula el MD5 resumen y devuelve el valor en forma de cadena hexadecimal de 32 caracteres.

Uso: `md5('hello')`

Tipo de argumento:

- STRING

Tipo de devolución:

- STRING

Ejemplo:

```
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 devuelve el resultado de la cadena hexadecimal de SHA-1.

Uso: `sha1('hello')`

Tipo de argumento:

- STRING

Tipo de devolución:

- STRING

Ejemplo:

```
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 devuelve el resultado de la cadena hexadecimal de la familia de funciones hash SHA-2 (SHA-224, SHA-256, SHA-384 y SHA-512). NumBits indica la longitud de bits deseada del resultado, que debe tener un valor de 224, 256, 384, 512

Uso:

- `sha2('hello',256)`
- `sha2('hello',512)`

Tipo de argumento:

- STRING, INTEGER

Tipo de devolución:

- STRING

Ejemplo:

```
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>                              |
+------------------------------------------------------------------+
```

Funciones de fecha y hora PPL

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten esta función PPL, consulte. the section called “Funciones”

DAY

Uso: `DAY(date)` extrae el día del mes para una fecha, en el rango del 1 al 31.

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: `DAYOFMONTH`, `DAY_OF_MONTH`

Ejemplo:

```
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

Uso: `DAYOFMONTH(date)` extrae el día del mes para una fecha, en el rango del 1 al 31.

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: `DAY`, `DAY_OF_MONTH`

Ejemplo:

```
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

Uso: DAY_OF_MONTH(DATE) extrae el día del mes para una fecha, en el rango del 1 al 31.

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: DAY, DAYOFMONTH

Ejemplo:

```
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

Uso: DAYOFWEEK(DATE) devuelve el índice de los días de la semana de una fecha (1 = domingo, 2 = lunes,..., 7 = sábado).

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: DAY_OF_WEEK

Ejemplo:

```
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

Uso: DAY_OF_WEEK(DATE) devuelve el índice de los días de la semana de una fecha (1 = domingo, 2 = lunes,..., 7 = sábado).

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: DAYOFWEEK

Ejemplo:

```
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

Uso: DAYOFYEAR(DATE) devuelve el día del año de una fecha, en el rango de 1 a 366.

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: DAY_OF_YEAR

Ejemplo:

```
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

Uso: DAY_OF_YEAR(DATE) devuelve el día del año de una fecha, en el rango de 1 a 366.

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: DAYOFYEAR

Ejemplo:

```
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

Uso: DAYNAME(DATE) devuelve el nombre del día de la semana para una fecha, incluidos lunes, martes, miércoles, jueves, viernes, sábado y domingo.

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: STRING

Ejemplo:

```
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

Uso: FROM_UNIXTIME devuelve una representación del argumento dado como marca de tiempo o valor de cadena de caracteres. Esta función realiza una conversión inversa de la función UNIX_TIMESTAMP.

Si proporciona un segundo argumento, FROM_UNIXTIME lo utiliza para dar al resultado un formato similar al de la función DATE_FORMAT.

Si la marca de tiempo está fuera del intervalo 1970-01-01 00:00:00 a 3001-01-18 23:59:59.999999 (0 a 32536771199.999999 tiempo de época), la función devuelve NULL.

Tipo de argumento: DOUBLE, STRING

Mapa del tipo de devolución:

DOUBLE -> TIMESTAMP

DOUBLE, STRING -> STRING

Ejemplos:

```
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

Uso: HOUR(TIME) extrae el valor hora del tiempo.

A diferencia de una hora del día estándar, el valor de hora de esta función puede tener un rango superior a 23. Como resultado, el valor devuelto de HOUR(TIME) puede ser superior a 23.

Tipo de argumento: STRING/TIME/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: HOUR_OF_DAY

Ejemplo:

```
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

Uso: HOUR_OF_DAY(TIME) extrae el valor de la hora del tiempo dado.

A diferencia de una hora del día estándar, el valor de hora de esta función puede tener un rango superior a 23. Como resultado, el valor devuelto de HOUR_OF_DAY(TIME) puede ser superior a 23.

Tipo de argumento: STRING/TIME/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: HOUR

Ejemplo:

```
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

Uso: LAST_DAY devuelve el último día del mes como un valor de fecha para el argumento de fecha indicado.

Tipo de argumento: DATE/STRING/TIMESTAMP/TIME

Tipo de devolución: DATE

Ejemplo:

```
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

Uso: LOCALTIMESTAMP() es un sinónimo de NOW().

Ejemplo:

```
> source=people | eval `LOCALTIMESTAMP()` = LOCALTIMESTAMP() | fields
 `LOCALTIMESTAMP()`
fetched rows / total rows = 1/1
+---------------------+
| LOCALTIMESTAMP()    |
|---------------------|
| 2022-08-02 15:54:19 |
+---------------------+
```

LOCALTIME

Uso: LOCALTIME() es un sinónimo de NOW().

Ejemplo:

```
> source=people | eval `LOCALTIME()` = LOCALTIME() | fields `LOCALTIME()`
fetched rows / total rows = 1/1
+---------------------+
| LOCALTIME()         |
|---------------------|
| 2022-08-02 15:54:19 |
+---------------------+
```

MAKE_DATE

Uso: MAKE_DATE devuelve un valor de fecha basado en los valores de año, mes y día dados. Todos los argumentos se redondean a números enteros.

Especificaciones: 1. MAKE_DATE(INTEGER, INTEGER, INTEGER) -> DATE

Tipo de argumento: INTEGER, INTEGER, INTEGER

Tipo de devolución: DATE

Ejemplo:

```
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

Uso: MINUTE(TIME) devuelve el componente de minutos del tiempo dado, como un número entero en el rango de 0 a 59.

Tipo de argumento: STRING/TIME/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: MINUTE_OF_HOUR

Ejemplo:

```
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

Uso: `MINUTE_OF_HOUR(TIME)` devuelve el componente de minutos del tiempo dado, como un número entero en el rango de 0 a 59.

Tipo de argumento: STRING/TIME/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: `MINUTE`

Ejemplo:

```
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

Uso: `MONTH(DATE)` devuelve el mes de la fecha dada como un entero, en el rango de 1 a 12 (donde 1 representa enero y 12 representa diciembre).

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: `MONTH_OF_YEAR`

Ejemplo:

```
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

Uso: `MONTHNAME(DATE)` devuelve el mes de la fecha dada como un entero, en el rango de 1 a 12 (donde 1 representa enero y 12 representa diciembre).

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: `MONTH_OF_YEAR`

Ejemplo:

```
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

Uso: `MONTH_OF_YEAR(DATE)` devuelve el mes de la fecha dada como un entero, en el rango de 1 a 12 (donde 1 representa enero y 12 representa diciembre).

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: `MONTH`

Ejemplo:

```
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

Uso: NOW devuelve la fecha y hora actuales como un TIMESTAMP valor en el formato 'YYYY-MM-DD hh:mm:ss'. El valor se expresa en la zona horaria del clúster.

> **Note**
>
> NOW( ) devuelve un tiempo constante que indica cuándo comenzó a ejecutarse la sentencia. Esto es diferente de SYSDATE( ), que devuelve la hora exacta de ejecución.

Tipo de devolución: TIMESTAMP

Especificación: NOW () -> TIMESTAMP

Ejemplo:

```
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

Uso: QUARTER(DATE) devuelve el trimestre del año de la fecha dada como un entero, en el rango de 1 a 4.

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Ejemplo:

```
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

Uso: `SECOND(TIME)` devuelve el segundo componente del tiempo dado como un entero, en el rango de 0 a 59.

Tipo de argumento: STRING/TIME/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: `SECOND_OF_MINUTE`

Ejemplo:

```
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

Uso: `SECOND_OF_MINUTE(TIME)` devuelve el segundo componente del tiempo dado como un entero, en el rango de 0 a 59.

Tipo de argumento: STRING/TIME/TIMESTAMP

Tipo de devolución: INTEGER

Sinónimos: SECOND

Ejemplo:

```
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

Uso: SUBDATE(DATE, DAYS) resta el segundo argumento (como DATE o DAYS) de la fecha dada.

Tipo de argumento: DATE/TIMESTAMP, LONG

Mapa de tipos de devolución: (DATE, LONG) -> DATE

Antónimos: ADDDATE

Ejemplo:

```
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

Uso: SYSDATE() devuelve la fecha y hora actuales como un TIMESTAMP valor en el formato 'YYYY-MM-DD hh:mm:ss.nnnnnn'.

SYSDATE() devuelve la hora exacta en la que se ejecuta. Esto difiere de NOW(), que devuelve un tiempo constante que indica cuándo comenzó a ejecutarse la sentencia.

Tipo de argumento opcional: INTEGER (0 a 6): especifica el número de dígitos para las fracciones de segundo del valor devuelto.

Tipo de devolución: TIMESTAMP

Ejemplo:

```
os> source=people | eval `SYSDATE()` = SYSDATE() | fields `SYSDATE()`
fetched rows / total rows = 1/1
+----------------------------+
| SYSDATE()                  |
|----------------------------|
| 2022-08-02 15:39:05.123456 |
+----------------------------+
```

`TIMESTAMP`

Uso: `TIMESTAMP(EXPR)` construye un tipo de marca de tiempo con la cadena `expr` de entrada como marca de tiempo.

Con un solo argumento, `TIMESTAMP(expr)` construye una marca de tiempo a partir de la entrada. Si `expr` es una cadena, se interpreta como una marca de tiempo. Para los argumentos que no son cadenas, la función convierte `expr` en una marca de tiempo utilizando la zona horaria UTC. Cuando `expr` es un valor `TIME`, la función aplica la fecha de hoy antes de la conversión.

Cuando se usa con dos argumentos, `TIMESTAMP(expr1, expr2)` agrega la expresión de hora (`expr2`) a la expresión de fecha o marca de tiempo (`expr1`) y devuelve el resultado como un valor de marca de tiempo.

Tipo de argumento: STRING/DATE/TIME/TIMESTAMP

Mapa del tipo de devolución:

(STRING/DATE/TIME/TIMESTAMP) -> MARCA DE TIEMPO

(STRING/DATE/TIME/TIMESTAMP, STRING/DATE/TIME/TIMESTAMP) -> MARCA DE TIEMPO

Ejemplo:

```
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

Uso: UNIX_TIMESTAMP convierte un argumento de fecha determinado a la hora de Unix (segundos desde la Época, que comenzó a principios de 1970). Si no se proporciona ningún argumento, devuelve la hora actual de Unix.

El argumento de fecha puede ser un DATE, una cadena TIMESTAMP o un número en uno de estos formatos: YYMMDD, YYMMDDhhmmss, YYYYMMDD, o YYYYMMDDhhmmss. Si el argumento incluye un componente de tiempo, puede incluir opcionalmente fracciones de segundo.

Si el formato del argumento no es válido o se encuentra fuera del intervalo comprendido entre 1970-01-01 00:00:00 a 3001-01-18 23:59:59.999999 (0 a 32536771199.999999 en tiempo de época), la función devuelve NULL.

La función acepta DATE, TIMESTAMP, o DOUBLE como tipos de argumentos o no acepta ningún argumento. Siempre devuelve un valor DOUBLE que representa la marca de tiempo de Unix.

Para la conversión inversa, puede utilizar la función FROM_UNIXTIME.

Tipo de argumento:/DOUBLE/DATE/TIMESTAMP<NONE>

Tipo de devolución: DOUBLE

Ejemplo:

```
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

Uso: `WEEK(DATE)` devuelve el número de semana de una fecha determinada.

Tipo de argumento: DATE/TIMESTAMP/STRING

Tipo de devolución: INTEGER

Sinónimos: `WEEK_OF_YEAR`

Ejemplo:

```
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

Uso: `WEEKDAY(DATE)` devuelve el índice del día de la semana correspondiente a la fecha (0 = lunes, 1 = martes,..., 6 = domingo).

Es similar a la función `dayofweek`, pero devuelve índices diferentes para cada día.

Tipo de argumento: STRING/DATE/TIME/TIMESTAMP

Tipo de devolución: INTEGER

Ejemplo:

```
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

Uso: `WEEK_OF_YEAR(DATE)` devuelve el número de la semana de la fecha indicada.

Tipo de argumento: DATE/TIMESTAMP/STRING

Tipo de devolución: INTEGER

Sinónimos: `WEEK`

Ejemplo:

```
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

Uso: `YEAR(DATE)` devuelve el año para la fecha, en el rango de 1000 a 9999, o 0 para la fecha "zero".

Tipo de argumento: STRING/DATE/TIMESTAMP

Tipo de devolución: INTEGER

Ejemplo:

```
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

Uso: `DATE_ADD(date, INTERVAL expr unit)` agrega el intervalo especificado a la fecha dada.

Tipo de argumento: DATE, INTERVAL

Tipo de devolución: DATE

Antónimos: `DATE_SUB`

Ejemplo:

```
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

Uso: `DATE_SUB(date, INTERVAL expr unit)` resta el intervalo de expiración de la fecha.

Tipo de argumento: DATE, INTERVAL

Tipo de devolución: DATE

Antónimos: `DATE_ADD`

Ejemplo:

```
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

Uso: devuelve un valor TIMESTAMP después de agregar un intervalo de tiempo específico a una fecha determinada.

Argumentos:

- intervalo: INTERVAL (SECOND, MINUTE, HOUR, DAY, WEEK, MONTH, QUARTER, YEAR)
- entero: INTEGER
- fecha: DATE, TIMESTAMP o STRING

Si proporciona un argumento de fecha STRING, formatéelo como válido TIMESTAMP. La función convierte automáticamente un argumento DATE en un TIMESTAMP.

Ejemplos:

```
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

Uso: TIMESTAMPDIFF(interval, start, end) devuelve la diferencia entre el inicio y el final date/times en unidades de intervalo especificadas.

Argumentos:

- intervalo: INTERVAL (SECOND, MINUTE, HOUR, DAY, WEEK, MONTH, QUARTER, YEAR)
- inicio: DATE, TIMESTAMP, o STRING
- final: DATE, TIMESTAMP, o STRING

La función convierte automáticamente los argumentos `TIMESTAMP` cuando sea apropiado. Formatee los argumentos `STRING` como válidos `TIMESTAMP`.

Ejemplos:

```
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`

Uso: `UTC_TIMESTAMP` devuelve la marca de tiempo UTC actual como un valor en 'YYYY-MM-DD hh:mm:ss'.

Tipo de devolución: TIMESTAMP

Especificación: UTC_TIMESTAMP() -> TIMESTAMP

Ejemplo:

```
> 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

Uso: CURRENT_TIMEZONE devuelve la zona horaria local actual.

Tipo de devolución: STRING

Ejemplo:

```
> source=people | eval `CURRENT_TIMEZONE()` = CURRENT_TIMEZONE() | fields
 `CURRENT_TIMEZONE()`
fetched rows / total rows = 1/1
+------------------------+
| CURRENT_TIMEZONE()     |
|------------------------|
| America/Chicago        |
+------------------------+
```

Expresiones PPL

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten esta función PPL, consulte. the section called “Funciones”

Las expresiones, en particular las expresiones de valores, devuelven un valor escalar. Las expresiones tienen distintos tipos y formas. Por ejemplo, hay valores literales como expresiones atómicas y expresiones aritméticas, de predicados y de funciones construidas sobre ellos. Puede utilizar expresiones en distintas cláusulas, como las expresiones aritméticas en los comandos `Filter` y `Stats`.

Operadores

Una expresión aritmética es una expresión formada por literales numéricos y operadores aritméticos binarios, de la siguiente manera:

1. +: suma.
2. -: resta.
3. *: multiplicación.
4. /: división (en el caso de los números enteros, el resultado es un número entero sin la parte fraccionaria)
5. %: módulo (se usa solo con números enteros; el resultado es el resto de la división)

Precedencia

Use paréntesis para controlar la precedencia de los operadores aritméticos. De lo contrario, los operadores de mayor prioridad se ejecutan primero.

Conversión de tipos

La conversión de tipos implícita se realiza al buscar las firmas de los operadores. Por ejemplo, en un entero + un número real coincide con la firma `+(double,double)`, lo que da como resultado un número real. Esta regla también se aplica a las llamadas de funciones.

Ejemplo de diferentes tipos de expresiones aritméticas:

```
os> source=accounts | where age > (25 + 5) | fields age ;
fetched rows / total rows = 3/3
+-------+
| age   |
|-------|
| 32    |
| 36    |
| 33    |
+-------+
```

Operadores de predicados

Un operador de predicado es una expresión que se evalúa como verdadera. La comparación entre el valor `MISSING` y el valor `NULL` sigue estas reglas:

- Un valor `MISSING` solo es igual a un valor `MISSING` y es menor que otros valores.

- Un valor `NULL` es igual a un valor `NULL`, es mayor que un valor `MISSING`, pero menor que todos los demás valores.

Operadores

Operadores de predicados

Name	Description (Descripción)
`>`	El operador mayor que
`>=`	Operador Mayor o igual que
`<`	El operador menor que
`!=`	Operador no igual
`<=`	Los operadores menor que o igual a
`=`	Operador igual
`LIKE`	Coincidencia de patrones simple
`IN`	Prueba de valor NULL
`AND`	Operador AND
`OR`	Operador OR
`XOR`	Operador XOR
`NOT`	Prueba de valor NOT NULL

Puede comparar fechas y horas. Al comparar diferentes tipos de fecha y hora (por ejemplo, `DATE` y `TIME`), ambos se convierten en `DATETIME`. Las siguientes reglas se aplican para la conversión:

- `TIME` se aplica a la fecha de hoy.
- `DATE` se interpreta a medianoche.

Operador de predicado básico

Ejemplo de operadores de comparación:

```
os> source=accounts | where age > 33 | fields age ;
fetched rows / total rows = 1/1
+-------+
| age   |
|-------|
| 36    |
+-------+
```

IN

Ejemplo del campo de prueba `IN` del operador en las listas de valores:

```
os> source=accounts | where age in (32, 33) | fields age ;
fetched rows / total rows = 2/2
+-------+
| age   |
|-------|
| 32    |
| 33    |
+-------+
```

OR

Ejemplo del operador `OR`:

```
os> source=accounts | where age = 32 OR age = 33 | fields age ;
fetched rows / total rows = 2/2
+-------+
| age   |
|-------|
| 32    |
| 33    |
+-------+
```

NOT

Ejemplo del operador `NOT`:

```
os> source=accounts | where age not in (32, 33) | fields age ;
fetched rows / total rows = 2/2
```

```
+-------+
| age   |
|-------|
| 36    |
| 28    |
+-------+
```

Funciones PPL de dirección IP

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten esta función PPL, consulte. the section called “Funciones”

`CIDRMATCH`

Uso: `CIDRMATCH(ip, cidr)` comprueba si la dirección IP especificada está dentro del rango de cidr dado.

Tipo de argumento:

- STRING, STRING
- Tipo de devolución: BOOLEAN

Ejemplo:

```
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`puede ser una dirección IPv4 o una IPv6 dirección.
> - `cidr`puede ser un bloque IPv4 o un IPv6 bloque.
> - `ip`y `cidr` deben ser ambos IPv4 o ambos IPv6.
> - `ip` y `cidr` deben ser ambos válidos y no estar vacíos ni ser nulos.

Nuevas funciones PPL JSON

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten esta función PPL, consulte. the section called “Funciones”

`JSON`

Uso: `json(value)` evalúa si una cadena se puede analizar en formato JSON. La función devuelve la cadena original si es un JSON válido o nula si no es válido.

Tipo de argumento: STRING

Tipo de devolución: STRING/NULL. Una expresión STRING de un formato de objeto JSON válido.

Ejemplos:

```
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

Uso: `json_object(<key>, <value>[, <key>, <value>]...)` devuelve un objeto JSON de los miembros de pares clave-valor.

Tipo de argumento:

- La **<key>** debe ser STRING.
- El **<value>** puede ser cualquier tipo de dato.

Tipo de devolución: JSON_OBJECT. StructType Expresión de un objeto JSON válido.

Ejemplos:

```
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

Uso: `json_array(<value>...)` crea una matriz JSON usando una lista de valores.

Tipo de argumento: un <value> puede ser cualquier tipo de valor, como una cadena, un número o un booleano.

Tipo de devolución: ARRAY. Una matriz de cualquier tipo de datos compatible para una matriz JSON válida.

Ejemplos:

```
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

Uso: to_json_string(jsonObject) devuelve una cadena JSON con un valor de objeto json determinado.

Tipo de argumento: JSON_OBJECT

Tipo de devolución: STRING

Ejemplos:

```
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

Uso: `array_length(jsonArray)` devuelve la cantidad de elementos en la matriz exterior.

Tipo de argumento: ARRAY. Un objeto ARRAY o JSON_ARRAY.

Tipo de devolución: INTEGER

Ejemplo:

```
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

Uso: `json_extract(jsonStr, path)` extrae un objeto JSON de una cadena JSON en función de la ruta JSON especificada. La función devuelve un valor nulo si la cadena JSON de entrada no es válida.

Tipo de argumento: STRING, STRING

Tipo de devolución: STRING

- Una expresión STRING de un formato de objeto JSON válido.
- NULL se devuelve en caso de un JSON no válido.

Ejemplos:

```
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

Uso: `json_keys(jsonStr)` devuelve todas las claves del objeto JSON más externo como una matriz.

Tipo de argumento: STRING. Una expresión STRING de un formato de objeto JSON válido.

Tipo de devolución: ARRAY[STRING]. La función devuelve NULL para cualquier otra cadena JSON válida, una cadena vacía o un JSON no válido.

Ejemplos:

```
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

Uso: `json_valid(jsonStr)` evalúa si una cadena JSON usa una sintaxis JSON válida y devuelve TRUE o FALSE.

Tipo de argumento: STRING

Tipo de devolución: BOOLEAN

Ejemplos:

```
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    |
+------------------+---------+
```

Funciones de Lambda PPL

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten esta función PPL, consulte. the section called “Funciones”

`EXISTS`

Uso: `exists(array, lambda)` evalúa si un predicado de Lambda es válido para uno o más elementos de la matriz.

Tipo de argumento: ARRAY, LAMBDA

Tipo de devolución: BOOLEAN. Devuelve `TRUE` si al menos un elemento de la matriz satisface el predicado Lambda; en caso contrario `FALSE`.

Ejemplos:

```
 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

Uso: `filter(array, lambda)` filtra la matriz de entrada utilizando la función de Lambda dada.

Tipo de argumento: ARRAY, LAMBDA

Tipo de devolución: ARRAY. Un ARRAY que contiene todos los elementos de la matriz de entrada que satisfacen el predicado lambda.

Ejemplos:

```
 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

Uso: `transform(array, lambda)` transforma los elementos de una matriz mediante la función de transformación de Lambda. El segundo argumento implica el índice del elemento si se utiliza la función de Lambda binaria. Esto es similar a lo que ocurre a `map` en la programación funcional.

Tipo de argumento: ARRAY, LAMBDA

Tipo de devolución: ARRAY. Un ARRAY que contiene el resultado de aplicar la función de transformación de Lambda a cada elemento de la matriz de entrada.

Ejemplos:

```
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

Uso: `reduce(array, start, merge_lambda, finish_lambda)` reduce una matriz a un único valor mediante la aplicación de funciones de Lambda. La función aplica la combinación merge_lambda al valor inicial y a todos los elementos de la matriz y, a continuación, la aplica al resultado `finish_lambda`.

Tipo de argumento: ARRAY, ANY, LAMBDA, LAMBDA

Tipo de devolución: ANY. El resultado final de aplicar las funciones de Lambda al valor inicial y a la matriz de entrada.

Ejemplos:

```
 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        |
+-----------+
```

Funciones matemáticas PPL

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten esta función PPL, consulte. the section called “Funciones”

ABS

Uso: `ABS(x)` calcula el valor absoluto de x.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: INTEGER/LONG/FLOAT/DOUBLE

Ejemplo:

```
os> source=people | eval `ABS(-1)` = ABS(-1) | fields `ABS(-1)`
fetched rows / total rows = 1/1
+-----------+
| ABS(-1)   |
|-----------|
```

```
| 1         |
+-----------+
```

ACOS

Uso: `ACOS(x)` calcula el arco coseno de x. Retorna `NULL` si x no está en el rango de -1 a 1.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `ACOS(0)` = ACOS(0) | fields `ACOS(0)`
fetched rows / total rows = 1/1
+--------------------+
| ACOS(0)            |
|--------------------|
| 1.5707963267948966 |
+--------------------+
```

ASIN

Uso: `asin(x)` calcula el arco sinusoidal de x. Retorna `NULL` si x no está en el rango de -1 a 1.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
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)` calcula el arco tangente de x. `atan(y, x)` Calcula el arco tangente de y/x, excepto que los signos de ambos argumentos determinan el cuadrante del resultado.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
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

Uso: `ATAN2(y, x)` calcula el arco tangente de y/x, excepto que los signos de ambos argumentos determinan el cuadrante del resultado.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
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

Uso: CBRT calcula la raíz cuadrada de un número.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE:

INTEGER/LONG/FLOAT/DOUBLE-> DOBLE

Ejemplo:

```
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

Uso: Un alias para la función `CEILING`. `CEILING(T)` toma el límite máximo del valor T.

Limitación: `CEILING` solo funciona según lo esperado cuando el tipo doble IEEE 754 muestra un decimal al guardarlo.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: LONG

Ejemplo:

```
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

Uso: `CONV(x, a, b)` convierte el número x de base a a una base b.

Tipo de argumento: x: STRING, a: INTEGER, b: INTEGER

Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `COS(x)` calcula el coseno de x, donde x se expresa en radianes.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `COS(0)` = COS(0) | fields `COS(0)`
fetched rows / total rows = 1/1
+----------+
| COS(0)   |
|----------|
| 1.0      |
+----------+
```

COT

Uso: `COT(x)` calcula la cotangente de x. Devuelve out-of-range un error si x es igual a 0.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `COT(1)` = COT(1) | fields `COT(1)`
fetched rows / total rows = 1/1
+--------------------+
| COT(1)             |
|--------------------|
| 0.6420926159343306 |
+--------------------+
```

CRC32

Uso: `CRC32` calcula un valor de comprobación de redundancia cíclica y devuelve un valor sin signo de 32 bits.

Tipo de argumento: STRING

Tipo de devolución: LONG

Ejemplo:

```
os> source=people | eval `CRC32('MySQL')` = CRC32('MySQL') | fields `CRC32('MySQL')`
fetched rows / total rows = 1/1
+------------------+
| CRC32('MySQL')   |
|------------------|
```

```
| 3259397556       |
+------------------+
```

DEGREES

Uso: `DEGREES(x)` convierte x de radianes a grados.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
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

Uso: `E()` devuelve el número de Euler.

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `E()` = E() | fields `E()`
fetched rows / total rows = 1/1
+------------------+
| E()              |
|------------------|
| 2.718281828459045 |
+------------------+
```

EXP

Uso: `EXP(x)` devuelve e elevado a x.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `EXP(2)` = EXP(2) | fields `EXP(2)`
fetched rows / total rows = 1/1
+------------------+
| EXP(2)           |
|------------------|
| 7.38905609893065 |
+------------------+
```

FLOOR

Uso: FLOOR(T) toma el mínimo del valor T.

Limitación: FLOOR solo funciona según lo esperado cuando el tipo doble IEEE 754 muestra un decimal cuando se almacena.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: LONG

Ejemplo:

```
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

Uso: `LN(x)` devuelve el logaritmo natural de x.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `LN(2)` = LN(2) | fields `LN(2)`
fetched rows / total rows = 1/1
+--------------------+
| LN(2)              |
|--------------------|
| 0.6931471805599453 |
+--------------------+
```

LOG

Uso: `LOG(x)` devuelve el logaritmo natural de x que es el logaritmo en base e de x. log(B, x) es equivalente a log(x)/log(B).

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
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

Uso: `LOG2(x)` equivale a `log(x)/log(2)`.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `LOG2(8)` = LOG2(8) | fields `LOG2(8)`
fetched rows / total rows = 1/1
+-----------+
| LOG2(8)   |
|-----------|
| 3.0       |
+-----------+
```

LOG10

Uso: `LOG10(x)` equivale a `log(x)/log(10)`.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `LOG10(100)` = LOG10(100) | fields `LOG10(100)`
fetched rows / total rows = 1/1
+--------------+
| LOG10(100)   |
|--------------|
| 2.0          |
+--------------+
```

MOD

Uso: `MOD(n, m)` calcula el resto del número n dividido por m.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: Tipo más amplio entre los tipos de n y m si m es un valor distinto de cero. Si m es igual a 0, devuelve NULL.

Ejemplo:

```
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

Uso: `PI()` devuelve la constante pi.

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `PI()` = PI() | fields `PI()`
fetched rows / total rows = 1/1
+-------------------+
| PI()              |
|-------------------|
| 3.141592653589793 |
```

```
+-------------------+
```

POW

Uso: `POW(x, y)` calcula el valor de x elevado a la potencia de y. Las entradas incorrectas devuelven un resultado `NULL`.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Sinónimos: `POWER(_, _)`

Ejemplo:

```
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

Uso: `POWER(x, y)` calcula el valor de x elevado a la potencia de y. Las entradas incorrectas devuelven un resultado `NULL`.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Sinónimos: `POW(_, _)`

Ejemplo:

```
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

Uso: `RADIANS(x)` convierte x de grados a radianes.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `RADIANS(90)` = RADIANS(90) | fields `RADIANS(90)`
fetched rows / total rows  = 1/1
+--------------------+
| RADIANS(90)        |
|--------------------|
| 1.5707963267948966 |
+--------------------+
```

RAND

Uso: `RAND()`/`RAND(N)` devuelve un valor de punto flotante aleatorio en el rango 0 <= value < 1,0. Si especifica el entero N, la función inicializa la semilla antes de la ejecución. Una consecuencia de este comportamiento es que, con un argumento N idéntico, `rand(N)` devuelve el mismo valor cada vez, lo que produce una secuencia repetible de valores de columna.

Tipo de argumento: INTEGER

Tipo de devolución: FLOAT

Ejemplo:

```
os> source=people | eval `RAND(3)` = RAND(3) | fields `RAND(3)`
fetched rows / total rows = 1/1
+------------+
| RAND(3)    |
|------------|
| 0.73105735 |
+------------+
```

ROUND

Uso: `ROUND(x, d)` redondea el argumento de x a d decimales. Si no especifica d, el valor predeterminado es 0.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Mapa del tipo de devolución:

- (INTEGER/LONG [,INTEGER]) -> LONG
- (FLOAT/DOUBLE [,INTEGER]) -> LONG

Ejemplo:

```
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

Uso: `SIGN` devuelve el signo del argumento como -1, 0 o 1, dependiendo de si el número es negativo, cero o positivo.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: INTEGER

Ejemplo:

```
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

Uso: `sin(x)` calcula el seno de x, donde x se expresa en radianes.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Tipo de devolución: DOUBLE

Ejemplo:

```
os> source=people | eval `SIN(0)` = SIN(0) | fields `SIN(0)`
fetched rows / total rows = 1/1
+----------+
| SIN(0)   |
|----------|
| 0.0      |
+----------+
```

SQRT

Uso: SQRT calcula la raíz cuadrada de un número no negativo.

Tipo de argumento: INTEGER/LONG/FLOAT/DOUBLE

Mapa del tipo de devolución:

- (No negativo) INTEGER/LONG/FLOAT/DOUBLE -> DOBLE
- (Negativo) INTEGER/LONG/FLOAT/DOUBLE -> NULO

Ejemplo:

```
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          |
+-----------+--------------+
```

Funciones de cadena PPL

> **Note**
>
> Para ver qué integraciones de fuentes de AWS datos admiten esta función PPL, consulte. the section called “Funciones”

CONCAT

Uso: `CONCAT(str1, str2, ...., str_9)` suma hasta 9 cadenas.

Tipo de argumento:

- STRING, STRING, ..., STRING
- Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `CONCAT_WS(sep, str1, str2)` concatena dos o más cadenas utilizando un separador específico entre ellas.

Tipo de argumento:

- STRING, STRING, ..., STRING
- Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `length(str)` devuelve la longitud de la cadena de entrada medida en bytes.

Tipo de argumento:

- STRING
- Tipo de devolución: INTEGER

Ejemplo:

```
os> source=people | eval `LENGTH('helloworld')` = LENGTH('helloworld') | fields
 `LENGTH('helloworld')`
fetched rows / total rows = 1/1
+------------------------+
| LENGTH('helloworld')   |
|------------------------|
| 10                     |
+------------------------+
```

LOWER

Uso: `lower(string)` convierte la cadena de caracteres en minúsculas.

Tipo de argumento:

- STRING
- Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `ltrim(str)` elimina los espacios iniciales de la cadena de entrada.

Tipo de argumento:

- STRING
- Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `POSITION(substr IN str)` devuelve la posición de la primera aparición de la subcadena en la cadena. Devuelve 0 si no se encuentra la subcadena dentro de la cadena. Devuelve un valor NULL si cualquier argumento de entrada es NULL.

Tipo de argumento:

- STRING, STRING
- Tipo de devolución INTEGER

Ejemplo:

```
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

Uso: `REVERSE(str)` devuelve la cadena invertida de la cadena de entrada.

Tipo de argumento:

- STRING
- Tipo de devolución: STRING

Ejemplo:

```
os> source=people | eval `REVERSE('abcde')` = REVERSE('abcde') | fields
 `REVERSE('abcde')`
fetched rows / total rows = 1/1
+--------------------+
| REVERSE('abcde')   |
|--------------------|
| edcba              |
+--------------------+
```

RIGHT

Uso: `right(str, len)` devuelve los caracteres situados más a la derecha de la cadena de entrada. Devuelve 0 si no se encuentra la subcadena dentro de la cadena. Devuelve un valor NULL si cualquier argumento de entrada es NULL.

Tipo de argumento:

- STRING, INTEGER
- Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `rtrim(str)` recorta los espacios finales de la cadena de entrada.

Tipo de argumento:

- STRING
- Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `substring(str, start)` o `substring(str, start, length)` devuelve una subcadena de la cadena de entrada. Sin especificar la longitud, devuelve la cadena completa desde la posición inicial.

Tipo de argumento:

- STRING, INTEGER, INTEGER
- Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `trim(string)` elimina los espacios en blanco iniciales y finales de la cadena de entrada.

Tipo de argumento:

- STRING
- Tipo de devolución: STRING

Ejemplo:

```
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

Uso: `upper(string)` convierte la cadena de entrada en mayúsculas.

Tipo de argumento:

- STRING
- Tipo de devolución: STRING

Ejemplo:

```
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            |
+----------------------+----------------------+
```

Función de conversión de tipos PPL

> **Note**
>
> Para ver qué integraciones AWS de fuentes de datos admiten esta función PPL, consulte. the section called “Funciones”

`TRIM`

Uso: `cast(expr as dateType)` convierte el valor `expr` en `dataType` y devuelve el valor del `dataType`.

Se aplican las siguientes reglas de conversión:

Reglas de conversión de tipos

Src/Target	STRING	NUMBER	BOOLEANO	TIMESTAMP	DATE	TIME
STRING		Note1	Note1	TIMESTAMP()	FECHA()	TIME()
NUMBER	Note1		v!=0	N/A	N/A	N/A

Src/Target	STRING	NUMBER	BOOLEANO	TIMESTAMP	DATE	TIME
BOOLEANO	Note1	v?1:0		N/A	N/A	N/A
TIMESTAMP	Note1	N/A	N/A		FECHA()	TIME()
DATE	Note1	N/A	N/A	N/A		N/A
TIME	Note1	N/A	N/A	N/A	N/A	

Ejemplo de conversión a cadena:

```
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 |
+---------+--------+------------+
```

Ejemplo de conversión a número:

```
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         |
+---------+-----------+
```

Ejemplo de conversión a fecha:

```
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 |
+------------+----------+---------------------+
```

Ejemplo de conversión a encadenado:

```
os> source=people | eval `cbool` = CAST(CAST(true as string) as boolean) | fields
 `cbool`
fetched rows / total rows = 1/1
+---------+
| cbool   |
|---------|
| True    |
+---------+
```

Supervisión de los dominios OpenSearch de Amazon Service

La supervisión es una parte importante para mantener la fiabilidad, la disponibilidad y el rendimiento de Amazon OpenSearch Service y tus demás AWS soluciones. AWS proporciona las siguientes herramientas para supervisar los recursos de su OpenSearch servicio, informar de problemas y tomar medidas automáticas cuando proceda:

Amazon CloudWatch

Amazon CloudWatch monitorea tus recursos de OpenSearch servicio en tiempo real. Puede recopilar métricas y realizar un seguimiento de ellas, crear paneles personalizados y definir alarmas que realizan advertencias o acciones cuando una métrica alcanza el umbral que se especifique. Para obtener más información, consulta la Guía del CloudWatch usuario de Amazon.

Amazon CloudWatch Logs

Amazon CloudWatch Logs te permite supervisar, almacenar y acceder a tus archivos de OpenSearch registro. CloudWatch Logs supervisa la información de los archivos de registro y puede notificarle cuando se alcanzan ciertos umbrales. Para obtener más información, consulta la Guía del usuario CloudWatch de Amazon Logs.

Amazon EventBridge

Amazon EventBridge ofrece una transmisión casi en tiempo real de los eventos del sistema que describen los cambios en sus dominios OpenSearch de servicio. Puede crear reglas que vigilen determinados eventos y activar acciones automatizadas en otros AWS servicios cuando se produzcan estos eventos. Para obtener más información, consulta la Guía del EventBridge usuario de Amazon.

AWS CloudTrail

AWS CloudTrail captura las llamadas a la API de configuración realizadas al OpenSearch Servicio como eventos. Puede entregar estos eventos al bucket de Amazon S3 que se especifique. Mediante esta información, puede identificar qué usuarios y cuentas realizaron solicitudes, la dirección IP fuente desde la que se realizaron las solicitudes y cuándo se produjeron. Para obtener más información, consulte la Guía del usuario de AWS CloudTrail .

Temas

- Supervisión de las métricas de los OpenSearch clústeres con Amazon CloudWatch
- Supervisión de OpenSearch registros con Amazon CloudWatch Logs
- Supervisión de los registros de auditoría en Amazon OpenSearch Service
- Monitorización OpenSearch de eventos del servicio con Amazon EventBridge
- Supervisión de las llamadas a la API de Amazon OpenSearch Service con AWS CloudTrail

Supervisión de las métricas de los OpenSearch clústeres con Amazon CloudWatch

Amazon OpenSearch Service publica los datos de tus dominios en Amazon CloudWatch. CloudWatch le permite recuperar las estadísticas sobre esos puntos de datos como un conjunto ordenado de datos de series temporales, conocidos como métricas. OpenSearch El servicio envía la mayoría de las métricas CloudWatch en intervalos de 60 segundos. Si utiliza los volúmenes de EBS magnéticos o de uso general, las métricas de volumen de EBS solo se actualizan cada cinco minutos. Todas las métricas acumuladas (por ejemplo, `ThreadpoolWriteRejected`, `ThreadpoolSearchRejected`) están en la memoria y perderán estado. Las métricas se restablecerán durante la caída de un nodo, el rebote de un nodo, el reemplazo de un nodo y blue/green el despliegue. Para obtener más información sobre Amazon CloudWatch, consulta la Guía del CloudWatch usuario de Amazon.

La consola OpenSearch de servicio muestra una serie de gráficos basados en los datos sin procesar de CloudWatch. Según sus necesidades, es posible que prefiera ver los datos del clúster en CloudWatch lugar de los gráficos de la consola. El servicio archiva las métricas durante dos semanas antes de eliminarlas. Las métricas se proporcionan sin coste adicional, pero CloudWatch siguen cobrando por la creación de paneles y alarmas. Para obtener más información, consulta los CloudWatchprecios de Amazon.

OpenSearch El servicio publica las siguientes métricas en CloudWatch:

- the section called “Métricas de clúster”
- the section called “Métricas de nodo maestro dedicado”
- the section called “Métricas de volumen de EBS”
- the section called “Métricas de la instancia”
- the section called “Métricas cálidas”

- the section called “Métricas de nodos coordinadores dedicados”
- the section called “Métricas de almacenamiento en frío”
- the section called “Métricas de alertas”
- the section called “Métricas de detección de anomalías”
- the section called “Métricas de búsqueda asíncrona”
- the section called “Métricas de SQL”
- the section called “Métricas k-NN”
- the section called “Métricas de búsqueda entre clústeres”
- the section called “Métricas de replicación entre clústeres”
- the section called “Aprender a clasificar métricas”
- the section called “Métricas del lenguaje de procesamiento de canalizaciones”

Visualización de las métricas en CloudWatch

CloudWatch Las métricas se agrupan primero por el espacio de nombres del servicio y, después, por las distintas combinaciones de dimensiones de cada espacio de nombres.

Para ver las métricas mediante la consola CloudWatch

1. Abra la CloudWatch consola en https://console.aws.amazon.com/cloudwatch/.
2. En el panel de navegación, busque Métricas y elija, Todas las métricas. Seleccione el espacio de OpenSearchService nombres ES/.
3. Seleccione una dimensión para ver las métricas correspondientes. Las métricas de los nodos individuales se encuentran en la dimensión `ClientId, DomainName, NodeId`. Las métricas de clúster se encuentran en la dimensión `Per-Domain, Per-Client Metrics`. Algunas métricas de nodos se agregan a nivel de clúster y, por lo tanto, se incluyen en ambas dimensiones. Las métricas de particiones se encuentran en la dimensión `ClientId, DomainName, NodeId, ShardRole`.

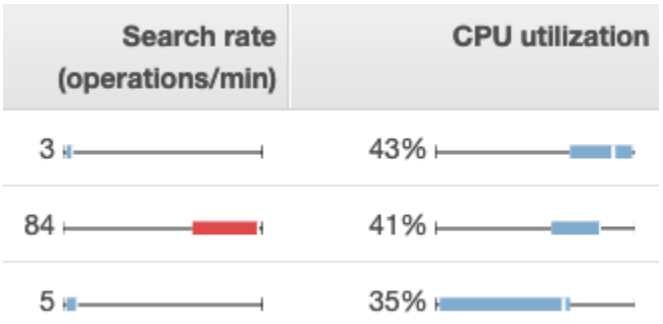
Para ver una lista de métricas mediante el AWS CLI

Use el siguiente comando:

```
aws cloudwatch list-metrics --namespace "AWS/ES"
```

Interpretación de las historias clínicas en OpenSearch Service

Para ver las métricas de OpenSearch Service, utilice las pestañas Estado del clúster y Estado de las instancias. La pestaña de estado de la instancia utiliza gráficos de recuadros para proporcionar at-a-glance visibilidad del estado de cada OpenSearch nodo:



- Cada cuadro de color muestra el rango de valores correspondiente al nodo durante el periodo de tiempo especificado.
- Los cuadros azules representan valores que son coherentes con otros nodos. Los cuadros rojos representan los valores atípicos.
- La línea blanca en cada cuadro muestra el valor actual del nodo.
- Los “bigotes” a cada lado del cuadro muestran los valores mínimo y máximo de todas las instancias a lo largo del periodo de tiempo.

Cuando se realizan cambios de configuración en el dominio, a menudo la lista de instancias individuales de las pestañas Estado del clúster e Estado de instancia duplica su tamaño durante un breve periodo antes de volver al número correcto. Para obtener una explicación de este comportamiento, consulte the section called “Cambios de configuración”.

Métricas de clúster

Amazon OpenSearch Service proporciona las siguientes métricas para los clústeres.

Métrica	Description (Descripción)
`ClusterStatus.green`	Un valor de 1 indica que todas las particiones de índice se han asignado a los nodos del clúster. Estadísticas pertinentes: máximo

Métrica	Description (Descripción)
`ClusterStatus.yellow`	El valor 1 indica que las particiones principales de todos los índices se han asignado a nodos del clúster, pero las particiones de réplica de al menos un índice están sin asignar. Para más información, consulte the section called “Estado amarillo del clúster”. Estadísticas pertinentes: Máximo
`ClusterStatus.red`	Un valor de 1 indica que las particiones principal y replicada de al menos un índice no se han asignado a los nodos del clúster. Para más información, consulte the section called “Estado rojo del clúster”. Estadísticas pertinentes: Máximo
`Shards.active`	El número total de particiones primarias y de réplicas activas. Estadísticas relevantes: Máximo, Suma
`Shards.unassigned`	El número de particiones que no se asignaron a los nodos del clúster. Estadísticas relevantes: Máximo, Suma
`Shards.delayedUnassigned`	El número de particiones cuya asignación de nodos se retrasó por la configuración de tiempo de espera. Estadísticas relevantes: Máximo, Suma
`Shards.activePrimary`	El número de particiones primarias activas. Estadísticas relevantes: Máximo, Suma
`Shards.initializing`	El número de particiones que se encuentran en inicialización. Estadísticas pertinentes: Suma
`Shards.relocating`	El número de particiones que se encuentran en reubicación. Estadísticas pertinentes: suma

Métrica	Description (Descripción)
`Nodes`	El número de nodos del clúster de OpenSearch servicios, incluidos los nodos principales dedicados y los nodos Warm. Para obtener más información, consulte the section called “Cambios de configuración”. Estadísticas pertinentes: Máximo
`SearchableDocuments`	El número total de documentos que admiten búsquedas para todos los nodos de datos del clúster. Estadísticas pertinentes: mínimo, máximo, promedio
`DeletedDocuments`	El número total de documentos marcados para su eliminación en todos los nodos de datos del clúster. Estos documentos ya no aparecen en los resultados de búsqueda, sino que OpenSearch solo eliminan los documentos eliminados del disco durante la combinación de segmentos. Esta métrica aumenta después de las solicitudes de eliminación y disminuye después de las combinaciones de segmentos. Estadísticas pertinentes: mínimo, máximo, promedio
`CPUUtilization`	El porcentaje de uso de CPU para los nodos de datos del clúster. Máximo muestra el nodo con el mayor uso de CPU. Promedio representa todos los nodos del clúster. Esta métrica también está disponible para los nodos individuales. Estadísticas relevantes: máximo, promedio

Métrica	Description (Descripción)
`FreeStorageSpace`	El espacio libre para los nodos de datos del clúster. `Sum` muestra el espacio libre total del clúster, pero debe dejar el periodo en un minuto para recibir un valor preciso. `Minimum` y `Maximum` muestran los nodos con el mínimo y máximo espacio libre, respectivamente. Esta métrica también está disponible para nodos individuales. OpenSearch El servicio arroja un `ClusterBlockException` valor cuando esta métrica alcanza`0`. Para efectuar la recuperación, debe eliminar índices, agregar instancias de mayor tamaño o bien agregar almacenamiento basado en EBS a las instancias existentes. Para obtener más información, consulte the section called “Falta de espacio de almacenamiento disponible”. La consola OpenSearch de servicio muestra este valor en GiB. La CloudWatch consola de Amazon lo muestra en MiB. **Note** `FreeStorageSpace` siempre será inferior a los valores proporcionados por OpenSearch `_cluster/stats` y `_cat/allocation` APIs . OpenSearch El servicio reserva un porcentaje del espacio de almacenamiento en cada instancia para las operaciones internas. Para más información, consulte Cálculo de requisitos de almacenamiento. Estadísticas pertinentes: mínimo, máximo, promedio, suma
`ClusterUsedSpace`	El espacio utilizado total para el clúster. Debe dejar el periodo en un minuto para obtener un valor preciso. La consola OpenSearch de servicio muestra este valor en GiB. La CloudWatch consola de Amazon lo muestra en MiB. Estadísticas pertinentes: mínimo, máximo

Métrica	Description (Descripción)
`ClusterIndexWritesBlocked`	Indica si el clúster acepta o bloquea las solicitudes de escritura entrantes. Un valor de 0 indica que el clúster acepta solicitudes. Un valor de 1 indica que el clúster bloquea las solicitudes. Algunos factores comunes son los siguientes: `FreeStorageSpace` es demasiado bajo o `JVMMemoryPressure` demasiado alta. Para solucionar este problema, considere la posibilidad de agregar más espacio de disco o de ampliar el clúster. Estadísticas pertinentes: máximo
`JVMMemoryPressure`	El porcentaje máximo del montón de Java utilizado para todos los nodos de datos del clúster. OpenSearch El servicio utiliza la mitad de la RAM de una instancia para el montón de Java, hasta un tamaño de pila de 32 GiB. Puede escalar las instancias verticalmente hasta 64 GiB de RAM y después escalarlas horizontalmente mediante el agregado de instancias. Consulte the section called “CloudWatch Alarmas recomendadas”. Estadísticas pertinentes: máximo **Note** La lógica de esta métrica cambió en el software del servicio R20220323. Para más información, consulte las notas de la versión.
`OldGenJVMMemoryPressure`	El porcentaje máximo de la pila de Java que se emplea para la “anterior generación” en todos los nodos de datos del clúster. Esta métrica también está disponible a nivel de nodo. Estadísticas pertinentes: máximo

Métrica	Description (Descripción)
AutomatedSnapshotFailure	El número de instantáneas automatizadas que han producido un error para el clúster. Un valor de 1 indica que no se ha tomado ninguna instantánea automatizada del dominio en las últimas 36 horas. Estadísticas pertinentes: mínimo, máximo
CPUCreditBalance	Los créditos de CPU restantes que están disponibles para los nodos de datos del clúster. Un crédito de CPU proporciona el rendimiento de un núcleo de CPU completo durante un minuto. Para más información, consulte Créditos CPU en la Guía para desarrolladores de Amazon EC2. Esta métrica solo está disponible para el tipo de instancias T2. Estadísticas pertinentes: mínimo
OpenSearchDashboardsHealthyNodes	Una comprobación del estado de los paneles de control. OpenSearch Si el mínimo, el máximo y el promedio son todos iguales a 1, Dashboards se comporta con normalidad. Si tiene 10 nodos con un máximo de 1, mínimo de 0 y promedio de 0,7, esto significa que 7 nodos (70 %) están en buen estado y 3 nodos (30 %) no lo están. Estadísticas pertinentes: mínimo, máximo, promedio
OpensearchDashboardsReportingFailedRequestSysErrCount	El número de solicitudes para generar informes de OpenSearch paneles de control que fallaron debido a problemas con el servidor o limitaciones de las funciones. Estadísticas pertinentes: suma
OpensearchDashboardsReportingFailedRequestUserErrCount	El número de solicitudes para generar informes de OpenSearch paneles de control que fallaron debido a problemas con los clientes. Estadísticas pertinentes: suma

Métrica	Description (Descripción)
OpensearchDashboardsReportingRequestCount	El número total de solicitudes para generar informes de OpenSearch paneles de control. Estadísticas pertinentes: suma
OpensearchDashboardsReportingSuccessCount	El número de solicitudes aceptadas para generar informes de OpenSearch paneles de control. Estadísticas pertinentes: suma
KMSKeyError	Un valor de 1 indica que la AWS KMS clave utilizada para cifrar los datos en reposo se ha desactivado. Para restablecer el dominio a operaciones normales, rehabilita la clave. La consola solo muestra esta métrica para los dominios que encriptan los datos en reposo. Estadísticas pertinentes: mínimo, máximo
KMSKeyInaccessible	Un valor de 1 indica que la AWS KMS clave utilizada para cifrar los datos en reposo se ha eliminado o se ha revocado su concesión al Servicio. OpenSearch No puede recuperar los dominios que están en este estado. Sin embargo, si tiene una instantánea manual, puede utilizarla para migrar los datos del dominio a un nuevo dominio. La consola solo muestra esta métrica para los dominios que encriptan los datos en reposo. Estadísticas pertinentes: mínimo, máximo

Métrica	Description (Descripción)
`InvalidHostHeaderRequests`	El número de solicitudes HTTP realizadas al OpenSearch clúster que incluían un encabezado de host no válido (o faltante). Las solicitudes válidas incluyen el nombre de host del dominio como valor del encabezado del host. OpenSearch El servicio rechaza las solicitudes no válidas de dominios de acceso público que no tengan una política de acceso restrictiva. Es recomendable que aplique una política de acceso restrictiva a todos los dominios. Si se muestran valores altos para esta métrica, confirme que sus clientes de OpenSearch incluyen el nombre de host del dominio (y no, por ejemplo, su dirección IP) en sus solicitudes. Estadísticas pertinentes: suma
`OpenSearchRequests (previously ElasticsearchRequests)`	El número de solicitudes realizadas al OpenSearch clúster. Estadísticas pertinentes: suma
`2xx, 3xx, 4xx, 5xx`	El número de solicitudes para el dominio que produjeron el código de respuesta HTTP especificado (2xx, 3xx, 4xx, 5xx). Estadísticas pertinentes: suma

Métrica	Description (Descripción)
`ThroughputThrottle`	Indica si los discos se han limitado o no. La limitación se produce cuando el rendimiento combinado de `ReadThroughputMicroBursting` y `WriteThroughputMicroBursting` es superior al rendimiento máximo, `MaxProvisionedThroughput`. `MaxProvisionedThroughput` es el valor más bajo del rendimiento de la instancia o del volumen aprovisionado. Un valor de 1 indica que los discos se han limitado. Un valor de 0 indica un comportamiento normal. Para obtener más información sobre el rendimiento de la instancia, consulte Instancias optimizadas para Amazon EBS. Para obtener información sobre el rendimiento del volumen, consulte Tipos de volúmenes de Amazon EBS. Estadísticas pertinentes: mínimo, máximo
`IopsThrottle`	Indica si se ha reducido o no el número de input/output operaciones por segundo (IOPS) en el dominio. La limitación se produce cuando las IOPS del nodo de datos superan el límite máximo permitido del volumen de EBS o de la instancia de EC2 del nodo de datos. Para obtener más información sobre IOPS de instancias, consulte Tipos de instancias optimizadas para Amazon EBS. Para obtener información sobre IOPS de volúmenes, consulte Amazon EBS volume types. Estadísticas pertinentes: mínimo, máximo
`HighSwapUsage`	Un valor de 1 indica que el intercambio debido a errores de página puede provocar picos en el uso del disco subyacente durante un periodo de tiempo específico. Estadísticas pertinentes: máximo

Métricas de nodo maestro dedicado

Amazon OpenSearch Service proporciona las siguientes métricas para los nodos maestros dedicados.

Métrica	Description (Descripción)
`MasterCPUUtilization`	El porcentaje máximo de recursos de CPU que se utilizan en los nodos maestros dedicados. Recomendamos aumentar el tamaño del tipo de instancia cuando esta métrica alcance el 60 por ciento. Estadísticas pertinentes: máximo
`MasterFreeStorageSpace`	Esta métrica no es pertinente y puede pasarse por alto. El servicio no usa los nodos maestros como nodos de datos.
`MasterJVMMemoryPressure`	El porcentaje máximo del montón de Java que se emplea para todos los nodos maestros dedicados del clúster. Recomendamos cambiar a un tipo de instancia mayor cuando esta métrica alcance el 85 por ciento. Estadísticas pertinentes: máximo **Note** La lógica de esta métrica cambió en el software del servicio R20220323. Para más información, consulte las notas de la versión.
`MasterOldGenJVMMemoryPressure`	El porcentaje máximo de la pila de Java que se emplea para la “anterior generación” según el nodo maestro. Estadísticas pertinentes: máximo
`MasterCPUCreditBalance`	Los créditos de CPU restantes que están disponibles para los nodos maestros dedicados del clúster. Un crédito de CPU proporciona el rendimiento de un núcleo de CPU completo durante un minuto. Para más información, consulte Créditos CPU en la

Métrica	Description (Descripción)
	Guía para desarrolladores de Amazon EC2. Esta métrica solo está disponible para el tipo de instancias T2. Estadísticas pertinentes: mínimo
`MasterReachableFromNode`	Comprobación de estado de las excepciones `MasterNotDiscovered` . Un valor de 1 indica un comportamiento normal. Un valor de 0 indica que `/_cluster/health/` ha dado error. Los errores significan que no se puede acceder al nodo maestro desde el nodo de origen. Suelen ser el resultado de un problema de conectividad de red o de AWS dependencia. Estadísticas pertinentes: máximo
`MasterSysMemoryUtilization`	El porcentaje de memoria del nodo maestro que está en uso. Estadísticas pertinentes: máximo

Métricas de nodos coordinadores dedicados

Amazon OpenSearch Service proporciona las siguientes métricas para los nodos coordinadores dedicados.

Métrica	Description (Descripción)
`CoordinatorCPUUtilization`	El porcentaje máximo de recursos de CPU que se utilizan en los nodos coordinadores dedicados. Recomendamos aumentar el tamaño del tipo de instancia cuando esta métrica alcance el 80 %. Estadísticas pertinentes: máximo
`CoordinatorJVMMemoryPressure`	El porcentaje máximo del montón de Java que se emplea para todos los nodos coordinadores dedicados del clúster. Recomendamos cambiar a un tipo de instancia mayor cuando esta métrica alcance el 85 por ciento.

Métrica	Description (Descripción)
	Estadísticas pertinentes: máximo
`CoordinatorOldGenJVMMemoryPressure`	El porcentaje máximo de la pila de Java que se emplea para la “anterior generación” según el nodo maestro. Estadísticas pertinentes: máximo
`CoordinatorSysMemoryUtilization`	El porcentaje de la memoria del nodo coordinador que está en uso. Estadísticas pertinentes: máximo
`CoordinatorFreeStorageSpace`	Esta métrica indica que el servicio no usa los nodos coordinadores como nodos de datos.

Métricas de volumen de EBS

Amazon OpenSearch Service proporciona las siguientes métricas para los volúmenes de EBS.

Métrica	Description (Descripción)
`ReadLatency`	La latencia en segundos para las operaciones de lectura en los volúmenes de EBS. Esta métrica también está disponible para los nodos individuales. Estadísticas pertinentes: mínimo, máximo, promedio
`WriteLatency`	La latencia en segundos para las operaciones de escritura en los volúmenes de EBS. Esta métrica también está disponible para los nodos individuales. Estadísticas pertinentes: mínimo, máximo, promedio
`ReadThroughput`	El rendimiento en bytes por segundo para las operaciones de lectura en los volúmenes de EBS. Esta métrica también está disponible para los nodos individuales. Estadísticas pertinentes: mínimo, máximo, promedio

Métrica	Description (Descripción)
`ReadThroughputMicroBursting`	Se tiene en cuenta el rendimiento, en bytes por segundo, de las operaciones de lectura en volúmenes de EBS cuando se tiene en cuenta la microrráfaga. Esta métrica también está disponible para los nodos individuales. La microrráfaga se produce cuando un volumen de EBS alcanza niveles altos de IOPS o rendimiento durante períodos de tiempo significativamente más cortos (menos de un minuto). Estadísticas pertinentes: mínimo, máximo, promedio
`WriteThroughput`	El rendimiento en bytes por segundo para las operaciones de escritura en los volúmenes de EBS. Esta métrica también está disponible para los nodos individuales. Estadísticas pertinentes: mínimo, máximo, promedio
`WriteThroughputMicroBursting`	Se tiene en cuenta el rendimiento, en bytes por segundo, de las operaciones de escritura en volúmenes de EBS cuando se tiene en cuenta la microrráfaga. Esta métrica también está disponible para los nodos individuales. La microrráfaga se produce cuando un volumen de EBS alcanza niveles altos de IOPS o rendimiento durante períodos de tiempo significativamente más cortos (menos de un minuto). Estadísticas pertinentes: mínimo, máximo, promedio
`DiskQueueDepth`	El número de solicitudes de entrada y salida (E/S) pendientes de un volumen de EBS. Estadísticas pertinentes: mínimo, máximo, promedio
`ReadIOPS`	El número de operaciones de entrada y salida (E/S) por segundo para las operaciones de lectura en los volúmenes de EBS. Esta métrica también está disponible para los nodos individuales. Estadísticas pertinentes: mínimo, máximo, promedio

Métrica	Description (Descripción)
`ReadIOPSMicroBursting`	El número de operaciones de entrada y salida (E/S) por segundo para las operaciones de lectura en los volúmenes de EBS cuando se tiene en cuenta la microrráfaga. Esta métrica también está disponible para los nodos individuales. La microrráfaga se produce cuando un volumen de EBS alcanza niveles altos de IOPS o rendimiento durante períodos de tiempo significativamente más cortos (menos de un minuto). Estadísticas pertinentes: mínimo, máximo, promedio
`WriteIOPS`	El número de operaciones de entrada y salida (E/S) por segundo para las operaciones de escritura en los volúmenes de EBS. Esta métrica también está disponible para los nodos individuales. Estadísticas pertinentes: mínimo, máximo, promedio
`WriteIOPS MicroBursting`	El número de operaciones de entrada y salida (E/S) por segundo para las operaciones de escritura en los volúmenes de EBS cuando se tiene en cuenta la microrráfaga. Esta métrica también está disponible para los nodos individuales. La microrráfaga se produce cuando un volumen de EBS alcanza niveles altos de IOPS o rendimiento durante períodos de tiempo significativamente más cortos (menos de un minuto). Estadísticas pertinentes: mínimo, máximo, promedio
`BurstBalance`	El porcentaje de los créditos de entrada y salida (E/S) que quedan en el bucket de ráfaga para un volumen de EBS. Un valor de 100 significa que el volumen ha acumulado el número máximo de créditos. Si este porcentaje cae por debajo del 70 %, consulte the section called “Bajo balance de ráfaga EBS”. El balance de ráfagas se mantiene en 0 para los dominios con tipos de volúmenes gp3, así como para los dominios con volúmenes gp2 que tengan un tamaño de volumen superior a 1000 GiB. Estadísticas pertinentes: mínimo, máximo, promedio

Métrica	Description (Descripción)
`VolumeStalledIOcheck`	El estado de sus volúmenes EBS para determinar cuándo están deteriorados. La métrica es un valor binario que devuelve un estado 0 (aprobado) o 1 (fallido) según si el volumen EBS puede completar las operaciones de entrada y salida. `VolumeStalledIOcheck` también está disponible para nodos individuales. Estadísticas pertinentes: mínimo, máximo, promedio

Métricas de la instancia

Amazon OpenSearch Service proporciona las siguientes métricas para cada instancia de un dominio. OpenSearch El servicio también agrega estas métricas de instancias para proporcionar información sobre el estado general del clúster. Puede verificar este comportamiento mediante la estadística Recuento de muestras en la consola. Tenga en cuenta que las métricas de la tabla siguiente disponen de estadísticas relevantes para el nodo y el clúster.

> **Important**
>
> Las diferentes versiones de Elasticsearch utilizan diferentes grupos de subprocesos para procesar las llamadas a la API `_index`. Elasticsearch 1.5 y 2.3 utilizan el grupo de subprocesos de indexación. Elasticsearch 5. x, 6.0 y 6.2 utilizan el grupo de subprocesos masivos. OpenSearch y Elasticsearch 6.3 y versiones posteriores usan el grupo de subprocesos de escritura. Actualmente, la consola OpenSearch de servicio no incluye un gráfico para el grupo de subprocesos masivos.
> Utilice `GET _cluster/settings?include_defaults=true` para verificar el tamaño del grupo de subprocesos y las colas para el clúster.

Métrica	Description (Descripción)
`FetchLatency`	La diferencia en el tiempo total, en milisegundos, que tardan todas las operaciones de obtención de particiones en un nodo entre el minuto N y el minuto (N - 1). Estadísticas de nodo pertinentes: promedio

Métrica	Description (Descripción)
	Estadísticas de clúster pertinentes: promedio, máximo
`FetchRate`	El número total de operaciones de obtención de particiones por minuto para todos los particiones en un nodo de datos. Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo, suma
`ScrollTotal`	El número total de operaciones de desplazamiento de particiones por minuto para todas las particiones en un nodo de datos. Estadísticas de nodo relevantes: Promedio, Máximo Estadísticas de clúster pertinentes: promedio, máximo, suma
`ScrollCurrent`	El número de operaciones de desplazamiento de particiones que se están ejecutando actualmente. Estadísticas de nodo relevantes: Promedio, Máximo Estadísticas de clúster pertinentes: promedio, máximo, suma
`OpenContexts`	El número de contextos de búsqueda abiertos. Estadísticas de nodo relevantes: Promedio, Máximo Estadísticas de clúster pertinentes: promedio, máximo, suma
`ThreadCount`	El número total de subprocesos que utiliza actualmente el OpenSearch proceso. Estadísticas de nodo relevantes: Promedio, Máximo Estadísticas de clúster pertinentes: promedio, máximo, suma

Métrica	Description (Descripción)
`ShardReactivateCount`	El número total de veces que todos los particiones se han activado desde un estado inactivo. Estadísticas de nodo relevantes: Suma, Máximo Estadísticas de clúster relevantes: Suma, Máximo
`ConcurrentSearchRate`	El número total de solicitudes de búsqueda con búsqueda de segmentos simultánea por minuto de todas las particiones de un nodo de datos. Una sola llamada a la API `_search` podría devolver resultados de muchas particiones diferentes. Si cinco de estas particiones se encuentran en un nodo, este notificaría el valor 5 para la métrica, incluso aunque el cliente solo realizara una solicitud. Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo, suma
`ConcurrentSearchLatency`	La diferencia en el tiempo total, en milisegundos, calculada por todas las búsquedas que usan la búsqueda de segmentos simultánea en un nodo entre el minuto N y el minuto (N-1). Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo
`IndexingLatency`	La diferencia en el tiempo total, en milisegundos, calculada por todas las operaciones de indexación en un nodo entre el minuto N y el minuto (N-1). Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo

Métrica	Description (Descripción)
`IndexingRate`	El número de operaciones de indexación por minuto. Una sola llamada a la API `_bulk` que agrega dos documentos y actualiza dos recuentos como cuatro operaciones, que podrían propagars e por uno o varios nodos. Si ese índice tiene una o más réplicas y se encuentra en un OpenSearch dominio sin instancias optimizad as, los demás nodos del clúster también registran un total de cuatro operaciones de indexación. En el caso de OpenSearch los dominios con instancias optimizadas, los demás nodos con réplicas no registran ninguna operación. Las eliminaciones de documentos no se tienen en cuenta en esta métrica. Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo, suma
`SearchLatency`	La diferencia en el tiempo total, en milisegundos, calculada por todas las búsquedas en un nodo entre el minuto N y el minuto (N-1). Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo
`SearchRate`	El número total de peticiones de búsqueda por minuto de todas las particiones de un nodo de datos. Una sola llamada a la API `_search` podría devolver resultados de muchas particiones diferentes. Si cinco de estas particiones se encuentran en un nodo, este notificaría el valor 5 para la métrica, incluso aunque el cliente solo realizara una solicitud. Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo, suma

Métrica	Description (Descripción)
`SegmentCount`	El número de segmentos de un nodo de datos. Cuantos más segmentos tenga, más tardará cada búsqueda. OpenSearch de vez en cuando fusiona segmentos más pequeños en uno más grande. Estadísticas de nodo pertinentes: máximo, promedio Estadísticas de clúster pertinentes: suma, máximo, promedio
`SysMemoryUtilization`	El porcentaje de memoria de la instancia que está en uso. Es normal obtener valores altos para esta métrica y, por lo general, no representan un problema con el clúster. Para obtener un mejor indicador de posibles problemas de rendimiento y estabilidad, consulte la métrica `JVMMemoryPressure` . Estadísticas de nodo pertinentes: mínimo, máximo, promedio Estadísticas de clúster pertinentes: mínimo, máximo, promedio
`JVMGCYoungCollectionCount`	Número de veces que se ha ejecutado la recopilación de elementos no utilizados “nueva generación”. Un número grande y en continuo crecimiento de ejecuciones es normal para las operaciones de clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
`JVMGCYoungCollectionTime`	La cantidad de tiempo en milisegundos que el clúster ha empleado en realizar la recopilación de elementos no utilizados “nueva generación”. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio

Métrica	Description (Descripción)
JVMGCOldCollectionCount	Número de veces que se ha ejecutado la recopilación de elementos no utilizados “generación anterior”. En un clúster con suficientes recursos, este número debe ser pequeño y crecer con poca frecuencia. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
JVMGCOldCollectionTime	La cantidad de tiempo en milisegundos que el clúster ha empleado en realizar la recopilación de elementos no utilizados “generación anterior”. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
OpenSearchDashboardsConcurrentConnections	El número de conexiones simultáneas activas a OpenSearch los paneles de control. Si este número crece continuamente, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
OpenSearchDashboardsHealthyNode	Una comprobación del estado del nodo individual de los OpenSearch paneles de control. Un valor de 1 indica un comportamiento normal. Un valor de 0 indica que Dashboards es inaccesible. Estadísticas de nodo pertinentes: mínimo Estadísticas de clúster pertinentes: mínimo, máximo, promedio

Métrica	Description (Descripción)
OpenSearchDashboardsHeapTotal	La cantidad de memoria dinámica asignada a los OpenSearch paneles en MiB. Los diferentes tipos de instancias EC2 pueden afectar a la asignación exacta de memoria. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
OpenSearchDashboardsHeapUsed	La cantidad absoluta de memoria dinámica que utilizan los OpenSearch paneles de control en MiB. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
OpenSearchDashboardsHeapUtilization	El porcentaje máximo de memoria de pila disponible que utilizan los paneles. OpenSearch Si este valor aumenta por encima del 80 %, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: mínimo, máximo, promedio
OpenSearchDashboardsOS1MinuteLoad	El promedio de carga de CPU en un minuto para los paneles. OpenSearch La carga de la CPU idealmente debería permanecer por debajo de 1,00. Aunque los picos temporales están bien, se recomienda aumentar el tamaño del tipo de instancias si esta métrica está constantemente por encima de 1,00. Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo

Métrica	Description (Descripción)
`OpenSearchDashboardsRequestTotal`	El recuento total de solicitudes HTTP realizadas a OpenSearch Dashboards. Si su sistema es lento o ve un número elevado de solicitudes de Dashboards, considere aumentar el tamaño del tipo de instancias. Estadísticas de nodos pertinentes: suma Estadísticas de clúster pertinentes: suma
`OpenSearchDashboardsResponseTimesMaxInMillis`	El tiempo máximo, en milisegundos, que tarda OpenSearch Dashboards en responder a una solicitud. Si las solicitudes tardan mucho tiempo en devolver resultados, considere aumentar el tamaño del tipo de instancias. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: máximo, promedio
`SearchTaskCancelled`	El número de cancelaciones del nodo coordinador. Estadísticas de nodos pertinentes: suma Estadísticas de clúster pertinentes: suma
`SearchShardTaskCancelled`	El número de cancelaciones del nodo de datos. Estadísticas de nodos pertinentes: suma Estadísticas de clúster pertinentes: suma,
`ThreadpoolForce_mergeQueue`	El número de tareas en cola en el grupo de subprocesos de combinación forzada. Si el tamaño de la cola es sistemáticamente elevado, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio

Métrica	Description (Descripción)
ThreadpoolForce_mergeRejected	El número de tareas rechazadas en el grupo de subprocesos de combinación forzada. Si este número crece continuamente, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma
ThreadpoolForce_mergeThreads	El tamaño del grupo de subprocesos de combinación forzada. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma
ThreadpoolIndexQueue	El número de tareas en cola en el grupo de subprocesos de indexación. Si el tamaño de la cola es sistemáticamente elevado, considere la posibilidad de escalar el clúster. El tamaño de la cola de indexación máximo es 200. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
ThreadpoolIndexRejected	El número de tareas rechazadas en el grupo de subprocesos de indexación. Si este número crece continuamente, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma
ThreadpoolIndexThreads	El tamaño del grupo de subprocesos de indexación. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma

Métrica	Description (Descripción)
ThreadpoolSearchQueue	El número de tareas en cola en el grupo de subprocesos de búsqueda. Si el tamaño de la cola es sistemáticamente elevado, considere la posibilidad de escalar el clúster. El tamaño máximo de la cola de búsqueda es 1000. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
ThreadpoolSearchRejected	El número de tareas rechazadas en el grupo de subprocesos de búsqueda. Si este número crece continuamente, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma
ThreadpoolSearchThreads	El tamaño del grupo de subprocesos de búsqueda. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma
Threadpoolsql-workerQueue	El número de tareas en cola en el grupo de subprocesos de búsqueda de SQL. Si el tamaño de la cola es sistemáticamente elevado, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
Threadpoolsql-workerRejected	El número de tareas rechazadas en el grupo de subprocesos de búsqueda de SQL. Si este número crece continuamente, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma

Métrica	Description (Descripción)
`Threadpoolsql-workerThreads`	El tamaño del grupo de subprocesos de búsqueda de SQL. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma
`ThreadpoolBulkQueue`	El número de tareas en cola en el grupo de subprocesos por lotes. Si el tamaño de la cola es sistemáticamente elevado, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
`ThreadpoolBulkRejected`	El número de tareas rechazadas en el grupo de subprocesos por lotes. Si este número crece continuamente, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma
`ThreadpoolBulkThreads`	El tamaño del grupo de subprocesos por lotes. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma
`ThreadpoolIndexSearcherQueue`	El número de tareas en cola del grupo de subprocesos del buscador de índices. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio

Métrica	Description (Descripción)
ThreadpoolIndexSearcherRejected	El número de tareas rechazadas en el grupo de subprocesos del buscador de índices. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma
ThreadpoolIndexSearcherThreads	El tamaño del grupo de subprocesos del buscador de índices. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma
ThreadpoolWriteThreads	El tamaño del grupo de subprocesos de escritura. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma
ThreadpoolWriteQueue	El número de tareas en cola en el grupo de subprocesos de escritura. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma

Métrica	Description (Descripción)
`ThreadpoolWriteRejected`	El número de tareas rechazadas en el grupo de subprocesos de escritura. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma **Note** Como el tamaño predeterminado de la cola de escritura se incrementó de 200 a 10 000 en la versión 7.1, esta métrica ya no es el único indicador de los rechazos de Service. OpenSearch Utilice las métricas `CoordinatingWriteRejected` , `PrimaryWriteRejected` y `ReplicaWriteRejected` para monitorear los rechazos en las versiones 7.1 y posteriores.
`CoordinatingWriteRejected`	El número total de rechazos se produjo en el nodo coordinador debido a la presión de indexación desde el último OpenSearch inicio del proceso de servicio. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma Esta métrica está disponible en la versión 7.1 y posteriores.
`PrimaryWriteRejected`	El número total de rechazos se produjo en los fragmentos principales debido a la presión de indexación desde el último inicio del proceso de servicio. OpenSearch Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma Esta métrica está disponible en la versión 7.1 y posteriores.

Métrica	Description (Descripción)
`ReplicaWriteRejected`	El número total de rechazos que se produjeron en los fragmentos de réplica se debió a la presión ejercida por la indexación desde el último inicio del proceso de servicio. OpenSearch Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma Esta métrica está disponible en la versión 7.1 y posteriores.
`WorkloadManagementEnabled`	Indica si la característica de administración de cargas de trabajo está habilitada. Un valor de 1 significa que está habilitada y un valor de 0 significa que está `monitor_only` o deshabilitada. Estadísticas de nodo relevantes: Máximo, Mínimo Estadísticas de clúster pertinentes: promedio, suma Esta métrica está disponible en la versión 7.1 y posteriores.
`SoftQueryGroupCount`	Número de grupos de consultas en modo flexible en el dominio. Estadísticas de nodo relevantes: Promedio, Máximo Estadísticas de clúster pertinentes: promedio, máximo, suma Esta métrica está disponible en la versión 7.1 y posteriores.
`EnforcedQueryGroupCount`	Número de grupos de consultas en modo aplicado en el dominio. Estadísticas de nodo relevantes: Promedio, Máximo Estadísticas de clúster pertinentes: promedio, máximo, suma Esta métrica está disponible en la versión 7.1 y posteriores.

Métricas cálidas

Amazon OpenSearch Service proporciona las siguientes métricas para la arquitectura de almacenamiento de varios niveles y UltraWarm

> **Note**
>
> Las métricas relacionadas con la indexación en caliente solo son aplicables a la arquitectura de almacenamiento de varios niveles

Métrica	Description (Descripción)
`WarmIndexingLatency`	La diferencia en el tiempo total, en milisegundos, calculada por todas las operaciones de indexación en un nodo caliente entre el minuto N y el minuto (N-1). Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo
`WarmIndexingRate`	El número de operaciones de indexación en caliente por minuto. Una sola llamada a la API `_bulk` que agrega dos documentos y actualiza dos recuentos como cuatro operaciones, que podrían propagarse por uno o varios nodos. Si ese índice tiene una o más réplicas y se encuentra en un OpenSearch dominio sin instancias optimizadas, los demás nodos del clúster también registran un total de cuatro operacion es de indexación. En el caso de OpenSearch los dominios con instancia s optimizadas, los demás nodos con réplicas no registran ninguna operación. Las eliminaciones de documentos no se tienen en cuenta en esta métrica. Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo, suma
`WarmThreadpoolIndexingQueue`	El número de tareas en cola en el grupo de subprocesos de indexació n. Si el tamaño de la cola es sistemáticamente elevado, considere la

Métrica	Description (Descripción)
	posibilidad de escalar el clúster. El tamaño de la cola de indexación máximo es 200. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, máximo, suma
`WarmThreadpoolIndexingRejected`	El número de tareas rechazadas en el grupo de subprocesos de indexación. Si este número crece continuamente, considere la posibilidad de escalar el clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma
`WarmThreadpoolIndexingThreads`	El tamaño del grupo de subprocesos de indexación. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma
`WarmCPUUtilization`	El porcentaje de uso de la CPU para los nodos Warm del clúster. Máximo muestra el nodo con el mayor uso de CPU. El promedio representa todos los nodos cálidos del clúster. Esta métrica también está disponible para los nodos Warm individuales. Estadísticas relevantes: máximo, promedio
`WarmFreeStorageSpace`	La cantidad de espacio de almacenamiento en caliente libre en MiB. Dado que Warm usa Amazon S3 en lugar de discos conectados, `Sum` es la única estadística relevante. Debe dejar el periodo en un minuto para obtener un valor preciso. Estadísticas pertinentes: suma

Métrica	Description (Descripción)
`WarmSearchableDocuments`	Número total de documentos en los que se pueden realizar búsquedas de todos los índices templados del clúster. Debe dejar el periodo en un minuto para obtener un valor preciso. Estadísticas pertinentes: suma
`WarmSearchLatency`	La diferencia en el tiempo total, en milisegundos, obtenida por todas las búsquedas en un Warm entre el minuto N y el minuto (N-1). Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo
`WarmSearchRate`	El número total de solicitudes de búsqueda por minuto para todos los fragmentos de un nodo Warm. Una sola llamada a la API `_search` podría devolver resultados de muchas particiones diferentes. Si cinco de estas particiones se encuentran en un nodo, este notificaría el valor 5 para la métrica, incluso aunque el cliente solo realizara una solicitud. Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: promedio, máximo, suma
`WarmStorageSpaceUtilization`	Cantidad total del espacio de almacenamiento templado, en MiB, que utiliza el clúster. Estadísticas pertinentes: máximo
`HotStorageSpaceUtilization`	Cantidad total del espacio de almacenamiento en caliente que utiliza el clúster. Estadísticas pertinentes: máximo
`WarmSysMemoryUtilization`	Porcentaje de la memoria del nodo maestro que está en uso. Estadísticas pertinentes: máximo

Métrica	Description (Descripción)
`HotToWarm Migration QueueSize`	Número de índices a la espera actualmente para migrar del almacenamiento caliente al almacenamiento templado. Estadísticas pertinentes: máximo
`WarmToHot Migration QueueSize`	Número de índices a la espera actualmente para migrar del almacenamiento templado al almacenamiento caliente. Estadísticas pertinentes: máximo
`HotToWarm Migration FailureCount`	El número total de errores de migración del almacenamiento en caliente al almacenamiento templado. Esta métrica solo está disponible para UltraWarm los nodos. Estadísticas pertinentes: suma
`HotToWarm Migration ForceMergeLatency`	Latencia media de la etapa de fusión de fuerzas del proceso de migración. Si esta etapa tarda demasiado de forma constante, considere la posibilidad de aumentarla. Esta métrica solo está disponible para UltraWarm los nodos. `index.ultrawarm.migration.force_merge.max_num_segments` . Estadísticas pertinentes: promedio
`HotToWarm Migration SnapshotLatency`	Latencia media de la etapa de instantánea del proceso de migración. Si esta etapa tarda demasiado tiempo, asegúrese de que las particiones estén correctamente dimensionadas y distribuidas por todo el clúster. Esta métrica solo está disponible para UltraWarm los nodos. Estadísticas pertinentes: promedio

Métrica	Description (Descripción)
`HotToWarmMigrationProcessingLatency`	La latencia media de migraciones exitosas del almacenamiento en caliente al almacenamiento templado, sin incluir el tiempo que estuvo en la cola. Este valor es la suma de la cantidad de tiempo que se tarda en completar las etapas de fusión de fuerza, instantánea y reubicación de particiones del proceso de migración. Esta métrica solo está disponible para UltraWarm los nodos. Estadísticas pertinentes: promedio
`HotToWarmMigrationSuccessCount`	El número total de migraciones exitosas del almacenamiento en caliente al almacenamiento templado. Estadísticas pertinentes: suma
`HotToWarmMigrationSuccessLatency`	La latencia media de migraciones exitosas del almacenamiento en caliente al almacenamiento templado, incluido el tiempo que estuvo en la cola. Estadísticas pertinentes: promedio
`WarmThreadpoolSearchThreads`	El tamaño del conjunto de hilos de búsqueda de Warm. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: promedio, suma
`WarmThreadpoolSearchRejected`	El número de tareas rechazadas en el conjunto de hilos de búsqueda de Warm. Si este número aumenta continuamente, considere la posibilidad de añadir más nodos de Warm. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma

Métrica	Description (Descripción)
`WarmThreadpoolSearchQueue`	El número de tareas en cola en el grupo de hilos de búsqueda de Warm. Si el tamaño de la cola es constantemente alto, considera añadir más nodos Warm. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
`WarmJVMMemoryPressure`	El porcentaje máximo del montón de Java utilizado para los nodos Warm. Estadísticas pertinentes: máximo **Note** La lógica de esta métrica cambió en el software del servicio R20220323. Para más información, consulte las notas de la versión.
`WarmOldGenJVMMemoryPressure`	El porcentaje máximo del montón de Java utilizado para la «generación anterior» por nodo Warm. Estadísticas pertinentes: máximo
`WarmJVMGCYoungCollectionCount`	El número de veces que la recolección de basura de la «generación joven» se ha ejecutado en nodos Warm. Un número grande y en continuo crecimiento de ejecuciones es normal para las operaciones de clúster. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio

Métrica	Description (Descripción)
`WarmJVMGCYoungCollectionTime`	La cantidad de tiempo, en milisegundos, que el clúster ha dedicado a la recolección de basura de la «generación joven» en los nodos Warm. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
`WarmJVMGCOldCollectionCount`	El número de veces que la recolección de basura de la «generación anterior» se ha ejecutado en nodos Warm. En un clúster con suficientes recursos, este número debe ser pequeño y crecer con poca frecuencia. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
`WarmConcurrentSearchRate`	El número total de solicitudes de búsqueda que utilizan la búsqueda de segmentos simultánea por minuto para todos los fragmentos de un nodo Warm. Una sola llamada a la API `_search` podría devolver resultados de muchas particiones diferentes. Si cinco de estas particiones se encuentran en un nodo, este notificaría el valor 5 para la métrica, incluso aunque el cliente solo realizara una solicitud. Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: suma, máximo, promedio
`WarmConcurrentSearchLatency`	La diferencia en el tiempo total, en milisegundos, obtenida por todas las búsquedas que utilizan la búsqueda de segmentos simultánea en un nodo cálido entre el minuto N y el minuto (N-1). Estadísticas de nodo pertinentes: promedio Estadísticas de clúster pertinentes: máximo, promedio

Métrica	Description (Descripción)
`WarmThreadpoolIndexSearcherQueue`	El número de tareas en cola del grupo de subprocesos del buscador de índices Warm. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, máximo, promedio
`WarmThreadpoolIndexSearcherRejected`	El número de tareas rechazadas en el grupo de subprocesos del buscador de índices Warm. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma
`WarmThreadpoolIndexSearcherThreads`	El tamaño del conjunto de subprocesos del buscador de índices Warm. Estadísticas de nodo pertinentes: máximo Estadísticas de clúster pertinentes: suma, promedio

Métricas de almacenamiento en frío

Amazon OpenSearch Service proporciona las siguientes métricas para el almacenamiento en frío.

Métrica	Description (Descripción)
`ColdStorageSpaceUtilization`	La cantidad total de espacio de almacenamiento en frío, en MiB, que el clúster utiliza. Estadísticas pertinentes: máximo
`ColdToWarmMigrationFailureCount`	El número total de errores de migración del almacenamiento en frío al almacenamiento templado. Estadísticas pertinentes: suma

Métrica	Description (Descripción)
`ColdToWarmMigrationLatency`	Cantidad de tiempo para que se completen de forma correcta las migraciones del almacenamiento en frío al almacenamiento templado. Estadísticas pertinentes: promedio
`ColdToWarmMigrationQueueSize`	Número de índices a la espera actualmente para migrar del almacenamiento frío al almacenamiento templado. Estadísticas pertinentes: máximo
`ColdToWarmMigrationSuccessCount`	El número total de migraciones exitosas del almacenamiento en frío al almacenamiento templado. Estadísticas pertinentes: suma
`WarmToColdMigrationFailureCount`	El número total de errores de migración del almacenamiento templado al almacenamiento en frío. Estadísticas pertinentes: suma
`WarmToColdMigrationLatency`	Cantidad de tiempo para que se completen de forma correcta las migraciones del almacenamiento templado al almacenamiento en frío. Estadísticas pertinentes: promedio
`WarmToColdMigrationQueueSize`	Número de índices a la espera actualmente para migrar del almacenamiento templado al almacenamiento frío. Estadísticas pertinentes: máximo
`WarmToColdMigrationSuccessCount`	El número total de migraciones exitosas del almacenamiento templado al almacenamiento en frío. Estadísticas pertinentes: suma

OpenSearch Métricas de instancias optimizadas (OR1)

Amazon OpenSearch Service proporciona las siguientes métricas para las OR1instancias.

Métrica	Description (Descripción)
`RemoteStorageUsedSpace`	Cantidad total del espacio de Amazon S3, en MiB, que utiliza el clúster. Estadísticas pertinentes: suma
`RemoteStorageWriteRejected`	El número total de solicitudes rechazadas en las particiones principales debido a la presión de replicación y almacenamiento remoto. Se calcula a partir del último inicio del proceso del OpenSearch servicio. Estadísticas pertinentes: suma
`ReplicationLagMaxTime`	El tiempo, en milisegundos, que las particiones de la réplica permanecen detrás de las particiones primarias. Estadísticas pertinentes: máximo

Métricas de alertas

Amazon OpenSearch Service proporciona las siguientes métricas para las alertas.

Métrica	Description (Descripción)
`AlertingDegraded`	Un valor de 1 significa que el índice de alerta es rojo o que uno o más nodos no van según lo programado. Un valor de 0 indica un comportamiento normal. Estadísticas pertinentes: máximo
`AlertingIndexExists`	Un valor de 1 significa que el índice `.opensearch-alerting-config` existe. Un valor de 0 significa que no existe. Hasta que utilice la característica de alerta por primera vez, este valor sigue siendo 0.

Métrica	Description (Descripción)
	Estadísticas pertinentes: máximo
`AlertingIndexStatus.green`	El estado de salud del índice. Un valor de 1 significa verde. Un valor de 0 significa que el índice no existe o no es verde. Estadísticas pertinentes: máximo
`AlertingIndexStatus.red`	El estado de salud del índice. Un valor de 1 significa rojo. Un valor de 0 significa que el índice no existe o no es rojo. Estadísticas pertinentes: máximo
`AlertingIndexStatus.yellow`	El estado de salud del índice. Un valor de 1 significa amarillo. Un valor de 0 significa que el índice no existe o no es amarillo. Estadísticas pertinentes: máximo
`AlertingNodesNotOnSchedule`	Un valor de 1 significa que algunos trabajos no se están ejecutando según lo programado. Un valor de 0 significa que todos los trabajos de alerta se están ejecutando según lo programado (o que no hay trabajos de alerta). Comprueba la consola OpenSearch de servicio o realiza una `_nodes/stats` solicitud para ver si algún nodo muestra un uso elevado de recursos. Estadísticas pertinentes: máximo
`AlertingNodesOnSchedule`	Un valor de 1 significa que todos los trabajos de alerta se están ejecutando según lo programado (o que no hay trabajos de alerta). Un valor de 0 significa que algunos trabajos no se están ejecutando según lo programado. Estadísticas pertinentes: máximo

Métrica	Description (Descripción)
`AlertingScheduledJobEnabled`	Un valor de 1 significa que la configuración del clúster `opensearch.scheduled_jobs.enabled` es verdadera. Un valor de 0 significa que es falsa y que los trabajos programados están deshabilitados. Estadísticas pertinentes: máximo

Métricas de detección de anomalías

Amazon OpenSearch Service proporciona las siguientes métricas para la detección de anomalías.

Métrica	Description (Descripción)
`ADPluginUnhealthy`	El valor 1 significa que el complemento de detección de anomalías no funciona correctamente, bien debido a un gran número de errores o bien porque uno de los índices que utiliza es rojo. Un valor de 0 indica que el complemento funciona como se esperaba. Estadísticas pertinentes: máximo
`ADExecuteRequestCount`	El número de solicitudes para detectar anomalías. Estadísticas pertinentes: suma
`ADExecuteFailureCount`	El número de solicitudes erróneas para detectar anomalías. Estadísticas pertinentes: suma
`ADHCExecuteFailureCount`	El número de errores de solicitud al detectar anomalías para los detectores de alta cardinalidad. Estadísticas pertinentes: suma
`ADHCExecuteRequestCount`	El número de solicitudes al detectar anomalías para detectores de alta cardinalidad. Estadísticas pertinentes: suma

Métrica	Description (Descripción)
ADAnomaly ResultsIn dexStatus IndexExists	Un valor de 1 significa que existe el índice al que apunta el alias .opensearch-anomaly-results . Hasta que no se utilice la detección de anomalías por primera vez, este valor permanece en 0. Estadísticas pertinentes: máximo
ADAnomaly ResultsIn dexStatus.red	Un valor de 1 significa que el índice al que apunta el alias .opensear ch-anomaly-results es rojo. Un valor de 0 significa que no lo es. Hasta que no se utilice la detección de anomalías por primera vez, este valor permanece en 0. Estadísticas pertinentes: máximo
ADAnomaly Detectors IndexStat usIndexExists	Un valor de 1 significa que el índice .opensearch-anomaly-detectors existe. Un valor de 0 significa que no existe. Hasta que no se utilice la detección de anomalías por primera vez, este valor permanece en 0. Estadísticas pertinentes: máximo
ADAnomaly Detectors IndexStat us.red	Un valor de 1 significa que el índice .opensearch-anomaly-detectors es rojo. Un valor de 0 significa que no lo es. Hasta que no se utilice la detección de anomalías por primera vez, este valor permanece en 0. Estadísticas pertinentes: máximo
ADModelsC heckpoint IndexStat usIndexExists	Un valor de 1 significa que el índice .opensearch-anomaly-checkpoints existe. Un valor de 0 significa que no existe. Hasta que no se utilice la detección de anomalías por primera vez, este valor permanece en 0. Estadísticas pertinentes: máximo

Métrica	Description (Descripción)
`ADModelsCheckpointIndexStatus.red`	Un valor de 1 significa que el índice `.opensearch-anomaly-checkpoints` es rojo. Un valor de 0 significa que no lo es. Hasta que no se utilice la detección de anomalías por primera vez, este valor permanece en 0. Estadísticas pertinentes: máximo

Métricas de búsqueda asíncrona

Amazon OpenSearch Service proporciona las siguientes métricas para la búsqueda asíncrona.

Estadísticas del nodo coordinador de búsqueda asíncrona (por nodo coordinador)

Métrica	Description (Descripción)
`AsynchronousSearchSubmissionRate`	El número de búsquedas asíncronas enviadas en el último minuto.
`AsynchronousSearchInitializedRate`	El número de búsquedas asíncronas inicializadas en el último minuto.
`AsynchronousSearchRunningCurrent`	El número de búsquedas asíncronas que se ejecutan actualmente.
`AsynchronousSearchCompletionRate`	El número de búsquedas asíncronas realizadas correctamente en el último minuto.
`AsynchronousSearchFailureRate`	El número de búsquedas asíncronas que se completaron y fallaron en el último minuto.

Métrica	Description (Descripción)
`AsynchronousSearchPersistRate`	El número de búsquedas asíncronas que persistieron en el último minuto.
`AsynchronousSearchPersistFailedRate`	El número de búsquedas asíncronas que no pudieron persistir en el último minuto.
`AsynchronousSearchRejected`	El número total de búsquedas asíncronas rechazadas desde el tiempo de actividad del nodo.
`AsynchronousSearchCancelled`	El número total de búsquedas asíncronas canceladas desde el tiempo de actividad del nodo.
`AsynchronousSearchMaxRunningTime`	Duración de la búsqueda asíncrona de ejecución más larga en un nodo en el último minuto.

Estadísticas del clúster de búsqueda asíncrona

Métrica	Description (Descripción)
`AsynchronousSearchStoreHealth`	El estado de la tienda en el índice persistente (ROJO/no ROJO) en el último minuto.
`AsynchronousSearchStoreSize`	El tamaño del índice del sistema en todas las particiones en el último minuto.
`AsynchronousSearch`	El número de respuestas almacenadas en el índice del sistema en el último minuto.

Métrica	Description (Descripción)
StoredResponseCount	

Métricas de ajuste automático

Amazon OpenSearch Service proporciona las siguientes métricas para Auto-Tune.

Métrica	Description (Descripción)
AutoTuneChangesHistoryHeapSize	El historial de cambios en MiB para los valores de ajuste del tamaño del montón.
AutoTuneChangesHistoryJVMYoungGenArgs	El historial de cambios de los argumentos de JVM YongGen .
AutoTuneFailed	Un valor booleano que indica si el cambio de ajuste automático produjo un error.
AutoTuneSucceeded	Un valor booleano que indica si el cambio de ajuste automático se realizó correctamente.
AutoTuneValue	El historial de cambios de cola (recuento) y el historial de cambios de ajustes de caché (en MiB) para los cambios no disruptivos.

Métricas de Multi-AZ con modo de espera

Amazon OpenSearch Service proporciona las siguientes métricas para Multi-AZ con modo de espera.

Métricas a nivel de nodo para los nodos de datos en zonas de disponibilidad activas

Métrica	Description (Descripción)
`CPUUtilization`	El porcentaje de uso de CPU para los nodos de datos del clúster. Máximo muestra el nodo con el mayor uso de CPU. Promedio represent a todos los nodos del clúster. Esta métrica también está disponible para los nodos individuales.
`FreeStoragеSpace`	El espacio libre para los nodos de datos del clúster. `Sum` muestra el espacio libre total del clúster, pero debe dejar el periodo en un minuto para recibir un valor preciso. `Minimum` y `Maximum` muestran los nodos con el mínimo y máximo espacio libre, respectivamente. Esta métrica también está disponible para nodos individuales. OpenSearch El servicio arroja un `ClusterBlockException` valor cuando esta métrica alcanza`0`. Para efectuar la recuperación, debe eliminar índices, agregar instancias de mayor tamaño o bien agregar almacenamiento basado en EBS a las instancias existentes. Para obtener más información, consulte the section called “Falta de espacio de almacenamiento disponible”. La consola OpenSearch de servicio muestra este valor en GiB. La CloudWatch consola de Amazon lo muestra en MiB.
`JVMMemory Pressure`	El porcentaje máximo del montón de Java utilizado para todos los nodos de datos del clúster. OpenSearch El servicio utiliza la mitad de la RAM de una instancia para el montón de Java, hasta un tamaño de pila de 32 GiB. Puede escalar las instancias verticalmente hasta 64 GiB de RAM y después escalarlas horizontalmente mediante el agregado de instancias. Consulte the section called “ CloudWatch Alarmas recomendadas”.
`SysMemory Utilization`	El porcentaje de memoria de la instancia que está en uso. Es normal obtener valores altos para esta métrica y, por lo general, no representan un problema con el clúster. Para obtener un mejor indicador de posibles problemas de rendimiento y estabilidad, consulte la métrica `JVMMemory Pressure` .
`IndexingLatency`	La diferencia en el tiempo total, en milisegundos, calculada por todas las operaciones de indexación en un nodo entre el minuto N y el minuto (N-1).

Métrica	Description (Descripción)
IndexingRate	El número de operaciones de indexación por minuto.
SearchLatency	La diferencia en el tiempo total, en milisegundos, calculada por todas las búsquedas en un nodo entre el minuto N y el minuto (N-1).
SearchRate	El número total de peticiones de búsqueda por minuto de todas las particiones de un nodo de datos.
ThreadpoolSearchQueue	El número de tareas en cola en el grupo de subprocesos de búsqueda. Si el tamaño de la cola es sistemáticamente elevado, considere la posibilidad de escalar el clúster. El tamaño máximo de la cola de búsqueda es 1000.
ThreadpoolWriteQueue	El número de tareas en cola en el grupo de subprocesos de escritura.
ThreadpoolSearchRejected	El número de tareas rechazadas en el grupo de subprocesos de búsqueda. Si este número crece continuamente, considere la posibilidad de escalar el clúster.
ThreadpoolWriteRejected	El número de tareas rechazadas en el grupo de subprocesos de escritura.

Métricas de nivel de clúster para clústeres en zonas de disponibilidad activas

Métrica	Description (Descripción)
DataNodes	El número total de particiones activas y en espera.
DataNodesShards.active	El número total de particiones primarias y de réplicas activas.
DataNodesShards.unassigned	El número de particiones que no se asignaron a los nodos del clúster.

Métrica	Description (Descripción)
`DataNodes Shards.in itializing`	El número de particiones que se encuentran en inicialización.
`DataNodes Shards.re locating`	El número de particiones que se encuentran en reubicación.

Métricas de rotación de zonas de disponibilidad

Si `ActiveReads.`*`Availability-Zone`* `= 1`, entonces la zona está activa. Si `ActiveReads.`*`Availability-Zone`* `= 0`, entonces la zona está en espera.

Métricas de un momento dado

Amazon OpenSearch Service proporciona las siguientes métricas para las búsquedas puntuales (PIT).

Estadísticas del nodo coordinador de PIT (por nodo coordinador)

Métrica	Description (Descripción)
`CurrentPo intInTime`	El número de contextos de búsqueda PIT activos en el nodo.
`TotalPoin tInTime`	El número de contextos de búsqueda PIT caducados desde el tiempo de actividad del nodo.
`AvgPointI nTimeAliveTime`	El keep-alive promedio de los contextos de búsqueda PIT desde el tiempo de actividad del nodo.
`HasActive PointInTime`	Un valor de 1 indica que hay contextos PIT activos en los nodos desde el tiempo de actividad del nodo. Un valor de 0 indica que no hay.
`HasUsedPo intInTime`	Un valor de 1 indica que hay contextos PIT vencidos en los nodos desde el tiempo de actividad del nodo. Un valor de 0 indica que no hay.

Métricas de SQL

Amazon OpenSearch Service proporciona las siguientes métricas para el soporte de SQL.

Métrica	Description (Descripción)
`SQLFailedRequestCountByCusErr`	El número de solicitudes a la API `_sql` que no se han realizado correctamente debido a un problema del cliente. Por ejemplo, una solicitud podría devolver el código de estado HTTP 400 debido a una excepción `IndexNotFoundException` . Estadísticas pertinentes: suma
`SQLFailedRequestCountBySysErr`	El número de solicitudes a la API `_sql` que no se han realizado correctamente debido a un problema del servidor o a una limitación de características. Por ejemplo, una solicitud podría devolver el código de estado HTTP 503 debido a una excepción `VerificationException` . Estadísticas pertinentes: suma
`SQLRequestCount`	El número de solicitudes a la API `_sql`. Estadísticas pertinentes: suma
`SQLDefaultCursorRequestCount`	Similar a `SQLRequestCount` , pero solo cuenta las solicitudes de paginación. Estadísticas pertinentes: suma
`SQLUnhealthy`	Un valor de 1 indica que, en respuesta a determinadas solicitudes, el complemento SQL devuelve códigos de respuesta 5xx o pasa un DSL de consulta no válida a OpenSearch. Las demás solicitudes deberían realizarse correctamente. Un valor de 0 indica que no hay errores recientes. Si aparece sistemáticamente un valor de 1, solucione los problemas en las solicitudes que sus clientes están realizando al complemento. Estadísticas pertinentes: máximo

Métricas k-NN

Amazon OpenSearch Service incluye las siguientes métricas para el complemento k-Nearest neighbor (k-NN).

Métrica	Description (Descripción)
`KNNCacheCapacityReached`	Métrica por nodo para saber si se alcanzó la capacidad de caché. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: máximo
`KNNCircuitBreakerTriggered`	Métrica por clúster para saber si se activa el disyuntor. Si algún nodo devuelve un valor de 1 para `KNNCacheCapacityReached` , este valor también devolverá 1. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: máximo
`KNNEvictionCount`	Métrica por nodo para el número de gráficos que se expulsaron de la caché debido a limitaciones de memoria o tiempo de inactividad. Las expulsiones explícitas que se producen debido a la eliminación de indexación no se cuentan. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: suma
`KNNGraphIndexErrors`	Métrica por nodo para el número de solicitudes para agregar el campo `knn_vector` de un documento a un gráfico que produjo un error. Estadísticas pertinentes: suma
`KNNGraphIndexRequests`	Métrica por nodo para el número de solicitudes para agregar el campo `knn_vector` de un documento a un gráfico. Estadísticas pertinentes: suma

Métrica	Description (Descripción)
KNNGraphMemoryUsage	Métrica por nodo para el tamaño actual de caché (tamaño total de todos los gráficos en memoria) en kilobytes. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: promedio
KNNGraphQueryErrors	Métrica por nodo para el número de consultas de gráficos que produjeron un error. Estadísticas pertinentes: suma
KNNGraphQueryRequests	Métrica por nodo para el número de consultas de gráficos. Estadísticas pertinentes: suma
KNNHitCount	Métrica por nodo para el número de aciertos de caché. Un acierto de caché se produce cuando un usuario consulta un gráfico que ya está cargado en la memoria. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: suma
KNNLoadExceptionCount	Métrica por nodo para el número de veces que se produjo una excepción al intentar cargar un gráfico en la caché. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: suma
KNNLoadSuccessCount	Métrica por nodo para el número de veces que el complemento cargó correctamente un gráfico en la caché. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: suma

Métrica	Description (Descripción)
KNNMissCount	Métrica por nodo para el número de errores de caché. Un error de caché se produce cuando un usuario consulta un gráfico que aún no está cargado en la memoria. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: suma
KNNQueryRequests	Métrica por nodo para el número de solicitudes de consulta que recibió el complemento k-NN. Estadísticas pertinentes: suma
KNNRemoteBuildEnabled	Valor binario que especifica si la función está habilitada. Estadísticas relevantes: binarias
KNNRemoteIndexBuildFailureCount	Número total de errores de construcción. Estadísticas pertinentes: suma
KNNRemoteIndexBuildSuccessCount	Número total de compilaciones correctas. Estadísticas pertinentes: suma
KNNScriptCompilationErrors	Métrica por nodo para el número de errores durante la compilación del script. Esta estadística solo es relevante para la búsqueda de scripts de puntuación k-NN. Estadísticas pertinentes: suma
KNNScriptCompilations	Métrica por nodo para el número de veces que se ha compilado el script k-NN. Por lo general, este valor debe ser 1 o 0, pero si se llena la caché que contiene los scripts compilados, el script k-NN podría ser recompilado. Esta estadística solo es relevante para la búsqueda de scripts de puntuación k-NN. Estadísticas pertinentes: suma

Métrica	Description (Descripción)
KNNScriptQueryErrors	Métrica por nodo para el número de errores durante las consultas de scripts. Esta estadística solo es relevante para la búsqueda de scripts de puntuación k-NN. Estadísticas pertinentes: suma
KNNScriptQueryRequests	Métrica por nodo para el número total de consultas de script. Esta estadística solo es relevante para la búsqueda de scripts de puntuación k-NN. Estadísticas pertinentes: suma
KNNTotalLoadTime	El tiempo en nanosegundos que k-NN tardó en cargar gráficos en la caché. Esta métrica solo es pertinente para la búsqueda aproximada de k-NN. Estadísticas pertinentes: suma
VectorIndexBuildAccelerationOCU	El número de unidades de OpenSearch cálculo (OCUs) utilizadas para acelerar la indexación vectorial. Estadísticas pertinentes: suma

Métricas de búsqueda entre clústeres

Amazon OpenSearch Service proporciona las siguientes métricas para la búsqueda entre clústeres.

Métricas de dominio de origen

Métrica	Dimensión	Description (Descripción)
CrossClusterOutboundConnections	ConnectionId	Número de nodos conectados. Si la respuesta incluye uno o más dominios omitidos, utilice esta métrica para rastrear las conexiones que no estén en buen estado. Si este número cae a 0, entonces la conexión no está en buen estado.

Métrica	Dimensión	Description (Descripción)
`CrossClusterOutboundRequests`	`ConnectionId`	El número de peticiones de búsqueda enviadas al dominio de destino. Utilícelo para comprobar si la carga de peticiones de búsqueda en clústeres está sobrecargando su dominio, correlacione cualquier pico en esta métrica con cualquier pico JVM/CPU.

Métrica de dominio de destino

Métrica	Dimensión	Description (Descripción)
`CrossClusterInboundRequests`	`ConnectionId`	El número de solicitudes de conexión entrantes recibidas desde el dominio de origen.

Añade una CloudWatch alarma en caso de que pierdas una conexión de forma inesperada. Para ver los pasos para crear una alarma, consulte Crear una CloudWatch alarma basada en un umbral estático.

Métricas de replicación entre clústeres

Amazon OpenSearch Service proporciona las siguientes métricas para la replicación entre clústeres.

Métrica	Description (Descripción)
`ReplicationRate`	La tasa media de operaciones de replicación por segundo. Esta métrica es similar a `IndexingRate` .
`LeaderCheckPoint`	Para lograr una conexión específica, la suma de los valores de punto de control líder entre todos los índices de replicación. Puede utilizar esta métrica para medir la latencia de replicación.
`FollowerCheckPoint`	Para lograr una conexión específica, la suma de los valores de punto de control seguidor entre todos los índices de replicación. Puede utilizar esta métrica para medir la latencia de replicación.

Métrica	Description (Descripción)
`ReplicationNumSyncingIndices`	El número de índices que tienen un estado de replicación de `SYNCING`.
`ReplicationNumBootstrappingIndices`	El número de índices que tienen un estado de replicación de `BOOTSTRAPPING` .
`ReplicationNumPausedIndices`	El número de índices que tienen un estado de replicación de `PAUSED`.
`ReplicationNumFailedIndices`	El número de índices que tienen un estado de replicación de `FAILED`.
`CrossClusterOutboundReplicationRequests`	El número de solicitudes de transporte de replicación en el dominio seguidor. Las solicitudes de transporte son internas y se producen cada vez que se llama una operación de la API de replicación. También se producen cuando el dominio seguidor sondea los cambios con respecto al dominio líder.
`CrossClusterInboundReplicationRequests`	El número de solicitudes de transporte de replicación en el dominio principal. Las solicitudes de transporte son internas y se producen cada vez que se llama una operación de API de replicación.
`AutoFollowNumSuccessStartReplication`	El número de índices de seguidores que ha creado correctamente una regla de replicación para una conexión específica.

Métrica	Description (Descripción)
`AutoFollowNumFailedStartReplication`	El número de índices de seguidores que no pudo crear una regla de replicación cuando había un patrón coincidente. Este problema puede deberse a un problema de red en el clúster remoto o a un problema de seguridad (es decir, el rol asociado no tiene permiso para iniciar la replicación).
`AutoFollowLeaderCallFailure`	Si ha habido alguna consulta fallida del índice de seguidores al índice de líderes para obtener nuevos datos. Un valor de 1 significa que ha habido 1 o más llamadas fallidas en el último minuto.

Aprender a clasificar métricas

Amazon OpenSearch Service proporciona las siguientes métricas para aprender a clasificar.

Métrica	Description (Descripción)
`LTRRequestTotalCount`	Recuento total de solicitudes de clasificación.
`LTRRequestErrorCount`	Recuento total de solicitudes fallidas.
`LTRStatus.red`	Comprueba si uno de los índices necesarios para ejecutar el complemento es rojo.
`LTRMemoryUsage`	Memoria total utilizada por el complemento.
`LTRFeatureMemoryUsageInBytes`	La cantidad de memoria, en bytes, utilizada por los campos de características de Learning to Rank (Aprender a clasificar).
`LTRFeatursetMemoryUsageInBytes`	La cantidad de memoria, en bytes, que utiliza todos los conjuntos de características de Learning to Rank (Aprender a clasificar).

Métrica	Description (Descripción)
`LTRModelMemoryUsageInBytes`	La cantidad de memoria, en bytes, utilizada por todos los modelos Learning to Rank (Aprender a clasificar).

Métricas del lenguaje de procesamiento de canalizaciones

Amazon OpenSearch Service proporciona las siguientes métricas para Piped Processing Language.

Métrica	Description (Descripción)
`PPLFailedRequestCountByCusErr`	El número de solicitudes a la API `_ppl` que no se han realizado correctamente debido a un problema del cliente. Por ejemplo, una solicitud podría devolver el código de estado HTTP 400 debido a una excepción `IndexNotFoundException` .
`PPLFailedRequestCountBySysErr`	El número de solicitudes a la API `_ppl` que no se han realizado correctamente debido a un problema del servidor o a una limitación de características. Por ejemplo, una solicitud podría devolver el código de estado HTTP 503 debido a una excepción `VerificationException` .
`PPLRequestCount`	El número de solicitudes a la API `_ppl`.

Supervisión de OpenSearch registros con Amazon CloudWatch Logs

Amazon OpenSearch Service expone los siguientes OpenSearch registros a través de Amazon CloudWatch Logs:

- Registros de errores
- Registros lentos de solicitudes de búsqueda
- Registros lentos de particiones
- Registros de auditoría

Los registros lentos de particiones de búsqueda, los registros lentos de particiones de índice y los registros de errores son útiles para la resolución de problemas de rendimiento y estabilidad. Los registros de auditoría realizan un seguimiento de la actividad de los usuarios con fines de conformidad. Todos los registros están deshabilitados de forma predeterminada. Si está activado, se aplicará el CloudWatch precio estándar.

> **Note**
>
> Los registros de errores solo están disponibles para las versiones 5.1 OpenSearch y posteriores de Elasticsearch. Los registros lentos están disponibles para todas las versiones OpenSearch y para las versiones de Elasticsearch.

Para sus registros, OpenSearch utiliza Apache Log4j 2 y sus niveles de registro integrados (del menor al más grave) de`TRACE`,, `DEBUG``INFO`, `WARN` y. `ERROR` `FATAL`

Si habilita los registros de errores, el OpenSearch servicio publica las líneas de registro de `WARN``ERROR`, y `FATAL` para. CloudWatch OpenSearch El servicio también publica varias excepciones a este `DEBUG` nivel, entre las que se incluyen las siguientes:

- `org.opensearch.index.mapper.MapperParsingException`
- `org.opensearch.index.query.QueryShardException`
- `org.opensearch.action.search.SearchPhaseExecutionException`
- `org.opensearch.common.util.concurrent.OpenSearchRejectedExecutionException`
- `java.lang.IllegalArgumentException`

Los registros de errores ayudan a solucionar problemas en muchas situaciones, entre las que se incluyen las siguientes:

- Problemas de compilación de script Painless
- Consultas no válidas
- Problemas de indexación
- Errores de instantáneas
- Fallas de migración de la administración de estados de índice

> **Note**
>
> No todos los errores se registran en los registros de errores.

> **Note**
>
> OpenSearch El servicio no registra todos los errores que se producen.

Temas

- Habilitación de la publicación de registros (consola)
- Habilitación de la publicación de registros (AWS CLI)
- Habilitación de la publicación de registros (AWS SDKs)
- Habilitación de la publicación de registros (CloudFormation)
- Establecimiento de umbrales de registros lentos de solicitudes de búsqueda
- Establecimiento de umbrales de registros lentos de particiones
- Prueba de registros lentos
- Visualización de registros

Habilitación de la publicación de registros (consola)

La consola OpenSearch de servicio es la forma más sencilla de permitir la publicación de registros en CloudWatch.

Para habilitar la publicación de registros en CloudWatch (consola)

1. Vaya a aws.amazon.com, luego elija Iniciar sesión y proporcione sus credenciales.
2. En Analytics, selecciona Amazon OpenSearch Service.
3. Seleccione el dominio que desea actualizar.
4. En la pestaña Registros, seleccione un tipo de registro y luego elija Habilitar.
5. Cree un nuevo grupo de CloudWatch registros o elija uno existente.

> **Note**
>
> Si prevé habilitar varios registros, recomendamos publicar cada uno en su propio grupo de registros. Esta separación hace que los registros sean más fáciles de analizar.

6. Seleccione una política de acceso que contenga los permisos adecuados o cree una política mediante el JSON que proporciona la consola:

   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:*"
    }
  ]
}
```

   Le recomendamos que agregue las claves de condición `aws:SourceAccount` y `aws:SourceArn` para protegerse contra la política problema del suplente confuso. La cuenta de origen es la propietaria del dominio y el ARN de origen es el ARN del dominio. Su dominio debe estar en el software de servicio R20211203 o posterior para agregar estas claves de condición.

   Por ejemplo, podría agregar el siguiente bloque de condición a la política:

```
"Condition": {
    "StringEquals": {
        "aws:SourceAccount": "account-id"
    },
```

```
    "ArnLike": {
        "aws:SourceArn": "arn:aws:es:region:account-id:domain/domain-name"
    }
}
```

> **Important**
>
> CloudWatch Logs admite 10 políticas de recursos por región. Si planea habilitar los registros para varios dominios de OpenSearch servicio, debe crear y reutilizar una política más amplia que incluya varios grupos de registros para evitar alcanzar este límite. Para obtener información sobre los pasos para actualizar la política, consulte the section called “Habilitación de la publicación de registros (AWS CLI)”.

7. Seleccione Habilitar.

   El estado del dominio cambia de Activo a En proceso. El estado debe volver a Activo antes de que la publicación de registro esté habilitada. Este cambio suele tardar 30 minutos, pero puede tardar más en función de la configuración del dominio.

Si ha habilitado uno de los registros lentos de particiones, consulte the section called “Establecimiento de umbrales de registros lentos de particiones”. Si habilitó registros de auditorías, consulte the section called “Paso 2: Active los registros de auditoría en los OpenSearch paneles”. Si habilitó únicamente los registros de errores, no tiene que realizar ningún paso de configuración adicional.

Habilitación de la publicación de registros (AWS CLI)

Para poder habilitar la publicación de registros, necesita un grupo de CloudWatch registros. Si aún no lo tiene, puede crear uno mediante el siguiente comando:

```
aws logs create-log-group --log-group-name my-log-group
```

Introduzca el siguiente comando para encontrar el ARN del grupo de registros y, a continuación, anótelo:

```
aws logs describe-log-groups --log-group-name my-log-group
```

Ahora puede conceder permisos OpenSearch al Servicio para que escriba en el grupo de registros. Debe proporcionar el ARN del grupo de registros cerca del final 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 admite 10 políticas de recursos por región. Si planea habilitar los registros lentos por fragmentos para varios dominios de OpenSearch servicio, debe crear y reutilizar una política más amplia que incluya varios grupos de registros para evitar que se alcance este límite.

Si tiene que revisar esta política más adelante, utilice el comando `aws logs describe-resource-policies`. Para actualizar la política, ejecute el mismo comando `aws logs put-resource-policy` con un nuevo documento de política.

Finalmente, puede utilizar la opción `--log-publishing-options` para habilitar la publicación. La sintaxis de la opción es la misma para los comandos `create-domain` y `update-domain-config`.

Parámetro	Valores válidos
`--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**
>
> Si prevé habilitar varios registros, recomendamos publicar cada uno en su propio grupo de registros. Esta separación hace que los registros sean más fáciles de analizar.

Ejemplo

En el siguiente ejemplo se habilita la publicación de los registros lentos de particiones de búsqueda e índice para el dominio especificado:

```
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}"
```

Para deshabilitar la publicación en CloudWatch, ejecute el mismo comando con`Enabled=false`.

Si ha habilitado uno de los registros lentos de particiones, consulte the section called "Establecimiento de umbrales de registros lentos de particiones". Si habilitó registros de auditorías, consulte the section called "Paso 2: Active los registros de auditoría en los OpenSearch paneles". Si habilitó únicamente los registros de errores, no tiene que realizar ningún paso de configuración adicional.

Habilitación de la publicación de registros (AWS SDKs)

Para poder habilitar la publicación de registros, primero debe crear un grupo de CloudWatch registros, obtener su ARN y conceder permisos de OpenSearch servicio para escribir en él. Las operaciones relevantes se documentan en la referencia de la API de Amazon CloudWatch Logs:

- `CreateLogGroup`
- `DescribeLogGroup`
- `PutResourcePolicy`

Puede acceder a estas operaciones mediante AWS SDKs.

Los AWS SDKs (excepto Android e iOS SDKs) admiten todas las operaciones que se definen en la referencia de la API de Amazon OpenSearch Service, incluida la `--log-publishing-options` opción para `CreateDomain` y`UpdateDomainConfig`.

Si ha habilitado uno de los registros lentos de particiones, consulte the section called "Establecimiento de umbrales de registros lentos de particiones". Si habilitó únicamente los registros de errores, no tiene que realizar ningún paso de configuración adicional.

Habilitación de la publicación de registros (CloudFormation)

En este ejemplo, solemos crear un grupo de registros llamado`opensearch-logs`, asignar los permisos adecuados y, CloudFormation a continuación, crear un dominio con la publicación de registros habilitada para los registros de las aplicaciones, los registros lentos de los fragmentos de búsqueda y la indexación de los registros lentos.

Para poder habilitar la publicación de registros, debe crear un grupo de CloudWatch registros:

```
Resources:
  OpenSearchLogGroup:
    Type: AWS::Logs::LogGroup
    Properties:
      LogGroupName: opensearch-logs
Outputs:
  Arn:
    Value:
      'Fn::GetAtt':
        - OpenSearchLogGroup
        - Arn
```

La plantilla genera el ARN del grupo de registros. En este caso, el ARN es `arn:aws:logs:us-east-1:123456789012:log-group:opensearch-logs`.

Con el ARN, cree una política de recursos que otorgue al OpenSearch Servicio permisos para escribir en el grupo de registros:

```
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:*\"}]}"
```

Por último, cree la siguiente CloudFormation pila, que genera un dominio de OpenSearch servicio con la publicación de registros. La política de acceso permite Cuenta de AWS al usuario realizar todas las solicitudes 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
```

Para obtener información detallada sobre la sintaxis, consulte las opciones de publicación de registros en la Guía del usuario de CloudFormation .

Establecimiento de umbrales de registros lentos de solicitudes de búsqueda

Los registros lentos de las solicitudes de búsqueda están disponibles para buscar en los dominios OpenSearch de servicio que se ejecutan en la versión 2.13 y versiones posteriores. Los umbrales de registros lentos de solicitudes de búsqueda se configuran para el tiempo total de procesamiento de las solicitudes. Esto es diferente de los registros lentos de solicitudes de particiones, que se configuran para el tiempo individual de procesamiento de las particiones.

Puede especificar los registros lentos de solicitudes de búsqueda con la configuración del clúster. Esto difiere de los registros lentos de particiones, que se habilitan con la configuración del índice. Por ejemplo, puedes especificar los siguientes ajustes a través de la API OpenSearch REST:

```
PUT domain-endpoint/_cluster/settings
{
  "transient": {
    "cluster.search.request.slowlog.threshold.warn": "5s",
    "cluster.search.request.slowlog.threshold.info": "2s"
  }
}
```

> **Note**
>
> El `PUT /_cluster/settings` punto final está restringido en Amazon OpenSearch Service. Al enviar esta solicitud, se mostrará el siguiente error: `Your request: '/_cluster/settings' payload is not allowed.` para configurar los umbrales de registro lento de los fragmentos, utilice en su lugar ajustes a nivel de índice. Por ejemplo:
>
> ```
> PUT domain-endpoint/index/_settings
> {
>   "index.search.slowlog.threshold.query.warn": "5s",
>   "index.search.slowlog.threshold.query.info": "2s"
> }
> ```

Establecimiento de umbrales de registros lentos de particiones

OpenSearch deshabilita los registros lentos de los fragmentos de forma predeterminada. Después de habilitar la publicación de registros lentos en fragmentos CloudWatch, aún debe especificar los umbrales de registro para cada índice. OpenSearch Estos umbrales definen exactamente lo que se debe registrar y el nivel de registro.

Por ejemplo, puede especificar estos ajustes a través de la API OpenSearch REST:

```
PUT domain-endpoint/index/_settings
{
  "index.search.slowlog.threshold.query.warn": "5s",
  "index.search.slowlog.threshold.query.info": "2s"
}
```

Prueba de registros lentos

Para comprobar que tanto los registros lentos de las solicitudes de búsqueda como los de fragmentos lentos se publican correctamente, considere la posibilidad de empezar con valores muy bajos para comprobar que los registros aparecen en ellos y CloudWatch, a continuación, aumentar los umbrales hasta alcanzar niveles más útiles.

Si los registros no aparecen, verifique lo siguiente:

- ¿Existe el grupo de CloudWatch registros? Compruebe la CloudWatch consola.
- ¿El OpenSearch servicio tiene permisos para escribir en el grupo de registros? Compruebe la consola OpenSearch de servicio.
- ¿El dominio del OpenSearch servicio está configurado para publicarse en el grupo de registros? Compruebe la consola de OpenSearch servicio, utilice la AWS CLI `describe-domain-config` opción o llame `DescribeDomainConfig` mediante una de las SDKs.
- ¿Los umbrales de OpenSearch registro son lo suficientemente bajos como para que sus solicitudes los superen?

  Para revisar los umbrales de registros lentos de solicitudes de búsqueda para un dominio, use el siguiente comando:

  ```
  GET domain-endpoint/_cluster/settings?flat_settings
  ```

  Para revisar los umbrales de registros lentos de particiones, use el siguiente comando:

```
GET domain-endpoint/index/_settings?pretty
```

Si desea deshabilitar los registros lentos de un índice, devuelva los umbrales que ha cambiado a sus valores predeterminados de `-1`.

Al deshabilitar la publicación para CloudWatch usar la consola de OpenSearch servicio, AWS CLI no se detiene la generación OpenSearch de registros; solo se detiene la publicación de esos registros. Asegúrese de verificar la configuración de su índice si ya no necesita registros lentos de particiones, y la configuración de su dominio si ya no necesita registros lentos de solicitudes de búsqueda.

Visualización de registros

Ver la aplicación e iniciar sesión lentamente CloudWatch es como ver cualquier otro CloudWatch registro. Para obtener más información, consulta Ver datos de registro en la Guía del usuario de Amazon CloudWatch Logs.

Aquí tiene algunas consideraciones para visualizar los registros:

- OpenSearch El servicio publica solo los primeros 255 000 caracteres de cada línea para CloudWatch. El contenido restante está truncado. Para los registros de auditoría, son 10 000 caracteres por mensaje.
- En CloudWatch, los nombres de los flujos de registro tienen los sufijos`-index-slow-logs`, `-search-slow-logs-application-logs`, y `-audit-logs` para ayudar a identificar su contenido.

Supervisión de los registros de auditoría en Amazon OpenSearch Service

Si tu dominio de Amazon OpenSearch Service utiliza un control de acceso detallado, puedes habilitar los registros de auditoría para tus datos. Los registros de auditoría son altamente personalizables y te permiten realizar un seguimiento de la actividad de los usuarios en tus OpenSearch clústeres, incluidas las autenticaciones correctas y fallidas, las solicitudes, los cambios de indexación y las consultas de búsqueda entrantes. OpenSearch La configuración predeterminada realiza un seguimiento de un conjunto común de acciones de usuario, pero recomendamos adaptar la configuración a las necesidades exactas.

Al igual que los registros de OpenSearch aplicaciones y los registros lentos, OpenSearch Service publica los registros de auditoría en CloudWatch Logs. Si está activado, se aplica el CloudWatch precio estándar.

> **Note**
>
> Para habilitar los registros de auditoría, su rol de usuario debe estar asignado al `security_manager` rol, lo que le da acceso a la API OpenSearch `plugins/_security` REST. Para obtener más información, consulte the section called “Modificar el usuario maestro”.

Temas

- Limitaciones
- Habilitación de los registros de auditoría
- Habilite el registro de auditorías mediante el AWS CLI
- Habilite el registro de auditoría con la API de configuración
- Categorías y capas de registro de auditoría
- Configuración de registros de auditoría
- Ejemplo de registro de auditoría
- Configuración de registros de auditoría mediante la API REST

Limitaciones

Los registros de auditoría presentan las siguientes limitaciones:

- Los registros de auditoría no incluyen peticiones de búsqueda entre clústeres rechazadas por la política de acceso al dominio de destino.
- El tamaño máximo de cada mensaje de registro de auditoría es de 10 000 caracteres. El mensaje del registro de auditoría se trunca si se supera este límite.

Habilitación de los registros de auditoría

La habilitación de registros de auditoría es un proceso de dos pasos. En primer lugar, debe configurar su dominio para publicar los registros de auditoría en CloudWatch Logs. A continuación,

habilita los registros de auditoría en los OpenSearch paneles de control y los configura para que se adapten a sus necesidades.

> **Important**
>
> Si encuentra un error al seguir estos pasos, consulte the section called “No se pueden habilitar los registros de auditoría” para obtener información para la resolución de problemas.

Paso 1: Habilitar los registros de auditoría y configurar una política de acceso

En estos pasos se describe cómo habilitar los registros de auditoría con la consola. También puede habilitarlos mediante la AWS CLI API de servicio o la API de OpenSearch servicio.

Para habilitar los registros de auditoría para un dominio de OpenSearch servicio (consola)

1. Seleccione el dominio para abrir su configuración y, a continuación, vaya a la pestaña Registros.
2. Seleccione Registros de auditoría y, luego Habilitar.
3. Cree un grupo de CloudWatch registros o elija uno existente.
4. Seleccione una política de acceso que contenga los permisos adecuados o cree una política mediante el JSON que proporciona la consola:

   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/*"
    }
```

```
  ]
}
```

Le recomendamos que agregue las claves de condición `aws:SourceAccount` y `aws:SourceArn` para protegerse contra la política problema del suplente confuso. La cuenta de origen es la propietaria del dominio y el ARN de origen es el ARN del dominio. Su dominio debe estar en el software de servicio R20211203 o posterior para agregar estas claves de condición.

Por ejemplo, podría agregar el siguiente bloque de condición a la política:

```
"Condition": {
    "StringEquals": {
        "aws:SourceAccount": "account-id"
    },
    "ArnLike": {
        "aws:SourceArn": "arn:aws:es:region:account-id:domain/domain-name"
    }
}
```

5. Seleccione Habilitar.

Paso 2: Active los registros de auditoría en los OpenSearch paneles

Después de habilitar los registros de auditoría en la consola de OpenSearch servicio, también debe habilitarlos en los OpenSearch paneles y configurarlos para que se adapten a sus necesidades.

1. Abre los OpenSearch paneles y selecciona Seguridad en el menú de la izquierda.
2. Seleccione Registros de auditoría.
3. Seleccione Habilitar registro de auditoría.

La interfaz de usuario de Dashboards ofrece un control total de la configuración del registro de auditoría en Configuración general y en Configuración de la conformidad. Para obtener una descripción de todas las opciones de configuración, consulte Configuración del registro de auditoría.

Habilite el registro de auditorías mediante el AWS CLI

El siguiente AWS CLI comando habilita los registros de auditoría en un dominio existente:

```
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}"
```

También puede habilitar registros de auditoría al crear un dominio. Para obtener más información, consulte la Referencia de comandos de la AWS CLI.

Habilite el registro de auditoría con la API de configuración

La siguiente solicitud a la API de configuración habilita los registros de auditoría en un dominio existente:

```
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
    }
  }
}
```

Para obtener más información, consulta la referencia de la API OpenSearch de Amazon Service.

Categorías y capas de registro de auditoría

La comunicación del clúster se produce a través de dos capas independientes: la capa REST y la capa de transporte.

- La capa REST cubre la comunicación con clientes HTTP como curl, Logstash, OpenSearch Dashboards, el cliente REST de alto nivel de Java y la biblioteca de solicitudes de Python, todas las solicitudes HTTP que llegan al clúster.
- La capa de transporte se encarga de la comunicación entre nodos. Por ejemplo, después de que una petición de búsqueda llega al clúster (en la capa REST), el nodo de coordinación que se encarga de la solicitud envía la consulta a otros nodos, recibe sus respuestas, reúne los documentos necesarios y los reúne en la respuesta final. Las operaciones como la asignación de particiones y el reequilibrio también se producen en la capa de transporte.

Puede habilitar o deshabilitar los registros de auditoría para capas enteras, así como las categorías de auditoría individuales de una capa. La siguiente tabla contiene un resumen de las categorías de auditoría y las capas para las que están disponibles.

Categoría	Description (Descripción)	Disponible para REST	Disponible para transporte
FAILED_LOGIN	Una solicitud contenía credenciales no válidas y la autenticación falló.	Sí	Sí
MISSING_PRIVILEGES	Un usuario no poseía los privilegios para realizar la solicitud.	Sí	Sí
GRANTED_PRIVILEGES	Un usuario poseía los privilegios para realizar la solicitud.	Sí	Sí
OPENSEARCH_SECURITY_INDEX_ATTEMPT	Una solicitud intentó modificar el índice `.opendistro_security` .	No	Sí
AUTHENTICATED	Una solicitud contenía credenciales válidas y la autenticación se realizó correctamente.	Sí	Sí
INDEX_EVENT	Una solicitud realizó una operación administrativa con un índice, como crearlo, establecer un alias o realizar una combinación de fuerza. La lista completa de `indices:a`	No	Sí

Categoría	Description (Descripción)	Disponible para REST	Disponible para transporte
	dmin/ acciones que incluye esta categoría está disponible en la documentación. OpenSearch		

Además de estas categorías estándar, el control de acceso detallado ofrece varias categorías adicionales diseñadas para cumplir con los requisitos de conformidad de los datos.

Categoría	Description (Descripción)
COMPLIANCE_DOC_READ	Una solicitud realizó un evento de lectura en un documento en un índice.
COMPLIANCE_DOC_WRITE	Una solicitud realizó un evento de escritura en un documento en un índice.
COMPLIANCE_INTERNAL_CONFIG_READ	Una solicitud realizó un evento de lectura en el índice `.opendistro_security` .
COMPLIANCE_INTERNAL_CONFIG_WRITE	Una solicitud realizó un evento de escritura en el índice `.opendistro_security` .

Puede utilizar cualquier combinación de categorías y atributos de mensaje. Por ejemplo, si envía una solicitud REST para indexar un documento, podría ver las siguientes líneas en los registros de auditoría:

- AUTHENTICATED en la capa REST (autenticación)
- GRANTED_PRIVILEGE en la capa de transporte (autorización)
- COMPLIANCE_DOC_WRITE (documento escrito en un índice)

Configuración de registros de auditoría

Los registros de auditoría cuentan con numerosas opciones de configuración.

Configuración general

La configuración general permite habilitar o deshabilitar categorías individuales o capas enteras. Recomendamos encarecidamente mantener GRANTED_PRIVILEGES y AUTHENTICATED como categorías excluidas. De lo contrario, estas categorías se registran en cada solicitud válida al clúster.

Name	Configuración del backend	Description (Descripción)
Capa REST	enable_rest	Habilite o deshabilite los eventos que se producen en la capa REST.
Categorías REST deshabilitadas	disabled_rest_categories	Especifique las categorías de auditoría que se omitirán en la capa REST. La modificación de estas categorías puede aumentar drásticamente el tamaño de los registros de auditoría.
Capa de transporte	enable_transport	Habilite o deshabilite eventos que ocurren en la capa de transporte.
Categorías de transporte deshabilitadas	disabled_transport_categories	Especifique las categorías de auditoría que deben omitirse en la capa de transporte. La modificación de estas categorías puede aumentar drásticamente el tamaño de los registros de auditoría.

La configuración de atributos permite personalizar la cantidad de detalles en cada línea de registro.

Name	Configuración del backend	Description (Descripción)
Solicitudes masivas	resolve_bulk_requests	Habilitar esta configuración genera un registro para cada documento de una solicitud masiva, lo que puede aumentar drásticamente el tamaño de los registros de auditoría.

Name	Configuración del backend	Description (Descripción)
Cuerpo de la solicitud	log_request_body	Incluya el cuerpo de la solicitud de las solicitudes.
Resolver índices	resolve_indices	Resolver alias a índices.

Utilice la configuración de ignorar para excluir un conjunto de usuarios o rutas de API:

Name	Configuración del backend	Description (Descripción)
Usuarios ignorados	ignore_users	Especifique los usuarios que desea excluir.
Solicitudes ignoradas	ignore_requests	Especifique los patrones de solicitud que desea excluir.

Configuración de la conformidad

La configuración de la conformidad permite ajustar el acceso a nivel del índice, del documento o del campo.

Name	Configuración del backend	Description (Descripción)
Registro de conformidad	enable_compliance	Para habilitar o deshabilitar el registro de conformidad.

Puede especificar la siguiente configuración para el registro de eventos de lectura y escritura.

Name	Configuración del backend	Description (Descripción)
Registro de configuración interno	internal_config	Habilite o deshabilite el registro de eventos en el índice `.opendistro_security` .

Puede especificar la siguiente configuración para el registro de eventos de lectura.

Name	Configuración del backend	Description (Descripción)
Metadatos de lectura	read_metadata_only	Incluya solo metadatos para eventos de lectura. No incluya ningún campo de documento.
Usuarios ignorados	read_ignore_users	No incluya ciertos usuarios para eventos de lectura.
Campos observados	read_watched_fields	Especifique los índices y campos que se van a observar en busca de eventos de lectura. Agregar campos vigilados genera un registro por cada acceso a los documentos, lo que puede aumentar drásticamente el tamaño de los registros de auditoría. Los campos observados admiten patrones de índice y patrones de campo: `{ "index-name-pattern": [ "field-name-pattern" ], "logs*": [ "message" ], "twitter": [ "id", "user*" ] }`

Puede especificar la siguiente configuración para eventos de escritura.

Name	Configuración del backend	Description (Descripción)
Metadatos de escritura	write_metadata_only	Incluya solo metadatos para eventos de escritura. No incluya ningún campo de documento.
Diferencias de registro	write_log_diffs	Si write_metadata_only es falso, incluya solo las diferencias entre los eventos de escritura.
Usuarios ignorados	write_ignore_users	No incluya ciertos usuarios para eventos de escritura.
Índices de seguimiento	write_watched_indices	Especifique los índices o patrones de índice para observar los eventos de escritura. Agregar campos vigilados genera un registro por cada acceso a los documentos, lo que puede aumentar drásticamente el tamaño de los registros de auditoría.

Ejemplo de registro de auditoría

Esta sección incluye un ejemplo de configuración, una petición de búsqueda y el registro de auditoría resultante para todos los eventos de lectura y escritura de un índice.

Paso 1: configurar los registros de auditoría

Después de habilitar la publicación de los registros de auditoría en un grupo de CloudWatch registros, vaya a la página de registro de auditorías del OpenSearch panel de control y seleccione Habilitar el registro de auditoría.

1. En Configuración general, seleccione Configurar y asegúrese de que la capa REST esté habilitada.
2. En Configuración de la conformidad, seleccione Configurar.
3. En Escritura, en Campos observados, agregue `accounts` para todos los eventos de escritura en este índice.
4. En Lectura, en Campos observados, agregue los campos `ssn` y `id-` del índice `accounts`:

```
{
  "accounts-": [
    "ssn",
    "id-"
  ]
}
```

Paso 2: realizar eventos de lectura y escritura

1. Vaya a los OpenSearch paneles, elija Herramientas de desarrollo e indexe un documento de muestra:

```
PUT accounts/_doc/0
{
  "ssn": "123",
  "id-": "456"
}
```

2. Para probar un evento de lectura, envíe la siguiente solicitud:

```
GET accounts/_search
{
  "query": {
    "match_all": {}
  }
}
```

Paso 3: observar los registros

1. Abra la CloudWatch consola en. https://console.aws.amazon.com/cloudwatch/
2. En el panel de navegación, seleccione Grupos de registro.
3. Seleccione el grupo de registros que especificó al habilitar registros de auditoría. Dentro del grupo de registros, OpenSearch Service crea un flujo de registros para cada nodo de su dominio.
4. En Flujos de registros, seleccione Buscar todos.
5. Para los eventos de lectura y escritura, consulte los registros correspondientes. Puede existir una demora de 5 segundos antes de que aparezca el registro.

Ejemplo de registro de auditoría de escritura


```
{
  "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"
  ]
}
```


Ejemplo de registro de auditoría de lectura


```
{
  "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"
  ]
}
```

Para incluir el cuerpo de la solicitud, vuelva a la configuración de conformidad en los OpenSearch paneles y desactive Escribir metadatos. Para excluir eventos de un usuario específico, agregue el usuario a Usuarios ignorados.

Para obtener una descripción de cada campo de registro de auditoría, consulte la Referencia de campo del registro de auditoría. Para obtener información sobre la búsqueda y el análisis de los datos del registro de auditoría, consulte Análisis de los datos de registro con CloudWatch Logs Insights en la Guía del usuario de Amazon CloudWatch Logs.

Configuración de registros de auditoría mediante la API REST

Recomendamos usar OpenSearch paneles de control para configurar los registros de auditoría, pero también puede usar la API REST de control de acceso detallada. Esta sección contiene una solicitud de ejemplo. La documentación completa sobre la API REST está disponible en la documentación. 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"
    ]
  }
}
```

Monitorización OpenSearch de eventos del servicio con Amazon EventBridge

Amazon OpenSearch Service se integra con Amazon EventBridge para notificarte determinados eventos que afectan a tus dominios. Los eventos de AWS los servicios se envían casi EventBridge

en tiempo real. Los mismos eventos también se envían a Amazon CloudWatch Events, la predecesora de Amazon EventBridge. Puede crear reglas sencillas para indicar qué eventos le resultan de interés, así como qué acciones automatizadas se van a realizar cuando un evento cumple una de las reglas. Entre las acciones que se pueden activar automáticamente se incluyen las siguientes:

- Invocar una función AWS Lambda
- Invocar un Run Command de Amazon EC2
- Desviar el evento a Amazon Kinesis Data Streams
- Activación de una máquina de estados AWS Step Functions
- Notificar un tema de Amazon SNS o una cola de Amazon SQS

Para obtener más información, consulta Cómo empezar con Amazon EventBridge en la Guía del EventBridge usuario de Amazon.

Temas

- Eventos de actualización del software de servicio
- Eventos de ajuste automático
- Eventos del estado del clúster
- Eventos del punto de conexión de VPC
- Eventos de retirada de nodos
- Eventos de retirada de nodos degradados
- Eventos de error de dominio
- Tutorial: Cómo escuchar los EventBridge eventos OpenSearch de Amazon Service
- Tutorial: Sending Amazon SNS alerts for available software updates

Eventos de actualización del software de servicio

OpenSearch El servicio envía los eventos EventBridge cuando se produce uno de los siguientes eventos de actualización del software del servicio.

Actualización del software de servicio disponible

OpenSearch El servicio envía este evento cuando hay disponible una actualización del software del servicio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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"
  }
}
```

Actualización del software de servicio programada

OpenSearch El servicio envía este evento cuando se ha programado una actualización del software del servicio. En el caso de las actualizaciones opcionales, recibirá la notificación en la fecha programada y tendrá la opción de volver a programarla en cualquier momento. En el caso de las actualizaciones obligatorias, recibirá la notificación tres días antes de la fecha programada y tendrá la opción de reprogramarla dentro del plazo obligatorio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Actualización del software de servicio reprogramada

OpenSearch El servicio envía este evento cuando se ha reprogramado una actualización opcional del software del servicio. Para obtener más información, consulte the section called “Actualizaciones opcionales frente a actualizaciones obligatorias”.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Actualización del software de servicio comenzada

OpenSearch El servicio envía este evento cuando se inicia una actualización del software del servicio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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.
  }
}
```

Actualización del software de servicio finalizada

OpenSearch El servicio envía este evento cuando se completa la actualización del software del servicio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Actualización del software de servicio cancelada

OpenSearch El servicio envía este evento cuando se cancela una actualización del software del servicio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
```

```
  }
}
```

Actualización programada del software de servicio cancelada

OpenSearch El servicio envía este evento cuando se cancela una actualización del software del servicio que estaba programada previamente para el dominio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Actualización del software de servicio no ejecutada

OpenSearch El servicio envía este evento cuando no puede iniciar una actualización del software del servicio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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]"
  }
}
```

Error al actualizar el software del servicio

OpenSearch El servicio envía este evento cuando se produce un error en la actualización del software del servicio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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].
  }
}
```

Se requiere actualización del software de servicio

OpenSearch El servicio envía este evento cuando se requiere una actualización del software del servicio. Para obtener más información, consulte the section called “Actualizaciones opcionales frente a actualizaciones obligatorias”.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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"
  }
}
```

Eventos de ajuste automático

OpenSearch El servicio envía los eventos EventBridge cuando se produce uno de los siguientes eventos de Auto-Tune.

Ajuste automático pendiente

OpenSearch El servicio envía este evento cuando Auto-Tune ha identificado recomendaciones de ajuste para mejorar el rendimiento y la disponibilidad del clúster. Solo verá este evento para dominios con el ajuste automático deshabilitado.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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}"
  }
}
```

Ajuste automático iniciado

OpenSearch El servicio envía este evento cuando Auto-Tune comienza a aplicar nuevas configuraciones a su dominio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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 requiere una implementación programada blue/green

OpenSearch El servicio envía este evento cuando Auto-Tune ha identificado recomendaciones de ajuste que requieren una implementación programada blue/green .

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Ajuste automático cancelado

OpenSearch El servicio envía este evento cuando se cancela la programación de Auto-Tune porque no hay recomendaciones de ajuste pendientes.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Ajuste automático completado

OpenSearch El servicio envía este evento cuando Auto-Tune ha completado la blue/green implementación y el clúster está operativo con la nueva configuración de JVM implementada.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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%}."
  }
}
```

Ajuste automático deshabilitado y cambios revertidos

OpenSearch El servicio envía este evento cuando se ha desactivado el ajuste automático y se han revertido los cambios aplicados.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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}"
  }
}
```

Ajuste automático deshabilitado y cambios conservados

OpenSearch El servicio envía este evento cuando se ha desactivado el ajuste automático y se han conservado los cambios aplicados.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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}"
  }
}
```

Eventos del estado del clúster

OpenSearch El servicio envía ciertos eventos EventBridge cuando el estado del clúster se ve comprometido.

Red cluster recovery started (Recuperación de clúster en rojo iniciada)

OpenSearch El servicio envía este evento después de que el estado del clúster haya estado en rojo de forma continua durante más de una hora. Intenta restaurar automáticamente uno o varios índices en rojo a partir de una instantánea para corregir el estado del clúster.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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]"
   }
}
```

Red cluster recovery partially completed (Recuperación de clúster en rojo completada parcialmente)

OpenSearch El servicio envía este evento cuando solo pudo restaurar un subconjunto de índices rojos a partir de una instantánea mientras intentaba corregir el estado de un clúster rojo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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."
   }
}
```

Red cluster recovery failed (Error de recuperación de clúster en rojo)

OpenSearch El servicio envía este evento cuando no puede restaurar ningún índice mientras intenta corregir el estado de un clúster rojo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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."
   }
}
```

Shards to be deleted (Se deben eliminar particiones)

OpenSearch El servicio envía este evento cuando intenta corregir automáticamente el estado del clúster rojo después de haber estado rojo de forma continua durante 14 días, pero uno o más índices permanecen en rojo. Al cabo de 7 días más (21 días en total si el número está en rojo de forma continua), el OpenSearch Servicio procederá a eliminar los fragmentos no asignados de todos los índices rojos.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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"
   }
}
```

Shards deleted (Particiones eliminadas)

OpenSearch El servicio envía este evento después de que el estado del clúster haya estado en rojo de forma continua durante 21 días. Procede a eliminar las particiones no asignadas (almacenamiento y computación) de todos los índices en rojo. Para más información, consulte the section called "Corrección automática de clústeres en rojo".

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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"
   }
}
```

Advertencia de alto recuento de particiones

OpenSearch El servicio envía este evento cuando el número medio de fragmentos en los nodos de datos activos supera el 90% del límite predeterminado recomendado de 1000. Si bien las versiones

posteriores de Elasticsearch OpenSearch admiten un límite máximo configurable de particiones por nodo, le recomendamos que no tenga más de 1000 particiones por nodo. Consulte Elección del número de particiones.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Límite de recuento de particiones superado

OpenSearch El servicio envía este evento cuando el número medio de particiones en los nodos de datos activos supera el límite predeterminado recomendado de 1000. Si bien las versiones posteriores de Elasticsearch OpenSearch admiten un límite máximo configurable de particiones por nodo, le recomendamos que no tenga más de 1000 particiones por nodo. Consulte Elección del número de particiones.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Espacio en disco reducido

OpenSearch El servicio envía este evento cuando uno o más nodos del clúster tienen menos del 25% del espacio de almacenamiento disponible o menos de 25 GB.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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"
  }
}
```

Vulneración del límite de espacio en disco mínimo

OpenSearch El servicio envía este evento cuando todos los nodos del clúster tienen menos del 10% del espacio de almacenamiento disponible o menos de 10 GB. Si todos los nodos superan el límite de disco bajo, cualquier índice nuevo da como resultado un clúster amarillo, y si todos los nodos se sitúan por debajo del límite de disco alto, aparece un clúster rojo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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".
  }
}
```

Balance de ráfaga de EBS por debajo del 70 %

OpenSearch El servicio envía este evento cuando el saldo de ráfagas de EBS en uno o más nodos de datos cae por debajo del 70%. El agotamiento del balance de ráfagas de EBS puede provocar una falta generalizada de disponibilidad de los clústeres y una limitación de I/O las solicitudes, lo que puede provocar altas latencias y tiempos de espera en las solicitudes de indexación y búsqueda. Para conocer los pasos para solucionar este problema, consulte the section called “Bajo balance de ráfaga EBS”.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Balance de ráfaga de EBS por debajo del 20 %

OpenSearch El servicio envía este evento cuando el saldo de ráfagas de EBS en uno o más nodos de datos cae por debajo del 20%. El agotamiento del balance de ráfagas de EBS puede provocar una falta generalizada de disponibilidad de los clústeres y una limitación de I/O las solicitudes, lo que puede provocar altas latencias y tiempos de espera en las solicitudes de indexación y búsqueda. Para conocer los pasos para solucionar este problema, consulte the section called “Bajo balance de ráfaga EBS”.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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.
  }
}
```

Limitación de rendimiento de disco

OpenSearch El servicio envía este evento cuando las solicitudes de lectura y escritura en su dominio se ven limitadas debido a las limitaciones de rendimiento de los volúmenes de EBS o de la instancia de EC2. Si recibe esta notificación, considere la posibilidad de ampliar sus volúmenes o instancias siguiendo las prácticas recomendadas. AWS Si su tipo de volumen es gp2, aumente el tamaño del volumen. Si su tipo de volumen es gp3, aprovisione más rendimiento. También puede comprobar que la base de instancias y el rendimiento máximo de EBS sean superiores o iguales al rendimiento del volumen aprovisionado, y puede escalar verticalmente en consecuencia.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
   }
}
```

Tamaño de partición grande

OpenSearch El servicio envía este evento cuando uno o más fragmentos del clúster superan los 50 GiB o los 65 GiB. Para garantizar un rendimiento y una estabilidad óptimos del clúster, reduzca el tamaño de las particiones.

Para obtener más información, consulte las prácticas recomendadas de particiones.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Alto uso de JVM

OpenSearch El servicio envía este evento cuando la `JVMMemoryPressure` métrica de tu dominio supera el 80%. Si supera el 92 % durante 30 minutos, se bloquearán todas las operaciones de escritura en el clúster. Para garantizar una estabilidad óptima del clúster, reduzca el tráfico al clúster o escale su dominio para proporcionar memoria suficiente para su carga de trabajo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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 insuficiente

OpenSearch El servicio envía este evento cuando la JVM máxima es superior al 70% y la diferencia entre el máximo y el mínimo es inferior al 30%. Esto puede indicar que la JVM no puede recuperar suficiente memoria para la carga de trabajo durante los ciclos de recopilación de elementos no utilizados. Esto puede provocar respuestas cada vez más lentas y latencias más altas y, en algunos casos, incluso la caída de nodos debido a que se ha agotado el tiempo de espera de las comprobaciones de estado. Para garantizar una estabilidad óptima del clúster, reduzca el tráfico al clúster o escale su dominio para proporcionar memoria suficiente para su carga de trabajo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Advertencia de enrutamiento indexado personalizada

OpenSearch El servicio envía este evento cuando el dominio está en proceso y contiene índices con una configuración personalizada de index.routing.location, lo que puede provocar que las implementaciones de color azul-verde se bloqueen. Compruebe que la configuración se aplique correctamente.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Fallo en el bloqueo de partición

OpenSearch El servicio envía este evento cuando tu dominio está en mal estado debido a fragmentos no asignados. `[ShardLockObtainFailedException]` Para obtener más información, consulta ¿Cómo resuelvo la excepción de bloqueo de fragmentos en memoria en Amazon OpenSearch Service?

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
}
```

Eventos del punto de conexión de VPC

OpenSearch El servicio envía ciertos eventos EventBridge relacionados con los puntos finales de la AWS PrivateLink interfaz.

Error al crear el punto de conexión de VPC

OpenSearch El servicio envía este evento cuando no puede crear un punto final de VPC solicitado. Este error puede ocurrir porque alcanzó el límite de la cantidad de puntos de conexión de VPC permitidos dentro de una región. También verá este error si no existe una subred ni un grupo de seguridad especificados.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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."
   }
}
```

Error al actualizar el punto de conexión de VPC

OpenSearch El servicio envía este evento cuando no puede eliminar un punto final de VPC solicitado.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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>."
   }
}
```

Error al eliminar el punto de conexión de VPC

OpenSearch El servicio envía este evento cuando no puede eliminar un punto final de VPC solicitado.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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."
    }
}
```

Eventos de retirada de nodos

OpenSearch El servicio envía los eventos EventBridge cuando se produce uno de los siguientes eventos de retirada de nodos.

Retirada de nodos programada

OpenSearch El servicio envía este evento cuando se ha programado la retirada de un nodo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Se completó la retirada del nodo

OpenSearch El servicio envía este evento cuando se ha completado la retirada de un nodo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Error al retirar el nodo

OpenSearch El servicio envía este evento cuando se produce un error al retirar un nodo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Eventos de retirada de nodos degradados

OpenSearch El servicio envía estos eventos cuando es necesario reemplazar un nodo debido a la degradación del hardware de un nodo.

Notificación de retirada de nodos degradados

OpenSearch El servicio envía este evento cuando la acción automática para retirar y reemplazar un nodo degradado se ha programado para su dominio.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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"
   }
}
```

Finalización de la retirada de nodos degradados

OpenSearch El servicio envía este evento cuando un nodo degradado se retira y se reemplaza por un nodo nuevo.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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"
   }
}
```

Errores de retirada de nodos degradados

OpenSearch El servicio envía este evento si se produce un error al retirar el nodo degradado.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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"
   }
}
```

Eventos de error de dominio

OpenSearch El servicio envía los eventos EventBridge cuando se produce uno de los siguientes errores de dominio.

Error en la validación de la actualización del dominio

OpenSearch El servicio envía este evento si detecta uno o más errores de validación al intentar actualizar o realizar un cambio de configuración en un dominio. Para conocer los pasos para solucionar estos errores, consulte the section called “Solución de errores de validación”.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
   "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"
   }
}
```

Clave KMS inaccesible

OpenSearch El servicio envía este evento cuando no puede acceder a tu AWS KMS clave.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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."
  }
}
```

Aislamiento de dominios

OpenSearch El servicio envía este evento cuando tu dominio queda aislado y no puede recibir, leer ni escribir solicitudes porque la red no puede acceder a él.

Ejemplo

A continuación se muestra un ejemplo de este evento:

```
{
  "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: Cómo escuchar los EventBridge eventos OpenSearch de Amazon Service

En este tutorial, configuras una AWS Lambda función sencilla que escucha los eventos de Amazon OpenSearch Service y los escribe en un flujo de registro de CloudWatch Logs.

Requisitos previos

En este tutorial se presupone que ya tienes un dominio OpenSearch de servicio. Si no creó un dominio, siga los pasos de Creación y administración de dominios para crear uno.

Paso 1: crear la función Lambda

En este procedimiento, se crea una función Lambda sencilla que sirva de destino para los mensajes de eventos de OpenSearch servicio.

Para crear una función de Lambda de destino

1. Abra la AWS Lambda consola en. https://console.aws.amazon.com/lambda/
2. Seleccione Crear función y Crear desde cero.
3. En Nombre de la función, ingrese event-handler.
4. En Tiempo de ejecución, seleccione Python 3.8.
5. Seleccione Crear función.
6. En la sección Código de función, edite el código de muestra de tal modo que coincida con el siguiente ejemplo:

```
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))
```

Esta es una función simple de Python 3.8 que imprime los eventos enviados por OpenSearch Service. Si todo está configurado correctamente, al final de este tutorial, los detalles del evento aparecen en el flujo de registro de CloudWatch Logs asociado a esta función de Lambda.

7. Seleccione Implementar.

Paso 2: registrar una regla de eventos

En este paso, creará una EventBridge regla que capture los eventos de sus dominios de OpenSearch servicio. Esta regla captura todos los eventos dentro de la cuenta, en la cual se define. Los propios mensajes de eventos contienen información acerca de la fuente del evento, incluido el dominio desde el que se originó. Puede utilizar esta información para filtrar y ordenar eventos de manera programática.

Para crear una EventBridge regla

1. Abra la EventBridge consola en https://console.aws.amazon.com/events/.
2. Seleccione Crear regla.
3. Nombre la regla event-rule.
4. Elija Siguiente.
5. Para el patrón de eventos, selecciona AWS services, Amazon OpenSearch Service y All Events. Este patrón se aplica a todos los dominios de tu OpenSearch Servicio y a todos los eventos del OpenSearch Servicio. También puede crear un patrón más específico para filtrar algunos resultados.
6. Pulse Siguiente.
7. Para el destino, seleccione Función de Lambda. En el menú desplegable de funciones, elija event-handler.
8. Pulse Siguiente.
9. Omita las etiquetas y vuelva a pulsar Siguiente.
10. Revise la configuración y elija Crear regla.

Paso 3: probar la configuración

La próxima vez que reciba una notificación en la sección Notificaciones de la consola de OpenSearch servicio, si todo está configurado correctamente, se activará la función Lambda y escribirá los datos del evento en un flujo de registro de CloudWatch Logs para la función.

Para probar la configuración

1. Abra la CloudWatch consola en. https://console.aws.amazon.com/cloudwatch/

2. En el panel de navegación, elija Registros y seleccione el grupo de registros para la función Lambda (por ejemplo,/aws/lambda/event-handler).
3. Seleccione una secuencia de registro para ver los datos de los eventos.

Tutorial: Sending Amazon SNS alerts for available software updates

En este tutorial, configuras una regla de EventBridge eventos de Amazon que captura las notificaciones de las actualizaciones de software de servicio disponibles en Amazon OpenSearch Service y te envía una notificación por correo electrónico a través de Amazon Simple Notification Service (Amazon SNS).

Requisitos previos

En este tutorial se presupone que ya tiene un dominio OpenSearch de servicio. Si no creó un dominio, siga los pasos de Creación y administración de dominios para crear uno.

Paso 1: crear y suscribirse a un tema de Amazon SNS

Configure un tema de Amazon SNS para servir como destino de eventos para la nueva regla de eventos.

Para crear un destino de Amazon SNS

1. Abra la consola Amazon SNS en https://console.aws.amazon.com/sns/ la versión 3/home.
2. Seleccione Temas y Crear tema.
3. Para el tipo de tarea, seleccione Estándar y nombre el trabajo software-update.
4. Seleccione Crear nuevo tema.
5. Una vez creado el tema, seleccione Crear suscripción.
6. En Protocolo, seleccione Correo electrónico. En Punto de conexión, ingrese una dirección de email a la que actualmente tenga acceso y elija Crear suscripción.
7. Consulte su cuenta de email y espere para recibir un mensaje de correo electrónico de confirmación de la suscripción. Cuando lo reciba, seleccione Confirmar suscripción.

Paso 2: registrar una regla de eventos

A continuación, registre una regla de eventos que solo capture los eventos de actualización del software de servicio.

Para crear una regla de eventos

1. Abra la consola en EventBridge . https://console.aws.amazon.com/events/
2. Seleccione Crear regla.
3. Nombre la regla softwareupdate-rule.
4. Elija Siguiente.
5. Para el patrón de eventos, selecciona AWS servicios, Amazon OpenSearch Service (clúster gestionado) y Amazon OpenSearch Service Software Update Notification. Este patrón coincide con cualquier evento de actualización del software del OpenSearch servicio realizado por Service. Para obtener más información sobre los patrones de eventos, consulta los patrones de EventBridge eventos de Amazon en la Guía del EventBridge usuario de Amazon.
6. Si lo desea, puede filtrar solo por gravedades específicas. Para conocer las gravedades de cada evento, consulte the section called “Eventos de actualización del software de servicio”.
7. Seleccione Siguiente.
8. Para el destino, seleccione Tema de SNS y seleccione software-update.
9. Seleccione Siguiente.
10. Omita las etiquetas y elija Siguiente.
11. Revise la configuración de la regla y elija Crear regla.

La próxima vez que reciba una notificación del OpenSearch Servicio sobre una actualización de software de servicio disponible, si todo está configurado correctamente, Amazon SNS debería enviarle una alerta por correo electrónico sobre la actualización.

Supervisión de las llamadas a la API de Amazon OpenSearch Service con AWS CloudTrail

Amazon OpenSearch Service se integra con AWS CloudTrail un servicio que proporciona un registro de las acciones realizadas por un usuario, un rol o un AWS servicio del OpenSearch Servicio. CloudTrail captura todas las llamadas a la API de configuración OpenSearch del Servicio como eventos.

> **Note**
>
> CloudTrail solo captura las llamadas a la API de configuración, como `CreateDomain` y`GetUpgradeStatus`. CloudTrail no captura las llamadas a OpenSearch APIs, como

`_search` y `_bulk`. Para estas llamadas, consulte the section called “Monitoreo de registros de auditoría”.

Las llamadas capturadas incluyen llamadas desde la consola de OpenSearch servicio o desde un AWS SDK. AWS CLI Si crea una ruta, puede habilitar la entrega continua de CloudTrail eventos a un bucket de Amazon S3, incluidos los eventos de OpenSearch Service. Si no configura una ruta, podrá ver los eventos más recientes en la CloudTrail consola, en el historial de eventos. Con la información recopilada por usted CloudTrail, puede determinar la solicitud que se realizó al OpenSearch Servicio, la dirección IP desde la que se realizó la solicitud, quién la hizo, cuándo se realizó y detalles adicionales.

Para obtener más información CloudTrail, consulte la Guía AWS CloudTrail del usuario.

Información OpenSearch de Amazon Service en CloudTrail

CloudTrail está habilitada en tu cuenta Cuenta de AWS al crear la cuenta. Cuando se produce una actividad en el OpenSearch Servicio, esa actividad se registra en un CloudTrail evento junto con otros eventos de AWS servicio en el historial de eventos. Se puede ver, buscar y descargar los últimos eventos de la cuenta de Cuenta de AWS . Para obtener más información, consulte Visualización de eventos con el historial de CloudTrail eventos.

Para tener un registro continuo de los eventos de tu Cuenta de AWS cuenta, incluidos los eventos del OpenSearch Servicio, crea un registro. Un rastro permite CloudTrail entregar archivos de registro a un bucket de Amazon S3. De forma predeterminada, cuando se crea un registro de seguimiento en la consola, el registro de seguimiento se aplica a todas las Regiones de AWS. La ruta registra los eventos de todas las regiones de la AWS partición y envía los archivos de registro al bucket de Amazon S3 que especifique. Además, puede configurar otros AWS servicios para analizar más a fondo los datos de eventos recopilados en los CloudTrail registros y actuar en función de ellos. Para obtener más información, consulte los siguientes temas:

- Creando una ruta para tu Cuenta de AWS
- AWS integraciones de servicios con Logs CloudTrail
- Configuración de las notificaciones de Amazon SNS para CloudTrail
- Recibir archivos de CloudTrail registro de varias regiones y recibir archivos de CloudTrail registro de varias cuentas

Todas las acciones de la API de configuración de OpenSearch servicios se registran CloudTrail y se documentan en la referencia de la API de Amazon OpenSearch Service.

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. La información de identidad del usuario le ayuda a determinar lo siguiente:

- Si la solicitud se realizó con credenciales de usuario root o AWS Identity and Access Management (IAM)
- si la solicitud se realizó con credenciales de seguridad temporales de un rol o fue un usuario federado
- Si la solicitud la realizó otro servicio AWS

Para obtener más información, consulte el Elemento userIdentity de CloudTrail .

Descripción de las entradas de los archivos de registro de Amazon OpenSearch Service

Un rastro es una configuración que permite la entrega de eventos como archivos de registro a un bucket de Amazon S3 que especifique. CloudTrail Los archivos de registro contienen una o más entradas de registro. Un evento representa una solicitud única de cualquier fuente e incluye información sobre la acción solicitada, la fecha y la hora de la acción, los parámetros de la solicitud, etc. CloudTrail Los archivos de registro no son un registro ordenado de las llamadas a la API pública, por lo que no aparecen en ningún orden específico.

El siguiente ejemplo muestra una entrada de CloudTrail registro que demuestra la `CreateDomain` operación:


```
{
  "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"
}
```

Supervisión operativa unificada con Cluster Insights

Amazon OpenSearch Service ahora incluye Cluster Insights, una solución de monitoreo que proporciona una visibilidad operativa completa de sus clústeres a través de un único panel de control. Esto elimina la complejidad de tener que analizar y correlacionar varios registros y métricas para identificar los posibles riesgos para la disponibilidad o el rendimiento de los clústeres. La solución automatiza la consolidación de los datos operativos críticos en todos los nodos, índices y fragmentos, lo que transforma la solución de problemas complejos en un proceso simplificado. Puede detectar problemas como fragmentos grandes y marcas de agua bajas en el disco, ver métricas detalladas a nivel de nodo, índice y fragmento, y acceder a las mejores prácticas de seguridad y resiliencia.

> **Note**
>
> Cluster Insights está disponible a través de la interfaz OpenSearch de usuario del servicio sin coste adicional para todos los usuarios que utilicen la OpenSearch versión 2.17 o posterior. Además, los dominios con OpenSearch las versiones 2.17 y 2.19 también deben incluir la última actualización de la versión del software del servicio.

Ventajas

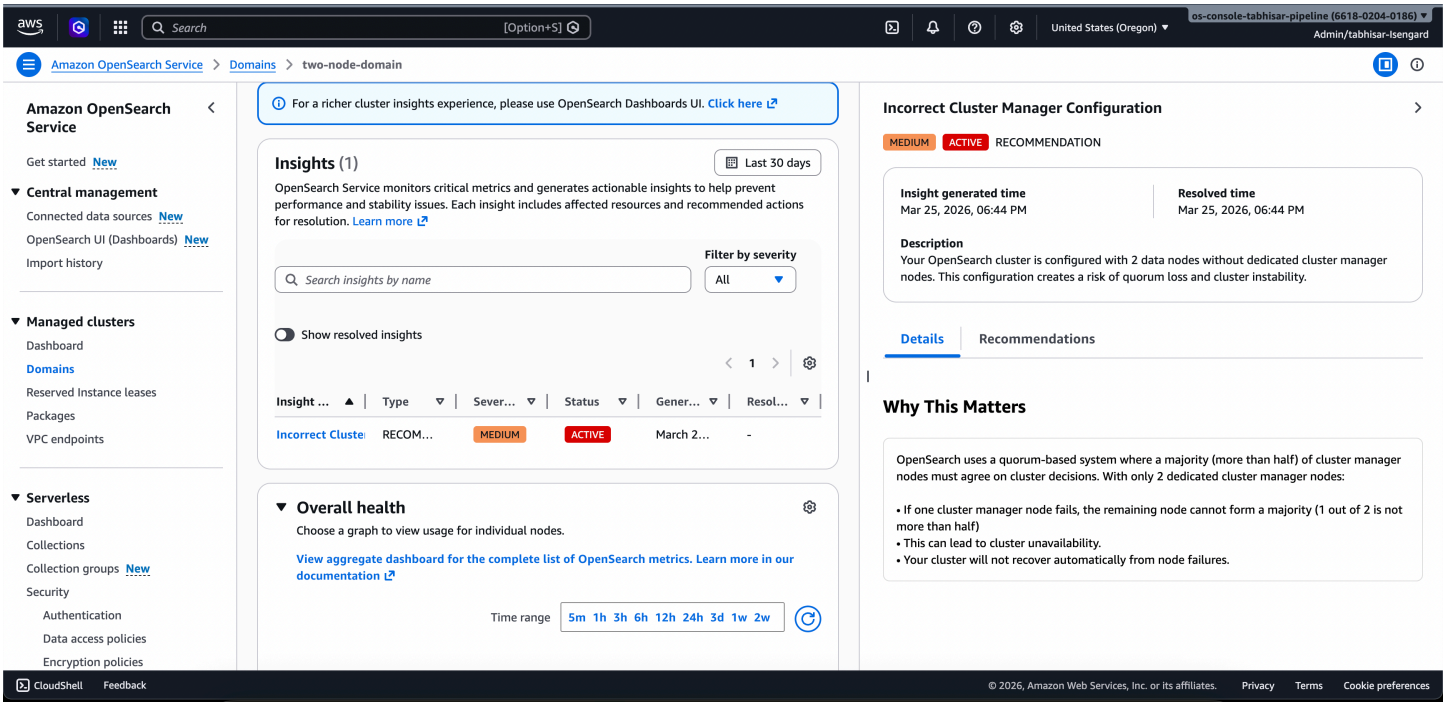
- Supervisión proactiva: supervise el estado del clúster de forma proactiva con métricas de rendimiento detalladas en todos los componentes, desde los nodos e índices individuales hasta los fragmentos y las consultas de búsqueda.
- Visibilidad unificada: consolide los datos de monitoreo en un único panel
- Recomendaciones prácticas: obtenga step-by-step orientación para la resolución de problemas
- Cobertura integral: supervise la seguridad, la estabilidad y la resiliencia en todos sus clústeres OpenSearch
- Optimización de consultas: identifique las consultas que consumen muchos recursos y optimice el rendimiento

Con Cluster Insights, puede mantener un rendimiento óptimo de los clústeres, reducir la sobrecarga operativa y garantizar la coherencia de las mejores prácticas en todos sus clústeres OpenSearch

Acceda a Cluster Insights a través de la consola

Revise las recomendaciones de rendimiento y resiliencia y realice los cambios de configuración necesarios, todo desde la misma consola. En la consola, en la pestaña Estado del clúster, puede acceder a Cluster Insights, que contiene una lista de todas las estadísticas activas. Haga clic en cualquier Insight para ver las recomendaciones.

Pantalla 1: Cluster Insights en la pestaña Cluster Health



Acceda a las estadísticas y métricas detalladas del clúster a través OpenSearch de la interfaz de usuario

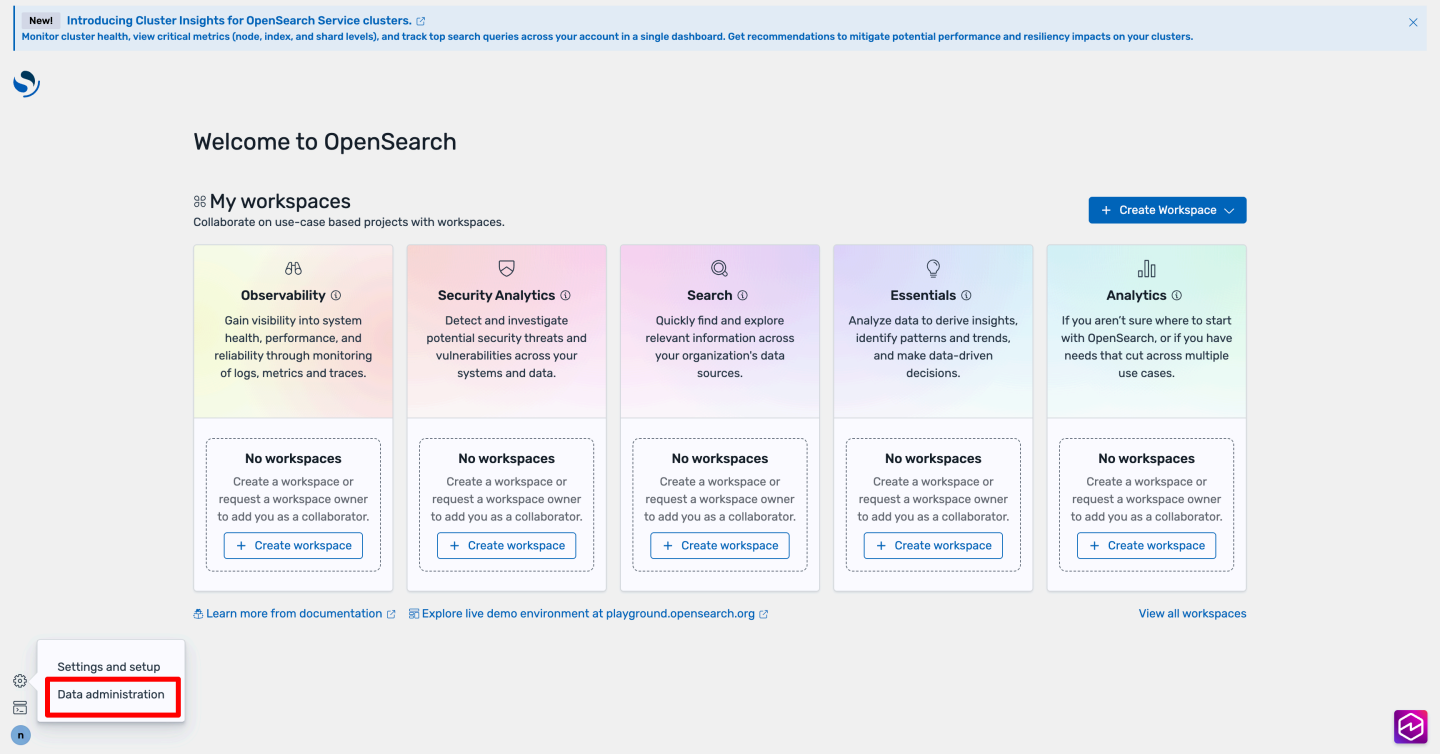
Puede ver la información de un clúster de OpenSearch servicios específico a través de la OpenSearch interfaz de usuario (paneles). En la OpenSearch interfaz de usuario, una aplicación es simplemente una estructura organizativa, como una carpeta. Cada aplicación puede conectarse a varios clústeres de OpenSearch servicios y mostrar información sobre ellos. El acceso a Cluster Insights requiere una función administrativa en la aplicación de OpenSearch interfaz de usuario.

> **Note**
>
> El acceso a Cluster Insights requiere una función administrativa en la aplicación de OpenSearch interfaz de usuario.

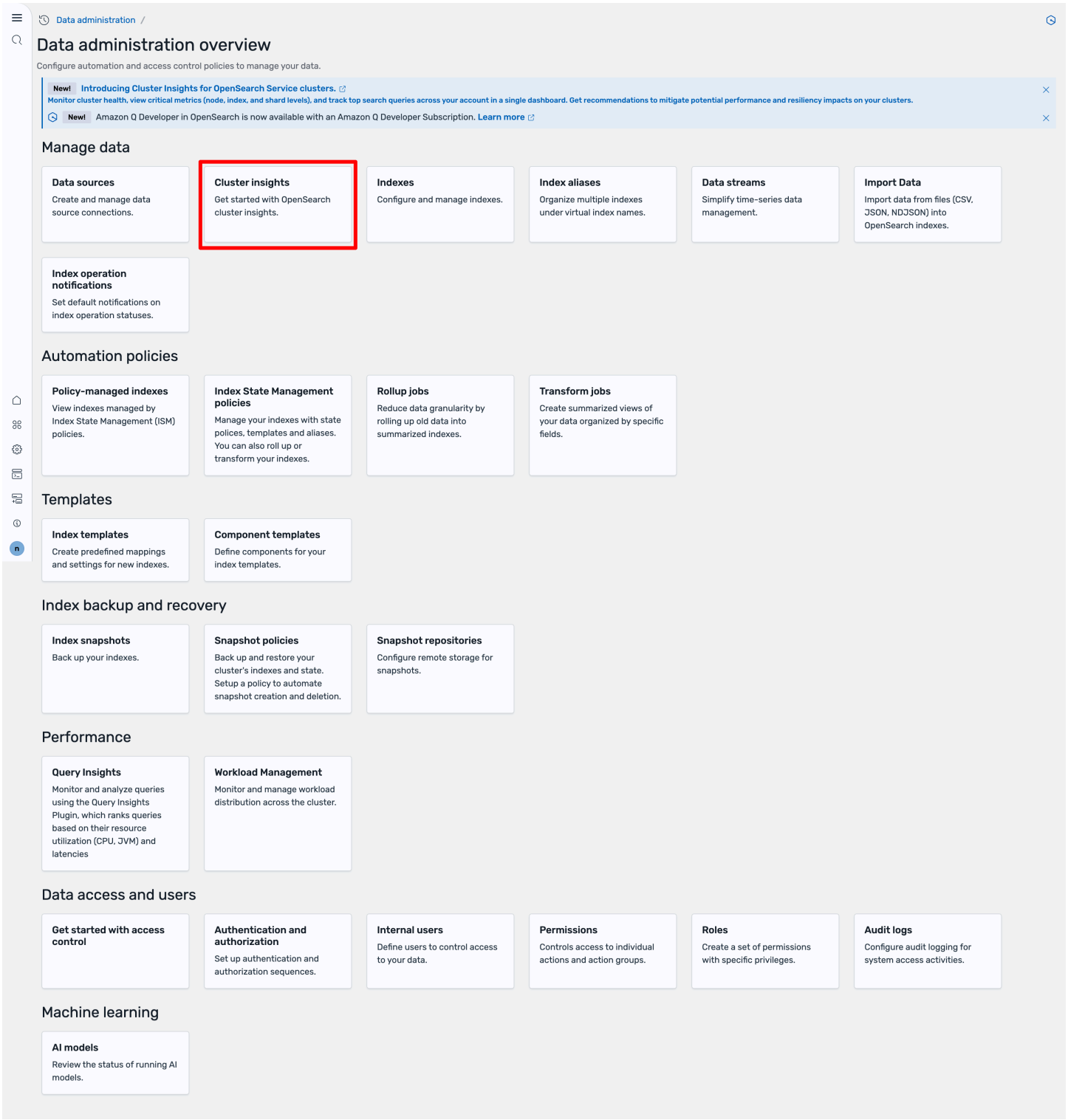
Cree y configure una aplicación para ver Cluster Insights

1. Abra la consola OpenSearch de servicio en https://console.aws.amazon.com/aos/home
2. Seleccione la OpenSearch interfaz de usuario (paneles) en el menú de navegación de la izquierda
3. Complete los siguientes pasos para crear y configurar una aplicación:
    1. Cree una aplicación OpenSearch de servicio
    2. Asocie fuentes de datos
4. Tras completar los dos pasos anteriores, podrás ver Cluster Insights en el panel de la OpenSearch interfaz de usuario, en la sección Configuración > Administrador de datos > Cluster Insights. El icono de configuración se encuentra en la parte inferior izquierda de la pantalla de la OpenSearch interfaz de usuario.

Pantalla 2: accede al administrador de datos desde OpenSearch la interfaz de usuario



Pantalla 3: Cluster Insights en la sección Administrar datos



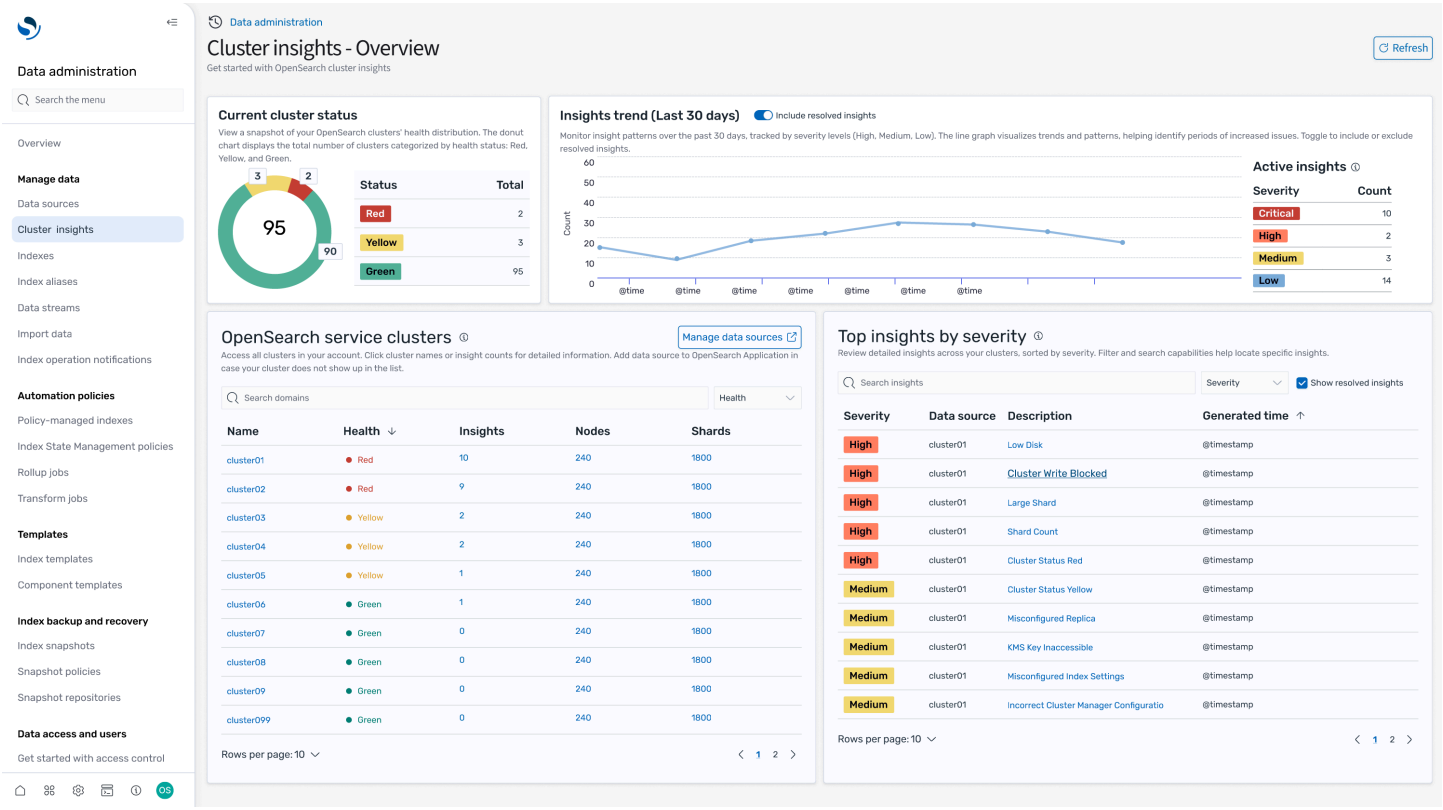
Comprensión de las perspectivas de los

En esta sección se describen los distintos conocimientos disponibles en Cluster Insights.

Panel de información general

La página de descripción general de Cluster Insights, como se muestra en la siguiente captura de pantalla, proporciona una visión general del estado del clúster a nivel de aplicación e incluye las siguientes secciones:

Pantalla 4: página de inicio de Cluster Insights en la aplicación de OpenSearch interfaz de usuario.



Estado actual del clúster

Un gráfico circular muestra el estado de salud del clúster:

- Verde: todos los fragmentos y réplicas principales están asignados a los nodos
- Amarillo: todos los fragmentos principales están asignados, pero algunas réplicas no
- Rojo: al menos un fragmento principal no está asignado a ningún nodo

Tendencia de Insights

El gráfico de tendencias rastrea los patrones de los problemas durante los últimos 30 días, lo que le ayuda a identificar los problemas emergentes y a monitorear el progreso de la resolución.

Perspectivas abiertas actuales

Un recuento organizado según la gravedad de la información pendiente de los últimos 30 días.

OpenSearch Clústeres de servicios

En esta sección se enumeran todos los OpenSearch clústeres con estadísticas clave, como el número de nodos, el número de fragmentos y las consultas activas.

Principales datos desglosados por gravedad

Puede revisar la información de todos los dominios de su aplicación. En esta sección se priorizan los problemas que requieren atención inmediata (críticos y de alta gravedad). Cada información incluye una descripción y recomendaciones específicas, que pueden ayudarle a centrarse primero en los problemas críticos.

Detalles de la información

Cada información de la sección Información principal por gravedad es interactiva y proporciona un análisis detallado. Por ejemplo, si elige la información sobre el tamaño de los fragmentos grandes:

1. Puede ver cuántos fragmentos superan el umbral y qué índices se ven afectados.
2. Un mapa de recursos identifica cada fragmento sobredimensionado con su índice, identificador y tamaño actual.
3. La pestaña de recomendaciones proporciona una guía de step-by-step corrección.
4. La pestaña Historial muestra un cronograma de las acciones de remediación de los recursos.

Detalles de clúster

Al seleccionar un clúster específico en la sección Clústeres de OpenSearch servicios, OpenSearch muestra información sobre ese clúster en las siguientes pestañas: estado del clúster, vista de nodos, vista de índice, vista de fragmentos y vista de consultas. La pestaña Estado del clúster muestra la siguiente información:

Descripción general de

La información clave incluye el estado del clúster, el recuento de fragmentos, el recuento de nodos, el recuento de índices y las estadísticas de los documentos.

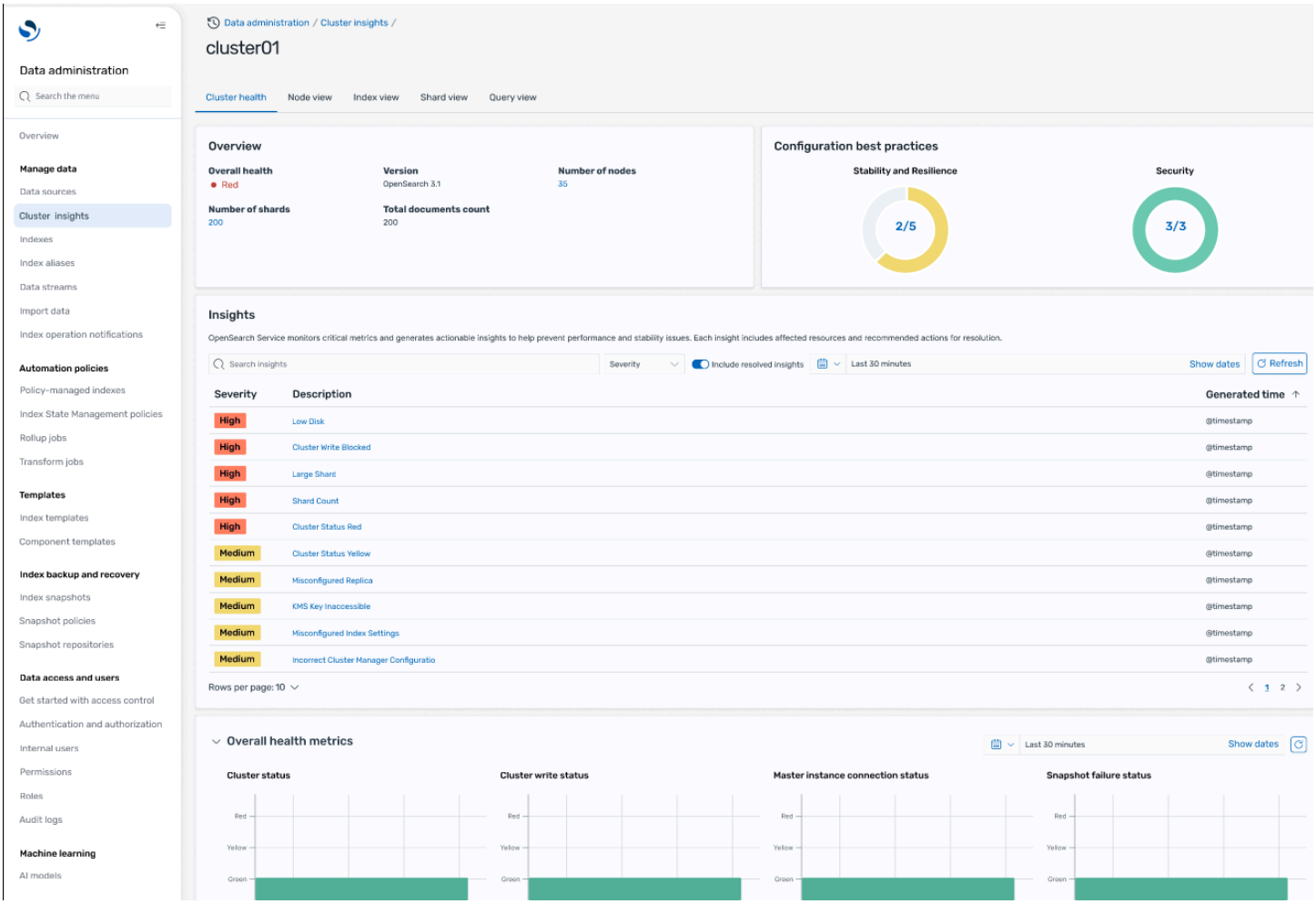
Prácticas recomendadas de configuración

Los gráficos anulares muestran el cumplimiento de las configuraciones recomendadas en materia de resiliencia y seguridad.

Información

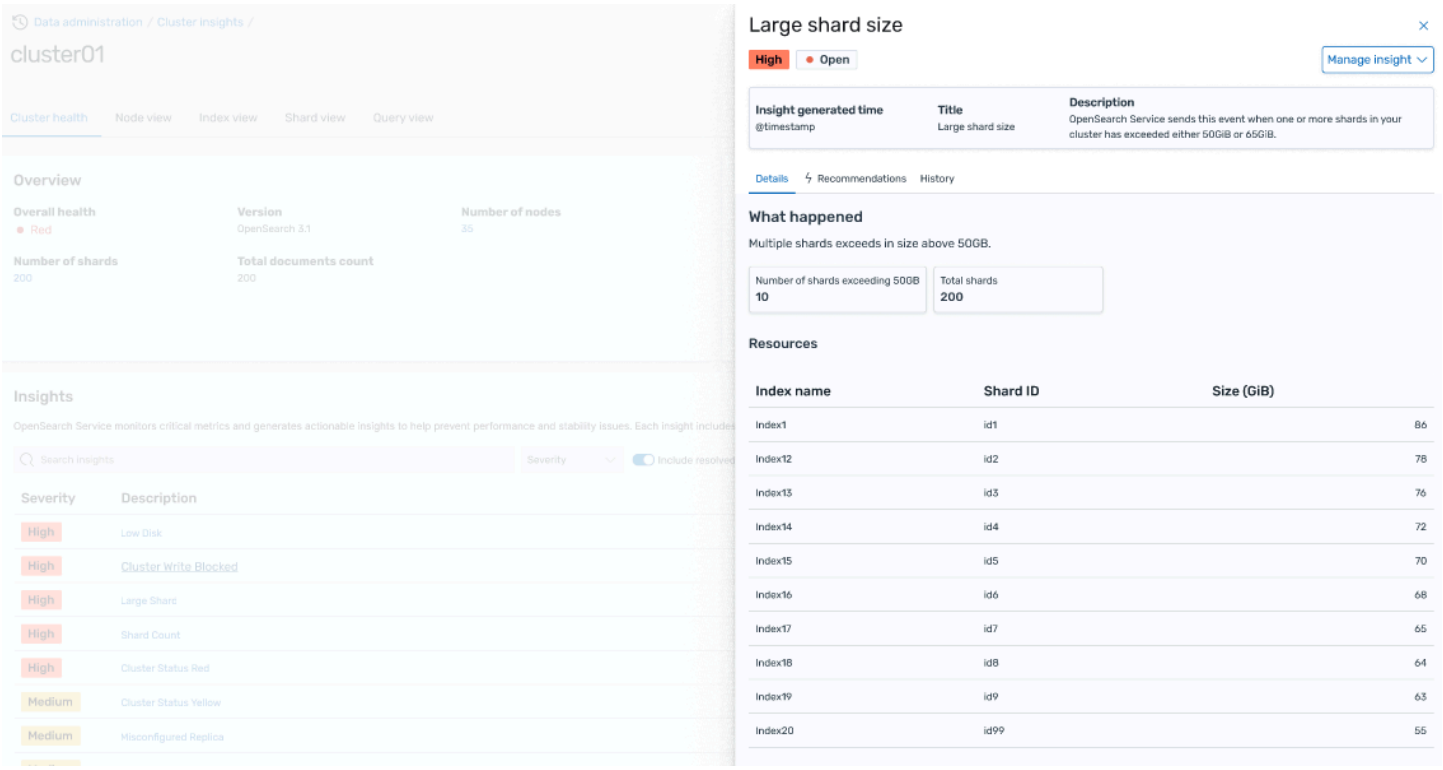
En una tabla se enumeran los datos recientes generados para el clúster, con el mismo desglose detallado y las mismas directrices de corrección disponibles en la página de información general.

La descripción general de Screen-5: Cluster Health proporciona métricas clave, mejores prácticas e información



Al hacer clic en cualquier información, puede ver los detalles, los recursos afectados y las recomendaciones. Además, también puedes ver el historial de los recursos fijos.

Pantalla 6: detalles de Insight. Le proporciona detalles, recomendaciones y un cronograma histórico.



Sección de métricas

Los gráficos interactivos de esta sección muestran las siguientes métricas del clúster:

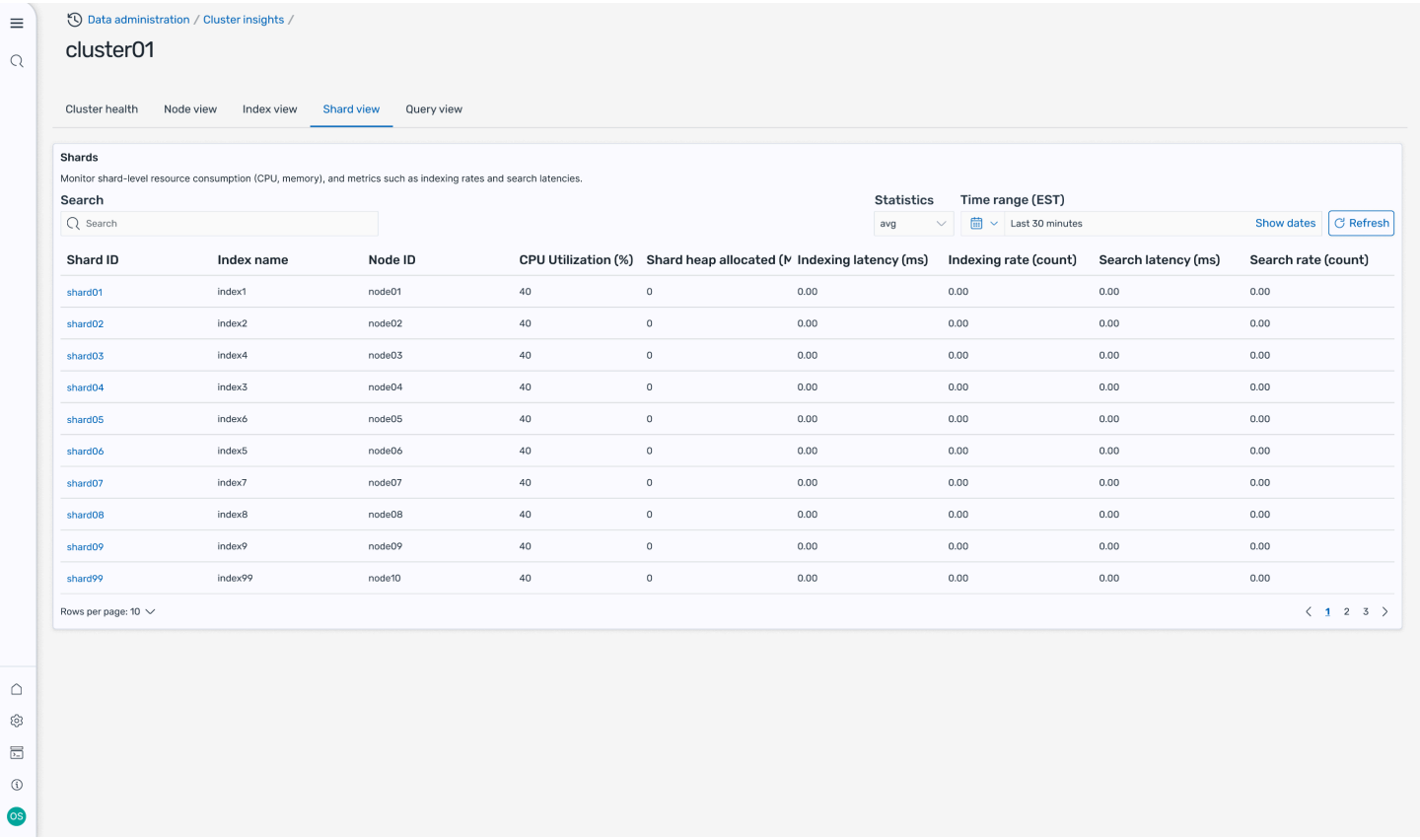
- Métricas generales del estado del clúster, como el estado del clúster, el estado de escritura y los documentos que permiten búsquedas
- KPIs (Indicadores clave de rendimiento), como las tasas y latencias de indexación y búsqueda
- Métricas de utilización de recursos, como la utilización de la JVM y la CPU

Vistas de nodos, índices y fragmentos

Las vistas de nodos, índices y fragmentos utilizan OpenSearch estadísticas para proporcionar una visibilidad detallada de las operaciones del clúster. En ella aparecen:

- Métricas en tiempo real, como el uso de la CPU y la presión de la memoria de la JVM
- Busque e indexe datos de rendimiento
- Puntos críticos de recursos en todos los componentes del clúster
- Diagnósticos granulares a nivel de nodo
- ¿Se ha asignado el montón de fragmentos superior?

Pantalla 7: métricas a nivel de nodo, índice y fragmento



Vista de consulta

> **Note**
>
> La función de visualización de consultas es compatible con las OpenSearch versiones 2.19 o posteriores.

La página de visualización de consultas le ayuda a supervisar las consultas que consumen muchos recursos con:

Configuración de acceso

La visualización de las N consultas principales requiere permisos de control de acceso detallados. Asegúrese de lo siguiente:

- El control de acceso detallado está habilitado en su dominio.

- Tu función de IAM (o usuario interno) está asignada a una OpenSearch función con los permisos de clúster necesarios para obtener información sobre las consultas.
- Para obtener acceso de administrador completo, asigne el ARN de su rol de IAM como rol de backend a los roles all_access y security_manager. Puede hacerlo en los OpenSearch paneles de control, en Seguridad > Roles > seleccione el rol > Usuarios mapeados > Administrar el mapeo, o bien mediante la API de seguridad (PUT _plugins/_ _access). security/api/rolesmapping/all

Sin una asignación de roles adecuada, los usuarios pueden recibir 403 errores prohibidos al intentar acceder a los datos de Query Insights. Para obtener más información, consulte Control de acceso detallado.

Paneles de control en vivo

Vea las estadísticas de ejecución, el uso de CPU y memoria y el progreso de finalización de cada consulta.

Consultas N principales

En una tabla de clasificación se muestran las consultas más importantes con detalles como:

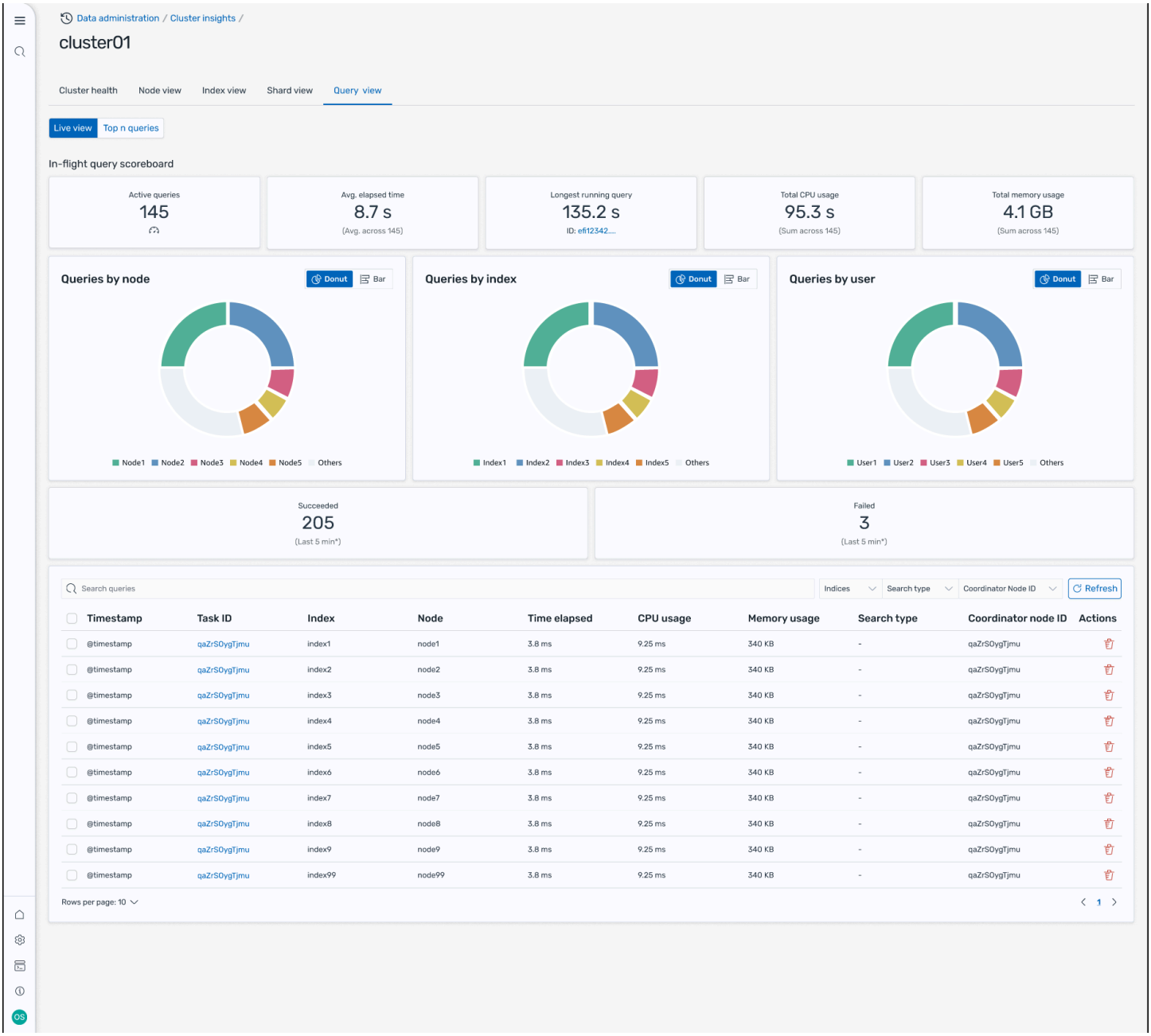
- Recuento de consultas
- Latencia, CPU y uso de memoria
- Tipo de búsqueda y nodo coordinador
- Índices objetivo y recuento de fragmentos

Detalles de consulta

Haga doble clic en cualquier consulta para ver:

- Carga útil y pasos de ejecución exactos de la consulta
- Desglose de la latencia de cada fase (expandir, consultar, recuperar)
- Recomendaciones de optimización

Pantalla-8: vista en directo durante el vuelo. También puedes ver las consultas de Top-N



Acceda a Insights a través de EventBridge los eventos de Amazon

Puedes monitorizar la información a través de EventBridge los eventos de Amazon. Para obtener más información, consulta las notificaciones.

Catálogo Insights

Cluster Insights ofrece la siguiente información útil sobre los clústeres afectados

Información	Description (Descripción)
Disco bajo	Se obtiene una información sobre la gravedad alta cuando el clúster FreeStorage cae por debajo del 25% de la capacidad total o de 25 GB, cualquiera que sea el umbral que se alcance primero.
La escritura en el clúster está bloqueada	Se obtiene una información sobre la gravedad crítica cuando el clúster FreeStorage cae por debajo del 20% de la capacidad total o de 20 GB, independientemente del umbral que se alcance primero.
Fragmento grande	Se genera una información de baja gravedad cuando el tamaño del fragmento está entre 50 GB y 100 GB. La gravedad aumenta a media cuando el tamaño del fragmento está entre 100 GB y 200 GB, y aumenta aún más a alta cuando el tamaño del fragmento supera los 200 GB.
Recuento de fragmentos	Se genera una información sobre la gravedad alta cuando el clúster supera los 25 fragmentos por GB de pila configurado o cuando el número de fragmentos por nodo alcanza o supera los 900. Para la OpenSearch versión 2.17 y versiones posteriores, esta información se generará en función de la configuración del administrador de clústeres, tal como se detalla en esta documentación.
Estado del clúster: rojo	Cuando la sustitución de un nodo degradado del clúster da como resultado un estado de clúster rojo al reemplazar un nodo degradado en el OpenSearch clúster, se obtiene una información sobre la gravedad crítica. Esto requiere una acción inmediata para restaurar el clúster a un estado verde.
Estado del clúster: amarillo	Se genera una información de gravedad media cuando la sustitución de un nodo degradado en el OpenSearch clúster da como resultado un estado de clúster amarillo.
Réplica mal configurada	Se genera una información de gravedad media cuando el clúster no tiene réplicas configuradas o cuando el número de réplicas supera el número total de nodos de datos.

Información	Description (Descripción)
No se puede acceder a la clave KMS	Cuando el OpenSearch clúster pierde el acceso a la clave KMS utilizada para el cifrado en reposo, se genera una información sobre la gravedad crítica. Esto impide que el clúster acceda a los datos cifrados y afecta a las operaciones del clúster, por lo que es necesario tomar medidas inmediatas para restablecer el acceso.
Configuración de índice mal configurada	Al configurar el índice «index.routing.location», se genera una información de gravedad media. require/include/exclude. <attribute>«existe en cualquier índice del OpenSearch clúster. Esta configuración puede impedir que las implementaciones en azul y verde se realicen correctamente durante las actualizaciones del clúster.
Configuración incorrecta del administrador de clústeres	Se genera una información sobre la gravedad crítica cuando el OpenSearch clúster tiene 2 nodos de datos sin nodos de administrador de clústeres dedicados o 2 nodos de administrador de clústeres dedicados. Estas configuraciones crean un riesgo de pérdida de quórum e inestabilidad del clúster.
Estrategia de fragmentación subóptima	(a) Se genera una información de gravedad media que indica que uno o más índices tienen un recuento de fragmentos inferior al 80% de los nodos de datos. Esto puede provocar una distribución de la carga de trabajo subóptima en todo el clúster. (b) Se genera una información de gravedad media que indica que uno o más índices tienen fragmentos con un 20% más de datos que el tamaño medio de los fragmentos por nodo. Esto puede provocar una distribución de la carga de trabajo subóptima en todo el clúster.

Información	Description (Descripción)
Sobrecarga de clústeres	(a) Un análisis exhaustivo de la sobrecarga del clúster indica que más del 70% de los nodos de datos del clúster están experimen tando un uso elevado de la CPU, la memoria, las E/S del disco, el rendimiento del disco o la utilización del disco. Esto indica que es posible que el clúster se esté acercando a los límites de capacidad . (b) Un análisis de la gravedad crítica indica que más del 70% de los nodos de datos utilizan muchos recursos en términos de CPU, memoria, E/S de disco, rendimiento del disco o utilización del disco, y se observan limitaciones o rechazos de las solicitudes. Esto requiere una acción inmediata para escalar el clúster.

Seguridad en Amazon OpenSearch Service

La seguridad en la nube AWS es la máxima prioridad. Como AWS cliente, usted se beneficia de una arquitectura de centro de datos y red diseñada para cumplir con los requisitos de las organizaciones más sensibles a la seguridad.

La seguridad es una responsabilidad compartida entre usted AWS y usted. El modelo de responsabilidad compartida la describe como seguridad de la nube y seguridad en la nube:

- Seguridad de la nube: AWS es responsable de proteger la infraestructura que ejecuta AWS los servicios en la AWS nube. AWS también le proporciona servicios que puede utilizar de forma segura. Auditores independientes prueban y verifican periódicamente la eficacia de nuestra seguridad en el marco de los programas de conformidad de AWS. Para obtener más información sobre los programas de conformidad que se aplican a Amazon OpenSearch Service, consulta AWS Servicios incluidos en el ámbito de aplicación por programa de conformidad.
- Seguridad en la nube: tu responsabilidad viene determinada por el AWS servicio que utilices. También es responsable de otros factores, incluida la confidencialidad de los datos, los requisitos de la empresa y la legislación y la normativa aplicables.

Esta documentación le ayuda a comprender cómo aplicar el modelo de responsabilidad compartida al utilizar el OpenSearch Servicio. Los siguientes temas muestran cómo configurar el OpenSearch Servicio para cumplir sus objetivos de seguridad y conformidad. También aprenderá a utilizar otros AWS servicios que le ayudan a supervisar y proteger los recursos OpenSearch del Servicio.

Temas

- Protección de datos en Amazon OpenSearch Service
- Identity and Access Management en Amazon OpenSearch Service
- Prevención de la sustitución confusa entre servicios
- Control de acceso detallado en Amazon Service OpenSearch
- Validación de conformidad para Amazon OpenSearch Service
- Resiliencia en Amazon OpenSearch Service
- Autenticación y autorización JWT para Amazon Service OpenSearch
- Seguridad de la infraestructura en Amazon OpenSearch Service
- Autenticación SAML para paneles OpenSearch

- IAM Identity Center Trusted Identity Propagation Support para OpenSearch
- Configuración de la autenticación de Amazon Cognito para paneles OpenSearch
- Uso de roles vinculados a servicios para Amazon Service OpenSearch

Protección de datos en Amazon OpenSearch Service

El AWS modelo de se aplica a protección de datos en Amazon OpenSearch Service. Como se describe en este modelo, AWS es responsable de proteger la infraestructura global en la que se ejecutan todos los Nube de AWS. Eres responsable de mantener el control sobre el contenido alojado en esta infraestructura. También eres responsable de las tareas de administración y configuración de seguridad para los Servicios de AWS que utiliza. Para obtener más información sobre la privacidad de los datos, consulte las Preguntas frecuentes sobre la privacidad de datos. Para obtener información sobre la protección de datos en Europa, consulte la publicación de blog sobre el Modelo de responsabilidad compartida de AWS y GDPR en el Blog de seguridad de AWS .

Con fines de protección de datos, le recomendamos que proteja Cuenta de AWS las credenciales y configure los usuarios individuales con AWS IAM Identity Center o AWS Identity and Access Management (IAM). De esta manera, solo se otorgan a cada usuario los permisos necesarios para cumplir sus obligaciones laborales. También recomendamos proteger sus datos de la siguiente manera:

- Utiliza la autenticación multifactor (MFA) en cada cuenta.
- Se utiliza SSL/TLS para comunicarse con AWS los recursos. Exigimos TLS 1.2 y recomendamos TLS 1.3.
- Configure la API y el registro de actividad de los usuarios con AWS CloudTrail. Para obtener información sobre el uso de CloudTrail senderos para capturar AWS actividades, consulte Cómo trabajar con CloudTrail senderos en la Guía del AWS CloudTrail usuario.
- Utilice soluciones de AWS cifrado, junto con todos los controles de seguridad predeterminados Servicios de AWS.
- Utiliza servicios de seguridad administrados avanzados, como Amazon Macie, que lo ayuden a detectar y proteger la información confidencial almacenada en Amazon S3.
- Si necesita módulos criptográficos validados por FIPS 140-3 para acceder a AWS través de una interfaz de línea de comandos o una API, utilice un punto final FIPS. Para obtener más información sobre los puntos de conexión de FIPS disponibles, consulte Estándar de procesamiento de la información federal (FIPS) 140-3.

Se recomienda encarecidamente no introducir nunca información confidencial o sensible, como por ejemplo, direcciones de correo electrónico de clientes, en etiquetas o campos de formato libre, tales como el campo Nombre. Esto incluye cuando trabaja con un OpenSearch Servicio u otro tipo de servicio Servicios de AWS mediante la consola, la API o. AWS CLI AWS SDKs Cualquier dato que introduzca en etiquetas o campos de formato libre utilizados para los nombres se pueden emplear para los registros de facturación o diagnóstico. Si proporciona una URL a un servidor externo, recomendamos encarecidamente que no incluya información de credenciales en la URL a fin de validar la solicitud para ese servidor.

Cifrado de datos en reposo para Amazon OpenSearch Service

OpenSearch Los dominios de servicio ofrecen el cifrado de los datos en reposo, una función de seguridad que ayuda a evitar el acceso no autorizado a los datos. La función utiliza AWS Key Management Service (AWS KMS) para almacenar y gestionar las claves de cifrado y el algoritmo del estándar de cifrado avanzado con claves de 256 bits (AES-256) para realizar el cifrado. Si está habilitada, la característica cifra los siguientes aspectos de un dominio:

- Todos los índices (incluidos los que están almacenados) UltraWarm
- OpenSearch registros
- Archivos de intercambio
- Todos los demás datos del directorio de la aplicación
- Instantáneas automatizadas

Los siguientes elementos no se cifran cuando habilita el cifrado de datos en reposo, pero puede realizar pasos adicionales para protegerlos:

- Instantáneas manuales: actualmente no puede utilizar AWS KMS claves para cifrar instantáneas manuales. No obstante, puede utilizar el cifrado del lado del servidor con claves administradas por S3 o claves de KMS para cifrar el bucket que utiliza como repositorio de instantáneas. Para obtener instrucciones, consulte the section called “Registrar un repositorio de instantáneas manuales”.
- Registros lentos y registros de errores: si publica registros y desea cifrarlos, puede cifrar su grupo de CloudWatch registros con la misma AWS KMS clave que el dominio del servicio. OpenSearch Para obtener más información, consulte Cifrar datos de registro en CloudWatch Logs utilizando AWS Key Management Service la Guía del usuario de Amazon CloudWatch Logs.

> **Note**
>
> No puedes habilitar el cifrado en reposo en un dominio existente si UltraWarm el almacenamiento en frío está habilitado en el dominio. Primero debes deshabilitar UltraWarm o almacenar en frío, habilitar el cifrado en reposo y, a continuación, volver a habilitar UltraWarm o almacenar en frío. Si desea conservar los índices UltraWarm o almacenarlos en frío, debe moverlos a un almacenamiento en caliente antes de inhabilitarlos UltraWarm o almacenarlos en frío.

OpenSearch El servicio solo admite claves KMS de cifrado simétrico, no asimétricas. Para conocer cómo crear claves simétricas, consulte Crear una clave de KMS en la Guía para desarrolladores de AWS Key Management Service .

Independientemente de si el cifrado en reposo está habilitado, todos los dominios cifran automáticamente los paquetes personalizados mediante AES-256 y claves administradas por el servicio. OpenSearch

Permisos

Para usar la consola de OpenSearch servicio para configurar el cifrado de los datos en reposo, debe tener permisos de lectura AWS KMS, como la siguiente política basada en la identidad:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:List*",
        "kms:Describe*"
      ],
      "Resource": "*"
    }
  ]
}
```

Si desea utilizar una clave distinta de la AWS clave propia, también debe tener permisos para crear concesiones para la clave. Estos permisos normalmente suelen adoptar la forma de una política con base en recursos que especifica al crear la clave.

Si quieres que tu clave sea exclusiva de OpenSearch Service, puedes añadir la ViaService condición kms: a esa política clave:

```
"Condition": {
  "StringEquals": {
    "kms:ViaService": "es.us-west-1.amazonaws.com"
  },
  "Bool": {
    "kms:GrantIsForAWSResource": "true"
  }
}
```

Para obtener más información, consulte Políticas de claves en AWS KMS en la Guía para desarrolladores de AWS Key Management Service .

Habilitar el cifrado de datos en reposo

El cifrado de los datos inactivos en dominios nuevos requiere Elasticsearch 5.1 OpenSearch o una versión posterior. Para habilitarlo en los dominios existentes, se requiere Elasticsearch OpenSearch 6.7 o una versión posterior.

Para habilitar el cifrado de los datos en reposo (consola)

1. Abre el dominio en la AWS consola y, a continuación, selecciona Acciones y Editar la configuración de seguridad.
2. En Cifrado, seleccione Habilitar el cifrado de datos en reposo.
3. Elige la AWS KMS clave que quieras usar y, a continuación, selecciona Guardar cambios.

También puede habilitar el cifrado a través de la API de configuración. La siguiente solicitud permite el cifrado de datos en reposo en un dominio existente:

```
{
   "ClusterConfig":{
      "EncryptionAtRestOptions":{
         "Enabled": true,
         "KmsKeyId":"arn:aws:kms:us-east-1:123456789012:alias/my-key"
```

```
        }
    }
}
```

Clave KMS desactivada o eliminada

Si deshabilitas o eliminas la clave que usaste para cifrar un dominio, el dominio deja de estar accesible. OpenSearch El servicio te envía una notificación informándote de que no puede acceder a la clave KMS. Vuelva a habilitar la clave inmediatamente para acceder a su dominio.

El equipo del OpenSearch Servicio no puede ayudarte a recuperar tus datos si se elimina la clave. AWS KMS elimina las claves solo después de un período de espera de al menos siete días. Si su clave está pendiente de eliminación, cancele la eliminación o tome una instantánea manual del dominio para evitar la pérdida de datos.

Deshabilitar el cifrado de datos en reposo

Después de configurar un dominio para cifrar los datos en reposo, no puede desactivar la configuración. En su lugar, puede realizar una instantánea manual del dominio existente, crear otro dominio, migrar los datos y eliminar el dominio antiguo.

Monitorear dominios que cifran los datos en reposo

Los dominios que cifran los datos en reposo tienen dos métricas adicionales: `KMSKeyError` y `KMSKeyInaccessible`. Estas métricas solo aparecen si el dominio se encuentra con un problema con la clave de cifrado. Para obtener una descripción completa de estas métricas, consulte the section called “Métricas de clúster”. Puede verlos mediante la consola de OpenSearch servicio o la consola de Amazon CloudWatch .

> **Tip**
>
> Cada métrica representa un problema importante para un dominio, por lo que te recomendamos que crees CloudWatch alarmas para ambos. Para obtener más información, consulte the section called “ CloudWatch Alarmas recomendadas”.

Otras consideraciones

- La rotación automática de claves preserva las propiedades de AWS KMS las claves, por lo que la rotación no afecta a la capacidad de acceder a OpenSearch los datos. Los dominios del

OpenSearch Servicio de Cifrado no admiten la rotación manual de claves, lo que implica crear una clave nueva y actualizar cualquier referencia a la clave anterior. Para obtener más información, consulta Rotar AWS KMS claves en la guía para AWS Key Management Service desarrolladores.

- Algunos tipos de instancias no admiten el cifrado de datos en reposo. Para obtener más información, consulte the section called “Tipos de instancias admitidas”.
- Los dominios que cifran los datos en reposo utilizan otro nombre de repositorio para sus instantáneas automatizadas. Para obtener más información, consulte the section called “Restaurar instantáneas”.
- Si bien recomendamos encarecidamente habilitar el cifrado en reposo, puede agregar sobrecarga de CPU adicional y unos pocos milisegundos de latencia. Sin embargo, la mayoría de los casos de uso no son sensibles a estas diferencias y la magnitud del impacto depende de la configuración del clúster, los clientes y el perfil de uso.

Node-to-node cifrado para Amazon OpenSearch Service

Node-to-node el cifrado proporciona una capa adicional de seguridad además de las funciones predeterminadas de Amazon OpenSearch Service.

Cada dominio OpenSearch de servicio, independientemente de si el dominio utiliza el acceso a la VPC, reside en su propia VPC dedicada. Esta arquitectura evita que los posibles atacantes intercepten el tráfico entre OpenSearch los nodos y mantiene el clúster seguro. Sin embargo, de forma predeterminada, el tráfico dentro de la VPC no está cifrado. Node-to-nodeel cifrado permite el cifrado TLS 1.2 para todas las comunicaciones dentro de la VPC.

Si envía datos al OpenSearch Servicio a través de HTTPS, el node-to-node cifrado ayuda a garantizar que sus datos permanezcan cifrados mientras OpenSearch los distribuye (y redistribuye) por todo el clúster. Si los datos llegan sin cifrar a través de HTTP, OpenSearch Service los cifra cuando llegan al clúster. Puedes exigir que todo el tráfico al dominio llegue a través de HTTPS mediante la consola o la API AWS CLI de configuración.

Node-to-node se requiere el cifrado si se habilita un control de acceso detallado.

Habilitar el cifrado node-to-node

Node-to-node El cifrado en dominios nuevos requiere cualquier versión de OpenSearch Elasticsearch 6.0 o posterior. Para habilitar el node-to-node cifrado en los dominios existentes se requiere cualquier versión de OpenSearch Elasticsearch 6.7 o posterior. Seleccione el dominio existente en la consola de AWS , Acciones y Editar la configuración de seguridad.

Como alternativa, puede usar la API de configuración AWS CLI or. Para obtener más información, consulta la referencia de AWS CLI comandos y la referencia de la API de OpenSearch servicio.

Deshabilitar el cifrado node-to-node

Después de configurar un dominio para que utilice el node-to-node cifrado, no podrá deshabilitar la configuración. En su lugar, puede realizar una instantánea manual del dominio cifrado, crear otro dominio, migrar los datos y eliminar el dominio antiguo.

Identity and Access Management en Amazon OpenSearch Service

Amazon OpenSearch Service ofrece varias formas de controlar el acceso a tus dominios. En este tema, se tratan los diversos tipos de política, la forma en que interactúan entre sí y cómo crear sus propias políticas personalizadas.

> **Important**
>
> La compatibilidad con VPC introduce algunas consideraciones adicionales en el control de acceso al OpenSearch servicio. Para obtener más información, consulte the section called "Acerca de las políticas de acceso en los dominios de VPC".

Tipos de políticas

OpenSearch El servicio admite tres tipos de políticas de acceso:

- the section called "Políticas basadas en recursos"
- the section called "Políticas basadas en identidades"
- the section called "Políticas basadas en IP"

Políticas basadas en recursos

Al crear un dominio, agrega una política basada en recursos, a veces denominada política de acceso al dominio. Estas políticas especifican qué acciones puede realizar una entidad principal en los subrecursos del dominio (con la excepción de la búsqueda entre clústeres). Los subrecursos incluyen OpenSearch índices y. APIs El elemento de la política JSON Principal (Entidad principal) en IAM especifica las cuentas, los usuarios o los roles que tienen permitido el acceso. El elemento de

la política JSON Resource (Recurso) especifica los subrecursos a los que pueden obtener acceso estas entidades principales.

Por ejemplo, la siguiente política basada en recursos concede a `test-user` acceso completo (`es:*`) a los subrecursos en `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/*"
    }
  ]
}
```

Dos consideraciones importantes se aplican a esta política:

- Estos privilegios se aplican únicamente a este dominio. A menos que cree políticas similares en otros dominios, `test-user` solo puede acceder a `test-domain`.
- El `/*` final en el elemento `Resource` es significativo e indica que las políticas basadas en recursos solo se aplican a los subrecursos del dominio, no al propio dominio. En las políticas basadas en recursos, la acción `es:*` es equivalente a `es:ESHttp*`.

  Por ejemplo, `test-user` puede realizar solicitudes frente a un índice (`GET https://search-test-domain.us-west-1.es.amazonaws.com/test-index`), pero no puede actualizar la configuración del dominio (`POST https://es.us-west-1.amazonaws.com/2021-01-01/opensearch/domain/test-domain/config`). Observe la diferencia entre los dos puntos de conexión. El acceso a la API de configuración requiere una política basada en identidad.

Puede especificar un nombre de índice parcial agregando un comodín. Este ejemplo identifica cualquier índice que empiecen con `commerce`:

```
arn:aws:es:us-west-1:987654321098:domain/test-domain/commerce*
```

En este caso, el comodín significa que `test-user` puede realizar solicitudes a índices dentro de `test-domain` que tengan nombres que comiencen con `commerce`.

Para restringir aún más `test-user`, puede aplicar la siguiente política:

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"
    }
  ]
}
```

Ahora `test-user` puede realizar solo una operación: búsquedas del índice `commerce-data`. Todos los demás índices del dominio son inaccesibles y, sin permisos para utilizar las acciones `es:ESHttpPut` o `es:ESHttpPost`, `test-user` no puede agregar ni modificar documentos.

A continuación, podría decidir configurar un rol para usuarios avanzados. Esta política da `power-user-role` acceso a los métodos HTTP GET y PUT para todos los elementos del URIs índice:

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/*"
    }
  ]
}
```

Si el dominio se encuentra en una VPC o utiliza un control de acceso detallado, puede utilizar una política de acceso al dominio abierto. De lo contrario, la política de acceso al dominio debe contener alguna restricción, ya sea por entidad principal o dirección IP.

Para obtener información sobre todas las acciones disponibles, consulte the section called "Referencia de los elementos de las políticas". Para obtener un control mucho más pormenorizado sobre sus datos, utilice una política de acceso al dominio abierto con control de acceso detallado.

Políticas basadas en identidades

A diferencia de las políticas basadas en recursos, que forman parte de cada dominio de OpenSearch servicio, las políticas basadas en la identidad se asocian a los usuarios o roles mediante el servicio AWS Identity and Access Management (IAM). Al igual que las políticas basadas en recursos, las políticas basadas en identidad especifican quién puede obtener acceso a un servicio, las acciones que puede realizar y, si corresponde, los recursos en los que se pueden realizar dichas acciones.

Aunque en realidad no tienen por qué serlo, las políticas basadas en identidad suelen ser más genéricas. Normalmente, solo determinan las acciones de la API de configuración que un usuario

puede realizar. Una vez implementadas estas políticas, puede usar políticas basadas en recursos (o un control de acceso detallado) en Service para ofrecer a los usuarios acceso a los índices y. OpenSearch OpenSearch APIs

> **Note**
>
> Los usuarios con la `AmazonOpenSearchServiceReadOnlyAccess` política AWS gestionada no pueden ver el estado del clúster en la consola. Para que puedan ver el estado del clúster (y otros OpenSearch datos), añada la `es:ESHttpGet` acción a una política de acceso y adjúntela a sus cuentas o funciones.

Dado que las políticas basadas en identidad se adjuntan a usuarios o roles (entidades principales), el JSON no especifica una entidad principal. La siguiente política concede acceso a las acciones que comienzan por `Describe` y `List`. Esta combinación de acciones proporciona acceso de solo lectura a las configuraciones de dominio, pero no a los datos almacenados en el dominio:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Action": [
        "es:Describe*",
        "es:List*"
      ],
      "Effect": "Allow",
      "Resource": "*"
    }
  ]
}
```

Un administrador puede tener acceso total al OpenSearch Servicio y a todos los datos almacenados en todos los dominios:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Action": [
        "es:*"
      ],
      "Effect": "Allow",
      "Resource": "*"
    }
  ]
}
```

Las políticas basadas en identidad permiten utilizar etiquetas para controlar el acceso a la API de configuración. La siguiente política, por ejemplo, permite a las entidades principales adjuntas ver y actualizar la configuración de un dominio si el dominio tiene la etiqueta `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"
        ]
      }
    }
  }]
}
```

También puedes usar etiquetas para controlar el acceso a la OpenSearch API. Las políticas basadas en etiquetas para la OpenSearch API solo se aplican a los métodos HTTP. Por ejemplo, la siguiente política permite a los directores adjuntos enviar solicitudes GET y PUT a la OpenSearch API si el dominio tiene la `environment:production` etiqueta:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Action": [
      "es:ESHttpGet",
      "es:ESHttpPut"
    ],
    "Effect": "Allow",
    "Resource": "*",
    "Condition": {
      "ForAnyValue:StringEquals": {
        "aws:ResourceTag/environment": [
          "production"
        ]
      }
    }
  }]
}
```

Para un control más detallado de la OpenSearch API, considera la posibilidad de utilizar un control de acceso más detallado.

> **Note**
>
> Después de añadir una o varias OpenSearch APIs políticas basadas en etiquetas, debe realizar una sola operación de etiquetado (como añadir, eliminar o modificar una etiqueta) para que los cambios surtan efecto en un dominio. Debe utilizar el software de servicio R20211203 o posterior para incluir las operaciones de la OpenSearch API en las políticas basadas en etiquetas.

OpenSearch El servicio admite las claves de condición `TagKeys` globales `RequestTag` y las claves de condición de la API de configuración, no de la API. OpenSearch Estas condiciones solo se aplican a las llamadas a la API que incluyen etiquetas dentro de la solicitud, como `CreateDomain`, `AddTags` y `RemoveTags`. La siguiente política permite a las entidades principales adjuntas crear dominios, pero solo si incluyen la etiqueta `team:it` en la solicitud:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": {
    "Effect": "Allow",
    "Action": [
      "es:CreateDomain",
      "es:AddTags"
    ],
    "Resource": "*",
    "Condition": {
      "StringEquals": {
        "aws:RequestTag/team": [
          "it"
        ]
      }
    }
  }
}
```

Para obtener más información sobre el uso de etiquetas para el control de acceso y las diferencias entre las políticas basadas en recursos y las políticas basadas en identidad, consulte Definición de permisos en función de los atributos con la autorización de ABAC en la Guía del usuario de IAM.

Políticas basadas en IP

Las políticas basadas en IP restringen el acceso a un dominio a una o más direcciones IP o bloques de CIDR. Técnicamente, las políticas basadas en IP no son un tipo de política distinto. En lugar de ello, son solo políticas basadas en recursos que especifican una entidad principal anónima e incluyen una Condición especial. Para obtener información, consulte Elementos de la política JSON de IAM: Condición en la Guía del usuario de IAM.

El principal atractivo de las políticas basadas en IP es que permiten realizar solicitudes sin firmar a un dominio de OpenSearch servicio, lo que permite utilizar clientes como curl y OpenSearch Dashboards o acceder al dominio a través de un servidor proxy. Para obtener más información, consulte the section called “Uso de un proxy para acceder al servicio desde los paneles OpenSearch”.

> **Note**
>
> Si ha habilitado acceso VPC para su dominio, no puede configurar una política basada en IP. En su lugar, puede utilizar `security groups` para controlar qué direcciones IP pueden tener acceso al dominio. Para obtener más información, consulte los temas siguientes:
>
> - the section called “Acerca de las políticas de acceso en los dominios de VPC”
> - Controle el tráfico a sus AWS recursos mediante grupos de seguridad en la Guía del usuario de Amazon VPC

La siguiente política concede a todas las solicitudes HTTP que se originan en el intervalo de IP especificado acceso 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/*"
    }
  ]
}
```

Si su dominio tiene un punto de conexión público y no utiliza un control de acceso detallado, recomendamos combinar direcciones IP y entidades principales de IAM. Esta política concede acceso HTTP `test-user` únicamente si la solicitud se origina en el intervalo IP especificado:

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/*"
  }]
}
```

Realizar y firmar solicitudes OpenSearch de servicio

Incluso si configuras una política de acceso completamente abierta y basada en recursos, todas las solicitudes a la API de configuración del OpenSearch servicio deben estar firmadas. Si sus políticas

especifican funciones o usuarios de IAM, las solicitudes que se envíen OpenSearch APIs también deben firmarse con la versión 4 de AWS Signature. El método de firma difiere en función de la API:

- Para realizar llamadas a la API OpenSearch de configuración del servicio, le recomendamos que utilice una de las AWS SDKs. Esto simplifica SDKs enormemente el proceso y puede ahorrarle una cantidad significativa de tiempo en comparación con la creación y firma de sus propias solicitudes. Los puntos de conexión de la API de configuración utilizan el siguiente formato:

  ```
  es.region.amazonaws.com/2021-01-01/
  ```

  Por ejemplo, la siguiente solicitud introduce un cambio de configuración en el dominio `movies`, pero es preciso identificarse (no recomendado):

  ```
  POST https://es.us-east-1.amazonaws.com/2021-01-01/opensearch/domain/movies/config
  {
    "ClusterConfig": {
      "InstanceType": "c5.xlarge.search"
    }
  }
  ```

  Si utilizas uno de ellos SDKs, como Boto 3, el SDK gestiona automáticamente la firma de las solicitudes:

  ```
  import boto3

  client = boto3.client(es)
  response = client.update_domain_config(
    DomainName='movies',
    ClusterConfig={
      'InstanceType': 'c5.xlarge.search'
    }
  )
  ```

  Para ver un código de muestra de Java, consulte the section called “Usando el AWS SDKs”.

- Para realizar llamadas al OpenSearch APIs, debes firmar tus propias solicitudes. OpenSearch APIs Utilizan el siguiente formato:

  ```
  domain-id.region.es.amazonaws.com
  ```

Por ejemplo, la siguiente solicitud busca en el índice `movies` de thor:

```
GET https://my-domain.us-east-1.es.amazonaws.com/movies/_search?q=thor
```

> **Note**
>
> El servicio ignora los parámetros transferidos a las URLs solicitudes HTTP POST firmadas con la versión 4 de Signature.

Cuando las políticas chocan

Surgen complejidades cuando las políticas están en desacuerdo o no hacen mención explícita a un usuario. Cómo funciona IAM en la Guía del usuario de IAM ofrece un breve resumen de la lógica de evaluación de políticas:

- De forma predeterminada, se deniegan todas las solicitudes.
- Un permiso explícito anula esta opción predeterminada.
- Una denegación explícita anula cualquier permiso concedido.

Por ejemplo, si una política basada en recursos le concede acceso a un subrecurso de dominio (un OpenSearch índice o una API), pero una política basada en la identidad le niega el acceso, se le deniega el acceso. Si una política basada en identidad concede acceso y una política basada en recursos no especifica si debe tener acceso o no, tiene permitido el acceso. Consulte la siguiente tabla de políticas que se entrecruzan para acceder a un resumen completo de resultados de subrecursos del dominio.

	Permitido en política basada en recursos	Denegado en política basada en recursos	Ni permitido ni denegado en política basada en recursos
Allowed in identity-based policy	Permitir	Denegar	Permitir

	Permitido en política basada en recursos	Denegado en política basada en recursos	Ni permitido ni denegado en política basada en recursos
Denied in identity-based policy	Denegar	Denegar	Denegar
Neither allowed nor denied in identity-based policy	Permitir	Denegar	Denegar

Referencia de los elementos de las políticas

OpenSearch El servicio es compatible con la mayoría de los elementos de política de la Referencia de elementos de política de IAM, con la excepción de. `NotPrincipal` La siguiente tabla muestra los elementos más comunes.

Elemento de política JSON	Resumen
`Version`	La versión actual del idioma de la política es `2012-10-17` . Todas las políticas de acceso deben especificar este valor.
`Effect`	Este elemento especifica si la instrucción permite o deniega el acceso a las acciones especificadas. Los valores válidos son `Allow` o `Deny`.
`Principal`	Este elemento especifica el rol Cuenta de AWS o el usuario de IAM al que se le permite o deniega el acceso a un recurso y puede adoptar varias formas: • AWS cuentas: `"Principal":{"AWS": ["123456789012"]}` o `"Principal":{"AWS": ["arn:aws:iam::123456789012:root"]}` • Usuarios de IAM: `"Principal":{"AWS": ["arn:aws:iam::123456789012:user/test-user"]}` • Roles de IAM: `"Principal":{"AWS": ["arn:aws:iam::123456789012:role/test-role"]}`

Elemento de política JSON	Resumen
	Important Especificar el comodín * habilita el acceso anónimo al dominio, lo que no se recomienda a menos que agregue una condición basada en IP, utilice la compatibilidad con la VPC o habilite un control de acceso detallado. Además, analice detenidamente las siguientes políticas para confirmar que no otorgan acceso amplio: • Políticas basadas en identidad asociadas a las entidades principales de AWS vinculadas (por ejemplo, los roles de IAM) • Políticas basadas en recursos asociadas a AWS los recursos asociados (por ejemplo, claves de AWS Key Management Service KMS)

Elemento de política JSON	Resumen
`Action`	OpenSearch El servicio usa `ESHttp*` acciones para los métodos OpenSearch HTTP. El resto de las acciones se aplican a la API de configuración. Determinadas acciones `es:` admiten permisos de nivel de recursos. Por ejemplo, puede otorgar a un usuario permisos para eliminar un dominio particular sin dar a dicho usuario permisos para eliminar cualquier dominio. Otras acciones se aplican solo al propio servicio. Por ejemplo, `es:ListDomainNames` no tiene sentido en el contexto de un dominio único y, por lo tanto, requiere un comodín. Para obtener una lista de todas las acciones disponibles y saber si se aplican a los subrecursos del dominio (`test-domain/*`), a la configuración del dominio (`test-domain`) o solo al servicio (*), consulte Acciones, recursos y claves de condición de Amazon OpenSearch Service en la Referencia de autorización de servicios Las políticas basadas en recursos difieren de los permisos en el nivel de recursos. Las políticas basadas en recursos son políticas JSON completas que se adjuntan a dominios. Los permisos de nivel de recurso permiten restringir acciones a dominios o subrecursos particulares. En la práctica, puede pensar en permisos de nivel de recursos como parte opcional de una política basada en recursos o identidades. Aunque los permisos de nivel de recurso para `es:CreateDomain` puedan parecer poco intuitivos (después de todo, ¿por qué darle a un usuario permisos para crear un dominio que ya existe?), la utilización de un comodín permite aplicar un esquema de nomenclatura sencillo para sus dominios, por ejemplo `"Resource": "arn:aws:es:us-west-1:987654321098:domain/my-team-name-*"` . Por supuesto, nada impide que incluya acciones junto a elementos de recursos menos restrictivos, como las siguientes:

Elemento de política JSON	Resumen

	JSON
	
	{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": [
	"es:ESHttpGet",
	"es:DescribeDomain"
	],
	"Resource": "*"
	}
	]
	}
	
	Para obtener más información acerca de las acciones y recursos de emparejamiento, consulte el elemento `Resource` en esta tabla.

Elemento de política JSON	Resumen
`Condition`	OpenSearch El servicio admite la mayoría de las condiciones que se describen en las claves de contexto de condiciones AWS globales de la Guía del usuario de IAM. Entre las excepciones más destacables se incluye la `aws:PrincipalTag` clave, que el OpenSearch servicio no admite. Al configurar una política basada en IP, debe especificar las direcciones IP o bloque de CIDR como condición, como en el ejemplo siguiente: `"Condition": { "IpAddress": { "aws:SourceIp": [ "192.0.2.0/32" ] } }` Como se indica en the section called “Políticas basadas en identidades” `aws:ResourceTag` `aws:RequestTag` , las claves de `aws:TagKeys` condición y las claves de condición se aplican tanto a la API de configura ción como a la OpenSearch APIs.

Elemento de política JSON	Resumen

`Resource`	OpenSearch El servicio utiliza `Resource` los elementos de tres formas básicas:
	
	• Para las acciones que se aplican al propio OpenSearch Servicio, por ejemplo`es:ListDomainNames` , o para permitir el acceso total, utilice la siguiente sintaxis:
	
	`"Resource": "*"`
	
	• Para las acciones que implican la configuración de un dominio, como `es:DescribeDomain` , puede utilizar la siguiente sintaxis:
	
	`"Resource": "arn:aws:es:` *region*`:`*aws-account-id*`:domain/`*domain-name*`"`
	
	• Para las acciones que se aplican a los subrecursos de un dominio, como `es:ESHttpGet` , puede utilizar la siguiente sintaxis:
	
	`"Resource": "arn:aws:es:` *region*`:`*aws-account-id*`:domain/`*domain-name*`/*"`
	
	No es necesario utilizar un comodín. OpenSearch El servicio le permite definir una política de acceso diferente para cada OpenSearch índice o API. Por ejemplo, podría limitar los permisos de un usuario al índice `test-index` :
	
	`"Resource": "arn:aws:es:` *region*`:`*aws-account-id*`:domain/`*domain-name*`/test-index"`
	
	En lugar de acceso completo a `test-index` , quizá prefiera limitar la política a solo la API de búsqueda:
	
	`"Resource": "arn:aws:es:` *region*`:`*aws-account-id*`:domain/`*domain-name*`/test-index/_search"`

Elemento de política JSON	Resumen
	Incluso puede controlar el acceso a documentos individuales: `"Resource": "arn:aws:es:  region:aws-account-id:domain/domain-name /test-index/test-type/1"` Básicamente, si OpenSearch expresa el subrecurso como un URI, puede controlar el acceso a él mediante una política de acceso. Para obtener un mayor control sobre los recursos a los que puede obtener acceso un usuario, consulte the section called “Control de acceso detallado”. Para conocer detalles sobre qué acciones admiten los permisos en el nivel de recursos, consulte el elemento `Action` en esta tabla.

Opciones avanzadas y consideraciones de la API

OpenSearch El servicio tiene varias opciones avanzadas, una de las cuales tiene implicaciones en el control de acceso:`rest.action.multi.allow_explicit_index`. En su configuración predeterminada como verdadero, permite a los usuarios omitir los permisos de subrecursos en determinadas circunstancias.

Por ejemplo, tomemos la siguiente política basada en recursos:

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/*"
    }
  ]
}
```

Esta política otorga acceso `test-user` completo a la API OpenSearch masiva `test-index` y a la misma. También permite solicitudes `GET` a `restricted-index`.

La siguiente solicitud de indexación, como cabría esperar, falla debido a un error de permisos:

```
PUT https://search-test-domain.us-west-1.es.amazonaws.com/restricted-index/movie/1
{
  "title": "Your Name",
  "director": "Makoto Shinkai",
  "year": "2016"
}
```

A diferencia de la API index, la API bulk permite crear, actualizar y eliminar muchos documentos en una única llamada. Sin embargo, con frecuencia especifica estas operaciones en el cuerpo de la solicitud, en lugar de la URL de solicitud. Dado que OpenSearch Service utiliza URLs para controlar el acceso a los subrecursos del dominio, `test-user` puede, de hecho, utilizar la API masiva para

realizar cambios en ella. `restricted-index` Aunque el usuario carezca de permisos POST en el índice, la siguiente solicitud tiene éxito:

```
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" }
```

En esta situación, la política de acceso no consigue cumplir su objetivo. Para evitar que los usuarios omitan este tipo de restricciones, puede cambiar `rest.action.multi.allow_explicit_index` a falso. Si este valor es falso, dejarán de funcionar todas las llamadas a bulk, mget y msearch APIs que especifiquen nombres de índice en el cuerpo de la solicitud. En otras palabras, las llamadas a `_bulk` ya no funcionan, pero las llamadas a `test-index/_bulk` sí. Este segundo punto de conexión contiene un nombre de índice, por lo que no es necesario especificar uno en el cuerpo de la solicitud.

OpenSearch Los paneles dependen en gran medida de mget y msearch, por lo que es poco probable que funcionen correctamente después de este cambio. Para una corrección parcial, puede dejar `rest.action.multi.allow_explicit_index` como true y denegar a determinados usuarios el acceso a una o más de estas opciones. APIs

Para obtener más información acerca de cómo cambiar esta configuración, consulte the section called “Configuración avanzada de clústeres”.

Del mismo modo, la siguiente política basada en recursos contiene dos problemas sutiles:

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/*"
    }
  ]
}
```

- A pesar de la denegación explícita, `test-user` puede seguir realizando llamadas como, por ejemplo, `GET https://search-test-domain.us-west-1.es.amazonaws.com/_all/_search` y `GET https://search-test-domain.us-west-1.es.amazonaws.com/*/_search` para acceder a los documentos en `restricted-index`.
- Dado que el elemento `Resource` hace referencia a `restricted-index/*`, `test-user` no tiene permisos para acceder directamente a los documentos del índice. El usuario, sin embargo, tiene permisos para eliminar todo el índice. Para evitar el acceso y la eliminación, la política debe especificar en su lugar `restricted-index*`.

En lugar de combinar permisos amplios y denegaciones delimitadas, el enfoque más seguro consiste en seguir el principio de privilegio mínimo y conceder solo los permisos requeridos para realizar una tarea. Para obtener más información sobre cómo controlar el acceso a índices u OpenSearch operaciones individuales, consulte. the section called “Control de acceso detallado”

Important

Especificar el comodín * permite el acceso anónimo al dominio. No se recomienda el uso del comodín. Además, analice detenidamente las siguientes políticas para confirmar que no otorgan acceso amplio:

- Políticas basadas en la identidad asociadas a las entidades AWS principales asociadas (por ejemplo, las funciones de IAM)
- Políticas basadas en recursos asociadas a los AWS recursos asociados (por ejemplo, claves de KMS) AWS Key Management Service

Configurar políticas de acceso

- Para obtener instrucciones sobre cómo crear o modificar políticas basadas en recursos e IP en OpenSearch Service, consulte. the section called “Configurar políticas de acceso”
- Para obtener instrucciones acerca de la creación o modificación de políticas basadas en identidad en IAM, consulte Definición de permisos de IAM personalizados con políticas administradas por el cliente en la Guía del usuario de IAM.

Políticas de muestra adicionales

Si bien este capítulo incluye muchos ejemplos de políticas, el control de AWS acceso es un tema complejo que se entiende mejor con ejemplos. Para más información, consulte Ejemplos de políticas de IAM basadas en identidad en la Guía del usuario de IAM.

Referencia de permisos OpenSearch de la API de Amazon Service

Cuando configura el control de acceso, escribe políticas de permiso que puede adjuntar a una identidad de IAM (políticas basadas en identidad). Para obtener más información detallada de referencia, consulte los siguientes temas en la Referencia de autorizaciones de servicio:

- Claves de condición, recursos y acciones del servicio. OpenSearch
- Claves de acciones, recursos y condiciones de OpenSearch Ingestion.

Esta referencia contiene información sobre las operaciones de la API que se pueden utilizar en una política de IAM. También incluye el AWS recurso para el que puede conceder los permisos y las claves de condición que puede incluir para un control de acceso detallado.

Las acciones se especifican en el campo `Action` de la política, el valor del recurso en el campo `Resource` de la política y las condiciones en el campo `Condition` de la política. Para especificar una acción para el OpenSearch Servicio, usa el `es:` prefijo seguido del nombre de la operación de la API (por ejemplo,). `es:CreateDomain` Para especificar una acción para la OpenSearch ingestión, usa el `osis:` prefijo seguido de la operación de API (por ejemplo,). `osis:CreatePipeline`

AWS políticas gestionadas para Amazon OpenSearch Service

Una política AWS gestionada es una política independiente creada y administrada por AWS. AWS Las políticas administradas están diseñadas para proporcionar permisos para muchos casos de uso comunes, de modo que pueda empezar a asignar permisos a usuarios, grupos y funciones.

Ten en cuenta que es posible que las políticas AWS administradas no otorguen permisos con privilegios mínimos para tus casos de uso específicos, ya que están disponibles para que los usen todos los AWS clientes. Se recomienda definir políticas administradas por el cliente específicas para sus casos de uso a fin de reducir aún más los permisos.

No puedes cambiar los permisos definidos en AWS las políticas administradas. Si AWS actualiza los permisos definidos en una política AWS administrada, la actualización afecta a todas las identidades principales (usuarios, grupos y roles) a las que está asociada la política. AWS es más probable que actualice una política AWS administrada cuando Servicio de AWS se lance una nueva o cuando estén disponibles nuevas operaciones de API para los servicios existentes.

Para obtener más información, consulte Políticas administradas por AWS en la Guía del usuario de IAM.

AmazonOpenSearchDirectQueryGlueCreateAccess

Concede a Amazon OpenSearch Service Direct Query Service acceso a `CreateDatabase`,`CreatePartition`,`CreateTable`, y `BatchCreatePartition` AWS Glue API.

Puede encontrar la AmazonOpenSearchDirectQueryGlueCreateAccesspolítica en la consola de IAM.

AmazonOpenSearchServiceFullAccess

Otorga acceso completo a las operaciones y los recursos de la API de configuración del OpenSearch servicio para un Cuenta de AWS.

Puede encontrar la AmazonOpenSearchServiceFullAccesspolítica en la consola de IAM.

AmazonOpenSearchServiceReadOnlyAccess

Otorga acceso de solo lectura a todos los recursos del OpenSearch Servicio para un. Cuenta de AWS

Puede encontrar la AmazonOpenSearchServiceReadOnlyAccesspolítica en la consola de IAM.

AmazonOpenSearchServiceRolePolicy

No puede asociar `AmazonOpenSearchServiceRolePolicy` a sus entidades IAM. Esta política está asociada a un rol vinculado al servicio que permite al OpenSearch Servicio acceder a los recursos de la cuenta. Para obtener más información, consulte the section called “Permisos”.

Puede encontrar la AmazonOpenSearchServiceRolePolicypolítica en la consola de IAM.

AmazonOpenSearchServiceCognitoAccess

Proporciona los permisos mínimos de Amazon Cognito necesarios para habilitar la autenticación de Cognito.

Puedes encontrar la AmazonOpenSearchServiceCognitoAccesspolítica en la consola de IAM.

AmazonOpenSearchIngestionServiceRolePolicy

No puede asociar `AmazonOpenSearchIngestionServiceRolePolicy` a sus entidades IAM. Esta política se adjunta a una función vinculada a un servicio que permite a OpenSearch Ingestion habilitar el acceso a la VPC para las canalizaciones de ingestión, crear etiquetas y publicar métricas relacionadas con la ingestión en su cuenta. CloudWatch Para obtener más información, consulte the section called “Cómo utilizar roles vinculados a servicios”.

Puedes encontrar la política en la consola de IAM. AmazonOpenSearchIngestionServiceRolePolicy

OpenSearchIngestionSelfManagedVpcePolicy

No puede asociar `OpenSearchIngestionSelfManagedVpcePolicy` a sus entidades IAM. Esta política se adjunta a un rol vinculado a un servicio que permite a OpenSearch Ingestion habilitar el acceso autogestionado a la VPC para los canales de ingestión, crear etiquetas y publicar métricas relacionadas con la ingestión en su cuenta. CloudWatch Para obtener más información, consulte the section called “Cómo utilizar roles vinculados a servicios”.

Puedes encontrar la política en la consola de IAM. OpenSearchIngestionSelfManagedVpcePolicy

AmazonOpenSearchIngestionFullAccess

Otorga acceso completo a las operaciones y los recursos de la API de OpenSearch ingestión para un. Cuenta de AWS

Puede encontrar la AmazonOpenSearchIngestionFullAccesspolítica en la consola de IAM.

AmazonOpenSearchIngestionReadOnlyAccess

Otorga acceso de solo lectura a todos los recursos de OpenSearch Ingestion durante un. Cuenta de AWS

Puede encontrar la AmazonOpenSearchIngestionReadOnlyAccesspolítica en la consola de IAM.

AmazonOpenSearchServerlessServiceRolePolicy

Proporciona los Amazon CloudWatch permisos mínimos necesarios para enviar datos de métricas OpenSearch sin servidor a. CloudWatch

Puede encontrar la AmazonOpenSearchServerlessServiceRolePolicypolítica en la consola de IAM.

OpenSearch Actualizaciones del servicio para las políticas AWS gestionadas

Consulte los detalles sobre las actualizaciones de las políticas AWS gestionadas del OpenSearch Servicio desde que este servicio comenzó a realizar un seguimiento de los cambios.

Cambio	Descripción	Fecha
Se actualizó el `AmazonOpenSearchIngestionServiceRolePolicy`	La actualización otorga permiso a OpenSearch Ingestion para modificar los puntos de enlace de VPC creados OpenSearch por para compartir canalizaciones entre sí. VPCs Para ver la política JSON, consulte la consola de IAM.	28 de agosto de 2025
Se actualizó el `AmazonOpenSearchServiceRolePolicy`	Se agregó la siguiente sentencia a la política. Cuando Amazon OpenSearch Service asume la función `AWSServiceRoleForAmazonOpenSearchService` vinculada al servicio, esta nueva declaración de la	31 de marzo de 2025

Cambio	Descripción	Fecha
	política permite OpenSearch actualizar el alcance de acceso de cualquier AWS IAM Identity Center aplicación que solo esté gestionada por. OpenSearch `{ "Effect": "Allow", "Action": "sso:PutApplicationAccessScope", "Resource": "arn:aws:sso::*:application/*/*", "Condition": { "StringEquals": { "aws:ResourceOrgID": "${aws:PrincipalOrgID}" } } }`	
Se actualizó `AmazonOpenSearchServerlessServiceRolePolicy`	Se agregó el Sid `AllowAOSSCloudwatchMetrics` a la política `AmazonOpenSearchServerlessServiceRolePolicy` . Un Sid es un ID de instrucción que actúa como identificador opcional para la instrucción de la política.	12 de julio de 2024

Cambio	Descripción	Fecha
Se ha agregado `OpenSearchIngestionSelfManagedVpcePolicy`	Una nueva política que permite a OpenSearch Ingestion habilitar el acceso autogestionado a las VPC para las canalizaciones de ingestión, crear etiquetas y publicar métricas relacionadas con la ingestión en su cuenta. CloudWatch Para ver la política JSON, consulte la consola de IAM.	12 de junio de 2024
Added`AmazonOpenSearchDirectQueryGlueCreateAccess`	Concede a Amazon OpenSearch Service Direct Query Service acceso a `CreateDatabase` `CreatePartition` ,`CreateTable` , y `BatchCreatePartition` AWS Glue API.	6 de mayo de 2024
Actualización de `AmazonOpenSearchServiceRolePolicy` y `AmazonElasticsearchServiceRolePolicy`	Se han añadido los permisos necesarios para que el rol vinculado al servicio pueda asignar y IPv6 desasignar direcciones. La política de Elasticsearch en desuso también se ha actualizado para garantizar la compatibilidad con versiones anteriores.	18 de octubre de 2023

Cambio	Descripción	Fecha
Se ha agregado `AmazonOpenSearchIngestionServiceRolePolicy`	Una nueva política que permite a OpenSearch Ingestion habilitar el acceso a la VPC para las canalizaciones de ingestión, crear etiquetas y publicar métricas relacionadas con la CloudWatch ingestión en su cuenta. Para ver la política JSON, consulte la consola de IAM.	26 de abril de 2023
Se ha agregado `AmazonOpenSearchIngestionFullAccess`	Una nueva política que otorga acceso total a las operaciones y los recursos de la API de OpenSearch Ingestion para un. Cuenta de AWS Para ver la política JSON, consulte la consola de IAM.	26 de abril de 2023
Se ha agregado `AmazonOpenSearchIngestionReadOnlyAccess`	Una nueva política que otorga acceso de solo lectura a todos los recursos de OpenSearch Ingestion para un. Cuenta de AWS Para ver la política JSON, consulte la consola de IAM.	26 de abril de 2023

Cambio	Descripción	Fecha
Se ha agregado `AmazonOpenSearchServerlessServiceRolePolicy`	Una nueva política que proporciona los permisos mínimos necesarios para enviar datos de métricas OpenSearch sin servidor a. Amazon CloudWatch Para ver la política JSON, consulte la consola de IAM.	29 de noviembre de 2022
Actualización de `AmazonOpenSearchServiceRolePolicy` y `AmazonElasticsearchServiceRolePolicy`	Se agregaron los permisos necesarios para que la función vinculada al servicio cree puntos de enlace de VPC OpenSearch gestionados por el servicio. Algunas acciones solo se pueden realizar cuando la solicitud contiene la etiqueta `OpenSearchManaged=true` . La política de Elasticsearch en desuso también se ha actualizado para garantizar la compatibilidad con versiones anteriores.	7 de noviembre de 2022

Cambio	Descripción	Fecha
Actualización de `AmazonOpenSearchServiceRolePolicy` y `AmazonElasticsearchServiceRolePolicy`	Se agregó soporte para la `PutMetricData` acción, que es necesario para publicar las métricas OpenSearch del clúster en Amazon CloudWatch. La política de Elasticsearch en desuso también se ha actualizado para garantizar la compatibilidad con versiones anteriores. Para ver la política JSON, consulte la consola de IAM.	12 de septiembre de 2022
Actualización de `AmazonOpenSearchServiceRolePolicy` y `AmazonElasticsearchServiceRolePolicy`	Se ha agregado compatibilidad con el tipo de recurso `acm`. La política proporciona el permiso mínimo AWS Certificate Manager (ACM) de solo lectura necesario para que el rol vinculado al servicio verifique y valide los recursos de ACM con el fin de crear y actualizar dominios personalizados habilitados para terminales. La política de Elasticsearch en desuso también se ha actualizado para garantizar la compatibilidad con versiones anteriores.	28 de julio de 2022

Cambio	Descripción	Fecha
Actualización de `AmazonOpenSearchServiceCognitoAccess` y `AmazonESCognitoAccess`	Se agregó soporte para la `UpdateUserPoolClient` acción, que es necesario para establecer la configuración del grupo de usuarios de Cognito durante la actualización de Elasticsearch a. OpenSearch Se han corregido los permisos de la acción `SetIdentityPoolRoles` para permitir el acceso a todos los recursos. La política de Elasticsearch en desuso también se ha actualizado para garantizar la compatibilidad con versiones anteriores.	20 de diciembre de 2021
Se ha actualizado `AmazonOpenSearchServiceRolePolicy`	Se ha agregado compatibilidad con el tipo de recurso `security-group` . La política proporciona los permisos mínimos de Amazon EC2 y Elastic Load Balancing necesarios para el rol vinculado a servicios a fin de habilitar el acceso mediante la VPC.	9 de septiembre de 2021

Cambio	Descripción	Fecha
• Se ha agregado `AmazonOpenSearchServiceFullAccess` • Se ha dado de baja `AmazonESFullAccess`	Esta nueva política tiene por objeto sustituir a la política anterior. Ambas políticas proporcionan acceso completo a la API de configuración del OpenSearch servicio y a todos los métodos HTTP del. OpenSearch APIs El control de acceso detallado y las políticas basadas en recursos aún pueden restringir el acceso.	7 de septiembre de 2021
• Se ha agregado `AmazonOpenSearchServiceReadOnlyAccess` • Se ha dado de baja `AmazonESReadOnlyAccess`	Esta nueva política tiene por objeto sustituir a la política anterior. Ambas políticas proporcionan acceso de solo lectura a la API de configuración del OpenSearch servicio (`es:Describe* es:List*`, `yes:Get*`) y no proporcionan acceso a los métodos HTTP para la. OpenSearch APIs	7 de septiembre de 2021
• Se ha agregado `AmazonOpenSearchServiceCognitoAccess` • Se ha dado de baja `AmazonESCognitoAccess`	Esta nueva política tiene por objeto sustituir a la política anterior. Ambas políticas proporcionan los permisos mínimos de Amazon Cognito necesarios para habilitar la autenticación de Cognito.	7 de septiembre de 2021

Cambio	Descripción	Fecha
• Se ha agregado AmazonOpenSearchServiceRolePolicy • Se ha dado de baja `AmazonElasticsearchServiceRolePolicy`	Esta nueva política tiene por objeto sustituir a la política anterior. Ambas políticas proporcionan los permisos mínimos de Amazon EC2 y Elastic Load Balancing necesarios para el rol vinculado a servicios a fin de habilitar el acceso mediante la VPC.	7 de septiembre de 2021
Comenzó el seguimiento de los cambios	Amazon OpenSearch Service ahora rastrea los cambios en las políticas AWS gestionadas.	7 de septiembre de 2021

Prevención de la sustitución confusa entre servicios

El problema de la sustitución confusa es un problema de seguridad en el que una entidad que no tiene permiso para realizar una acción puede obligar a una entidad con más privilegios a realizar la acción. En AWS, la suplantación de identidad entre servicios puede provocar el confuso problema de un diputado. La suplantación entre servicios puede producirse cuando un servicio (el servicio que lleva a cabo las llamadas) llama a otro servicio (el servicio al que se llama). El servicio que lleva a cabo las llamadas se puedes manipular para utilizar sus permisos a fin de actuar en función de los recursos de otro cliente de una manera en la que no debe tener permiso para acceder. Para evitarlo, AWS proporciona herramientas que lo ayudan a proteger sus datos para todos los servicios con entidades principales de servicio a las que se les ha dado acceso a los recursos de su cuenta.

Recomendamos utilizar las claves de contexto de condición `aws:SourceAccount`global `aws:SourceArn`y las claves de contexto en las políticas de recursos para limitar los permisos que Amazon OpenSearch Service concede a otro servicio al recurso. Si el valor de `aws:SourceArn` no contiene el ID de cuenta, como un ARN de bucket de Amazon S3, debe utilizar ambas claves de contexto de condición global para limitar los permisos. Si utiliza claves de contexto de condición global y el valor de `aws:SourceArn` contiene el ID de cuenta, el valor de `aws:SourceAccount` y

la cuenta en el valor de `aws:SourceArn` deben utilizar el mismo ID de cuenta cuando se utiliza en la misma instrucción de política. Utiliza `aws:SourceArn` si desea que solo se asocie un recurso al acceso entre servicios. Utiliza `aws:SourceAccount` si quiere permitir que cualquier recurso de esa cuenta se asocie al uso entre servicios.

El valor de `aws:SourceArn` debe ser el ARN del dominio del OpenSearch servicio.

La forma más eficaz de protegerse contra el problema de la sustitución confusa es utilizar la clave de contexto de condición global de `aws:SourceArn` con el ARN completo del recurso. Si no conoce el ARN completo del recurso o si especifica varios recursos, utiliza la clave de condición de contexto global `aws:SourceArn` con comodines (*) para las partes desconocidas del ARN. Por ejemplo, `arn:aws:es:*:123456789012:*`.

El siguiente ejemplo muestra cómo se pueden utilizar las claves de contexto de condición `aws:SourceAccount` global `aws:SourceArn` y las claves de contexto de OpenSearch Service para evitar el confuso problema de los adjuntos.

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"
         }
      }
   }
}
```

Control de acceso detallado en Amazon Service OpenSearch

El control de acceso detallado ofrece formas adicionales de controlar el acceso a tus datos en Amazon Service. OpenSearch Por ejemplo, puede ser conveniente que, según quién realice la solicitud, una búsqueda devuelva los resultados de un solo índice. Es posible que desee ocultar determinados campos en los documentos o excluir algunos documentos por completo.

El control de acceso detallado ofrece las siguientes características:

- Control de acceso con base en roles
- Seguridad en el nivel de índice, de documento y de campo
- OpenSearch Paneles de control multiusuario
- Autenticación básica HTTP para paneles y paneles OpenSearch OpenSearch

Temas

- Un panorama más amplio: control de acceso detallado y seguridad del servicio OpenSearch
- Conceptos clave
- Acerca del usuario maestro
- Habilitar el control de acceso detallado
- Acceder a los OpenSearch paneles de control como usuario maestro
- Administrar permisos
- Configuraciones recomendadas
- Limitaciones
- Modificar el usuario maestro
- Usuarios maestros adicionales
- Instantáneas manuales
- Integraciones
- Diferencias en la API de REST
- Tutorial: Configurar un dominio con un usuario maestro de IAM y la autenticación de Amazon Cognito
- Tutorial: Configurar un dominio con la base de datos de usuarios interna y la autenticación básica de HTTP

Un panorama más amplio: control de acceso detallado y seguridad del servicio OpenSearch

La seguridad OpenSearch de Amazon Service consta de tres capas principales:

Red

La primera capa de seguridad es la red, que determina si las solicitudes llegan a un dominio OpenSearch de servicio. Si elige Acceso público al crear un dominio, las solicitudes procedentes de cualquier cliente conectado a Internet pueden llegar al punto de conexión del dominio. Si elige Acceso a la VPC, los clientes deben conectarse a la VPC (y los grupos de seguridad asociados deben permitirlo) para que una solicitud llegue al punto de conexión. Para obtener más información, consulte the section called “Compatibilidad con VPC”.

Política de acceso al dominio

La segunda capa de seguridad es la política de acceso al dominio. Una vez que una solicitud llega a un punto de conexión del dominio, la política de acceso con base en recursos permite o deniega el acceso de la solicitud a un URI determinado. La política de acceso acepta o rechaza las solicitudes en el «límite» del dominio, antes de que lleguen a OpenSearch propiamente dicho.

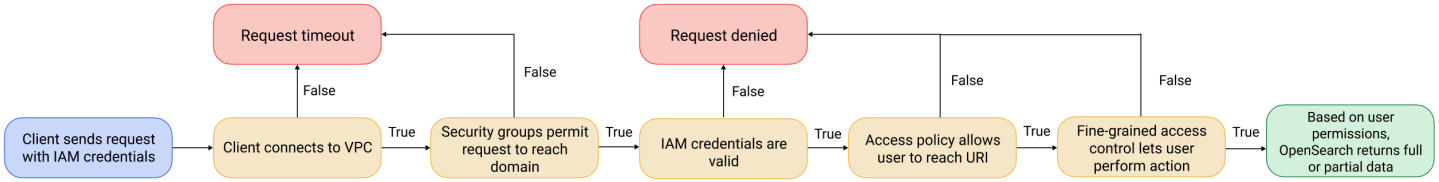
Control de acceso detallado

La tercera y última capa de seguridad es el control de acceso detallado. Una vez que una política de acceso con base en recursos ha permitido que una solicitud llegue a un punto de conexión del dominio, el control de acceso detallado evalúa las credenciales del usuario y lo autentica o bien deniega la solicitud. Si el control de acceso detallado autentica al usuario, obtiene todos los roles mapeados a ese usuario y utiliza el conjunto completo de permisos para determinar cómo gestionar la solicitud.

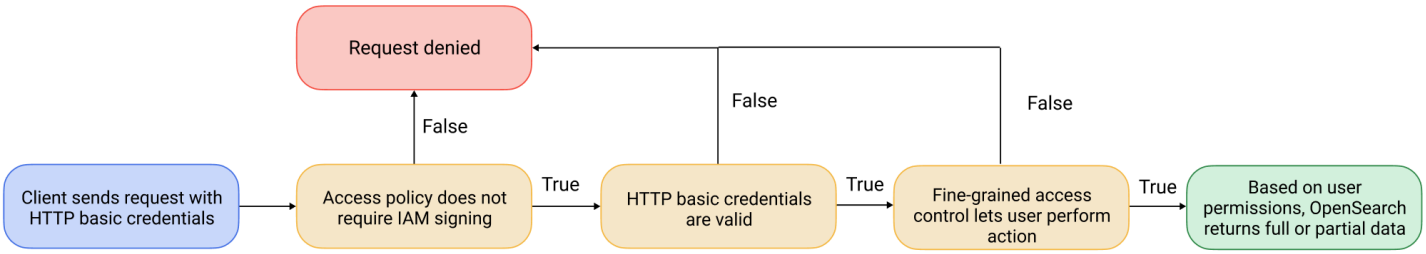
> **Note**
>
> Si una política de acceso basada en recursos contiene funciones o usuarios de IAM, los clientes deben enviar las solicitudes firmadas mediante la versión 4 de AWS Signature. Como tales, las políticas de acceso pueden entrar en conflicto con el control de acceso detallado, sobre todo si se utiliza la base de datos de usuarios interna y la autenticación básica de HTTP. No se puede firmar una solicitud con un nombre de usuario y contraseña y

también con credenciales de IAM. En general, si habilita el control de acceso detallado, se recomienda utilizar una política de acceso al dominio que no requiera solicitudes firmadas.

Este primer diagrama ilustra una configuración habitual: un dominio de acceso a la VPC con el control de acceso detallado habilitado, una política de acceso con base en IAM y un usuario maestro de IAM.



Este segundo diagrama ilustra otra configuración habitual: un dominio de acceso público con el control de acceso detallado habilitado, una política de acceso que no utiliza principales de IAM y un usuario maestro en la base de datos de usuarios interna.



Ejemplo

Supongamos que se realiza una solicitud `GET` a `movies/_search?q=thor`. ¿El usuario tiene permisos para buscar en el índice `movies`? En caso afirmativo, ¿el usuario tiene permisos para ver todos los documentos que contiene? ¿La respuesta debería omitir o anonimizar algún campo? Para el usuario maestro, la respuesta podría tener este aspecto:

```
{
  "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
        }
      },
      ...
    ]
  }
}
```

Si un usuario con permisos más limitados emite exactamente la misma solicitud, la respuesta podría tener el siguiente aspecto:

```
{
  "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 respuesta contiene menos aciertos y menos campos para cada acierto. Además, el campo `release_date` está anonimizado. Si un usuario sin permisos realiza la misma solicitud, el clúster devuelve un error:

```
{
  "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
}
```

Si un usuario proporciona credenciales no válidas, el clúster devuelve una excepción `Unauthorized`.

Conceptos clave

Al comenzar con el control de acceso detallado, tenga en cuenta los siguientes conceptos:

- Roles: la forma principal de utilizar el control de acceso detallado. En este caso, se trata de roles distintos de los roles de IAM. Los roles contienen cualquier combinación de permisos: para todo el clúster, específicos del índice, en el nivel de documento y en el nivel de campo.
- Asignación: después de configurar un rol, se asigna a uno o más usuarios. Por ejemplo, puede mapear tres roles a un solo usuario: un rol que proporciona acceso a Dashboards, otro que

proporciona acceso de solo lectura a `index1` y otro que proporciona acceso de escritura a `index2`. Si lo prefiere, puede incluir todos esos permisos en un solo rol.

- Usuarios: personas o aplicaciones que realizan solicitudes al clúster. OpenSearch Los usuarios tienen credenciales, ya sean claves de acceso de IAM o un nombre de usuario y una contraseña, que especifican al realizar las solicitudes.

Acerca del usuario maestro

El usuario maestro de OpenSearch Service es una combinación de nombre de usuario y contraseña, o un usuario principal de IAM, que tiene todos los permisos para acceder al OpenSearch clúster subyacente. Se considera que un usuario es un usuario maestro si tiene todos los accesos al OpenSearch clúster y la capacidad de crear usuarios internos, funciones y asignaciones de funciones en los paneles de control. OpenSearch

Un usuario maestro creado en la consola OpenSearch de servicio o mediante la CLI se asigna automáticamente a dos funciones predefinidas:

- `all_access`: proporciona acceso total a todas las operaciones del clúster, permiso para escribir en todos los índices del clúster y permiso para escribir en todos los inquilinos.
- `security_manager`: proporciona acceso al complemento de seguridad y administración de los usuarios y los permisos.

Con estas dos funciones, el usuario obtiene acceso a la pestaña Seguridad de los OpenSearch paneles, donde puede administrar los usuarios y los permisos. Si crea otro usuario interno y solo lo asigna al rol `all_access`, el usuario no tendrá acceso a la pestaña Seguridad. Para crear usuarios maestros adicionales, puede asignarlos de forma explícita a los roles `all_access` y `security_manager`. Para obtener instrucciones, consulte the section called “Usuarios maestros adicionales”.

Al crear un usuario maestro para su dominio, puede especificar una entidad principal de IAM existente o crear un usuario maestro en la base de datos de usuarios interna. Tenga en cuenta lo siguiente cuando tenga que tomar una decisión sobre el uso:

- Principal de IAM: si elige un principal de IAM para su usuario maestro, todas las solicitudes al clúster deben firmarse con la versión 4 de AWS Signature.

OpenSearch El servicio no tiene en cuenta ninguno de los permisos del director de IAM. El rol o usuario de IAM sirve únicamente para la autenticación. Las políticas de ese usuario o rol no influyen en la autorización del usuario maestro. La autorización se gestiona a través de los distintos permisos del complemento OpenSearch de seguridad.

Por ejemplo, no puede asignar ningún permiso de IAM a un responsable de IAM y, siempre que la máquina o la persona puedan autenticarse ante ese usuario o función, tendrán el poder de usuario maestro de Service. OpenSearch

Recomendamos IAM si quiere usar los mismos usuarios en varios clústeres, si quiere usar Amazon Cognito para acceder a los paneles o si OpenSearch tiene clientes que admiten la firma de Signature Version 4.

- Base de datos de usuarios interna: si crea una base de datos maestra en la base de datos de usuarios interna (con una combinación de nombre de usuario y contraseña), puede utilizar la autenticación básica HTTP (además de las credenciales de IAM) para hacer solicitudes al clúster. La mayoría de los clientes admiten la autenticación básica, incluida la función curl, que también admite la versión 4 de AWS Signature con la opción --aws-sigv4. La base de datos interna de usuarios se almacena en un OpenSearch índice, por lo que no puede compartirla con otros clústeres.

  Recomendamos la base de datos de usuarios interna si no se requiere reutilizar los usuarios en varios clústeres, si se desea utilizar la autenticación básica de HTTP para obtener acceso a Dashboards (en lugar de Amazon Cognito) o si se tienen clientes que solo admiten la autenticación básica. La base de datos de usuarios interna es la forma más sencilla de empezar a usar OpenSearch Service.

Habilitar el control de acceso detallado

Habilite un control de acceso detallado mediante la consola o la API de AWS CLI configuración. Para ver los pasos, consulte Creación y administración de dominios.

El control de acceso detallado requiere OpenSearch Elasticsearch 6.7 o una versión posterior. También requiere HTTPS para todo el tráfico al dominio, el cifrado de los datos en reposo y el cifrado. node-to-node Según cómo configure las características avanzadas del control de acceso detallado, el procesamiento adicional de sus solicitudes puede requerir recursos de cómputo y memoria en nodos de datos individuales. Después de habilitar el control de acceso detallado, no puede desactivarlo.

Habilitar el control de acceso detallado en los dominios existentes

Puedes habilitar un control de acceso detallado en los dominios existentes que ejecuten Elasticsearch 6.7 OpenSearch o una versión posterior.

Para habilitar el control de acceso detallado en un dominio existente (consola)

1. Seleccione el dominio y elija Acciones y Editar la configuración de seguridad.
2. Seleccione Habilitar el control de acceso detallado,
3. Seleccione cómo crear el usuario maestro:
   - Si desea utilizar IAM para la gestión de usuarios, seleccione Establecer ARN de IAM como usuario maestro y especifique el ARN para un rol de IAM.
   - Si desea utilizar la base de datos de usuarios interna, seleccione Crear usuario maestro y especifique un nombre de usuario y una contraseña.
4. (Opcional) Seleccione Habilitar el período de migración para la política de acceso abierta o basada en IP. Esta configuración permite un período de transición de 30 días durante el cual los usuarios existentes pueden seguir accediendo al dominio sin interrupciones, y las políticas de acceso basadas en IP y abiertas existentes seguirán funcionando con su dominio. Durante este período de migración, recomendamos a los administradores crear los roles necesarios y asignarlos a los usuarios para el dominio. Si utiliza políticas basadas en la identidad en lugar de una política de acceso abierta o basada en la IP, puede desactivar esta configuración.

   También debe actualizar sus clientes para que trabajen con un control de acceso de grano fino durante el periodo de migración. Por ejemplo, si asignas las funciones de IAM a un control de acceso detallado, debes actualizar tus clientes para empezar a firmar las solicitudes con la versión 4 de Signature. AWS Si configura la autenticación básica HTTP con un control de acceso detallado, debe actualizar los clientes para proporcionar las credenciales de autenticación básicas adecuadas en las solicitudes.

   Durante el período de migración, los usuarios que accedan al punto final de OpenSearch Dashboards para el dominio accederán directamente a la página Discover en lugar de a la página de inicio de sesión. Los administradores y los usuarios maestros pueden elegir Inicio de sesión para iniciar sesión con credenciales de administrador y configurar asignaciones de roles.

> **Important**
>
> OpenSearch El servicio desactiva automáticamente el período de migración después de 30 días. Recomendamos finalizarlo tan pronto como cree los roles necesarios y los asigne a los usuarios. Una vez finalizado el período de migración, no puedes volver a habilitarlo.

5. Seleccione Guardar cambios.

El cambio desencadena una implementación azul-verde durante el cual el estado del clúster se vuelve rojo, pero todas las operaciones de clúster no se ven afectadas.

Para habilitar el control de acceso detallado en un dominio (CLI) existente

Establezca `AnonymousAuthEnabled` a `true` para habilitar el período de migración con un control de acceso detallado:

```
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}'
```

Acerca del default_role

El control de acceso detallado requiere asignación de roles. Si tu dominio usa políticas de acceso basadas en la identidad, OpenSearch Service asigna automáticamente a tus usuarios a un nuevo rol llamado default_role para ayudarte a migrar correctamente a los usuarios existentes. Esta asignación temporal garantiza que los usuarios puedan seguir enviando correctamente solicitudes GET y PUT firmadas por IAM hasta que cree sus propias asignaciones de roles.

La función no añade ninguna vulnerabilidad o fallo de seguridad a tu dominio de servicio. OpenSearch Recomendamos eliminar el rol predeterminado tan pronto como configure sus propios roles y los asigne en consecuencia.

Escenarios de migración

La siguiente tabla describe el comportamiento de cada método de autenticación antes y después de habilitar el control de acceso detallado en un dominio existente, y los pasos que deben seguir los administradores para asignar correctamente sus usuarios a los roles:

Método de autenticación	Antes de habilitar el control de acceso detallado	Después de habilitar el control de acceso detallado	Tareas administrativas
Políticas basadas en identidades	Todos los usuarios que cumplan la política de IAM pueden acceder al dominio.	No es necesario habilitar el período de migración. OpenSearch El servicio asigna automáticamente a todos los usuarios que cumplen la política de IAM al default_role para que puedan seguir accediendo al dominio.	1. Cree asignaciones de roles personalizadas en el dominio. 2. Elimine las default_role.
Políticas basadas en IP	Todos los usuarios de las direcciones IP permitidas o los bloques de CIDR pueden acceder al dominio.	Durante el período de migración de 30 días, todos los usuarios de las direccion es IP permitida s o los bloques de CIDR pueden seguir accediendo al dominio.	1. Cree asignaciones de roles personalizadas en el dominio. 2. Actualice sus clientes para proporcionar credenciales de autenticación básicas o credencia les de IAM, según la configuración de asignación de roles. 3. Desactive el período de migración. Los usuarios de las direcciones IP permitidas o los bloques de CIDR que envían solicitudes sin autentica ción básica o credenciales de IAM perderán el acceso al dominio.
Políticas de acceso abierto	Todos los usuarios en Internet	Durante el periodo de migración de 30 días, todos los usuarios	1. Cree asignaciones de roles en el dominio.

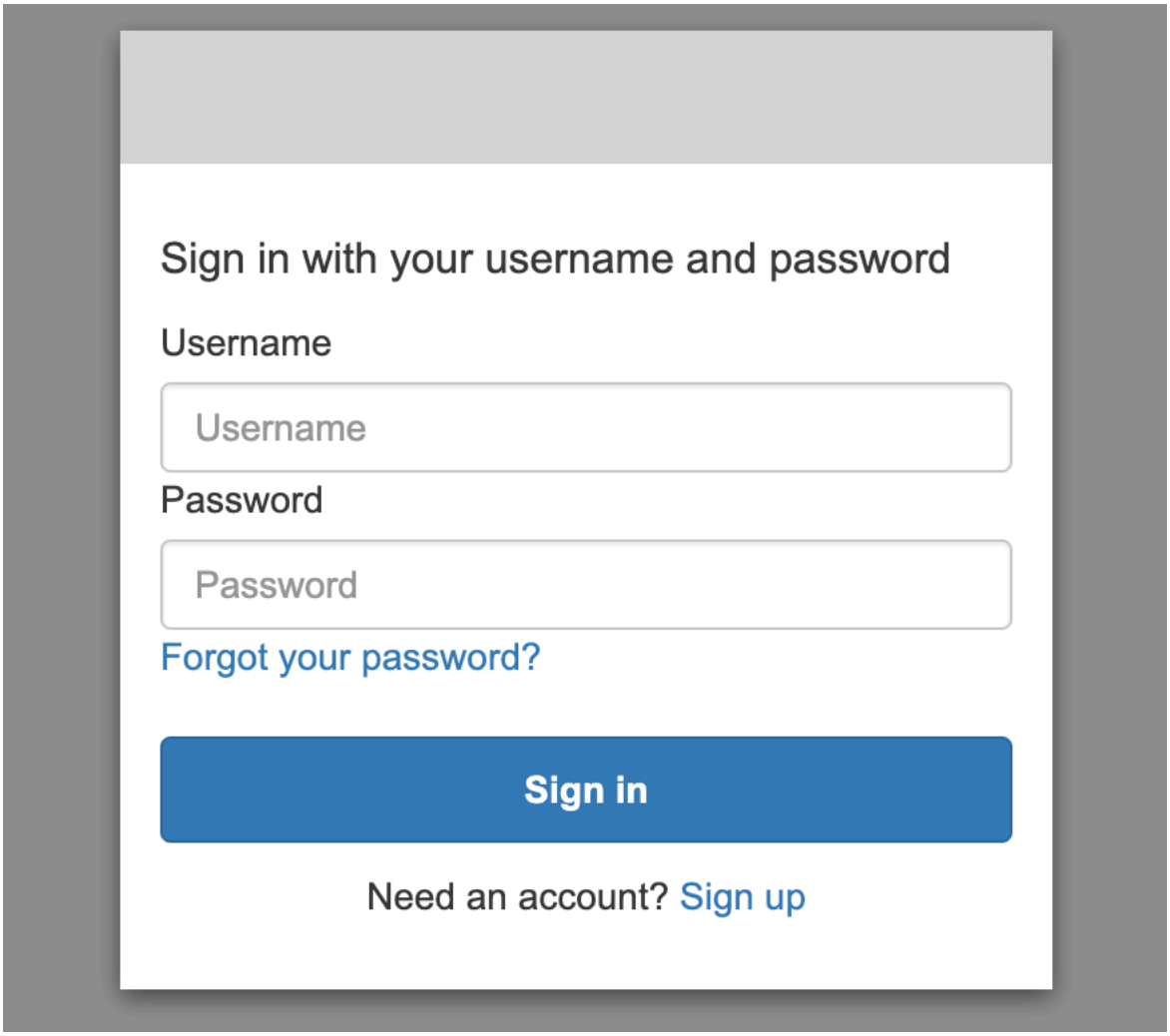
Método de autenticación	Antes de habilitar el control de acceso detallado	Después de habilitar el control de acceso detallado	Tareas administrativas
	pueden acceder al dominio.	de Internet pueden seguir accediendo al dominio.	2. Actualice sus clientes para proporcionar credenciales de autenticación básicas o credencia les de IAM, según la configuración de asignación de roles. 3. Desactive el período de migración. Los usuarios que envíen solicitudes sin autenticación básica o credencia les de IAM perderán el acceso al dominio.

Acceder a los OpenSearch paneles de control como usuario maestro

El control de acceso detallado tiene un complemento de OpenSearch paneles de control que simplifica las tareas de administración. Puede utilizar Dashboards para administrar usuarios, roles, mapeos, grupos de acciones e inquilinos. Sin embargo, la página de inicio de sesión de OpenSearch Dashboards y el método de autenticación subyacente varían según la forma en que administre los usuarios y configure su dominio.

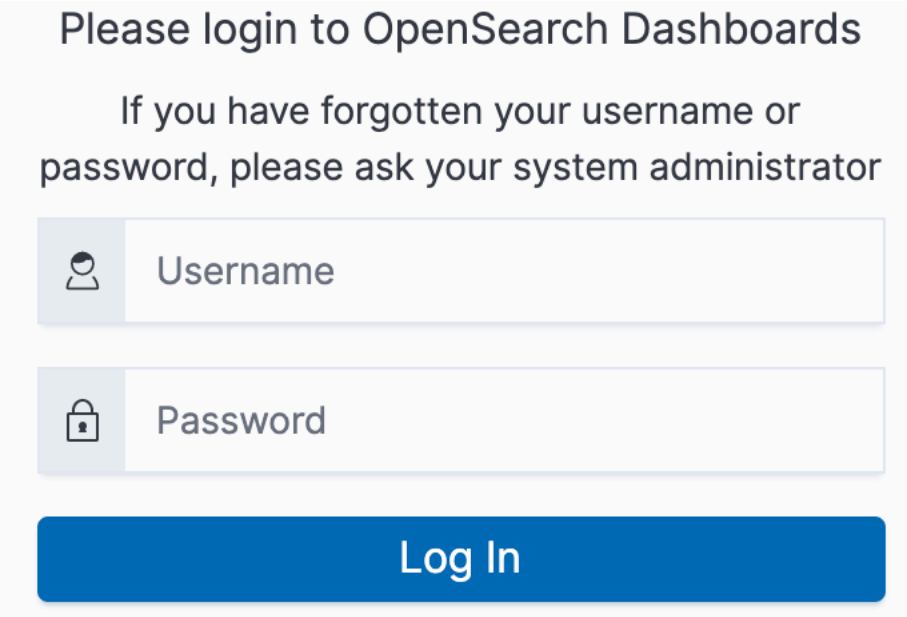
- Si desea utilizar IAM para la administración de usuarios, utilice the section called “Autenticación de Amazon Cognito para paneles OpenSearch ” para acceder a Dashboards. De lo contrario, Dashboards muestra una página de inicio de sesión no funcional. Consulte the section called “Limitaciones”.

  Con la autenticación de Amazon Cognito, uno de los roles asumidos del grupo de identidades debe coincidir con el rol de IAM especificado para el usuario maestro. Para obtener más información acerca de esta configuración, consulte the section called “(Opcional) Configuración de acceso pormenorizado” y the section called “Tutorial: Control de acceso detallado con la autenticación de Cognito”.



- Si decide utilizar la base de datos de usuarios interna, puede iniciar sesión en Dashboards con su nombre de usuario maestro y contraseña. Debe obtener acceso a Dashboards a través de HTTPS. La autenticación de Amazon Cognito y SAML para Dashboards reemplazan esta pantalla de inicio de sesión.

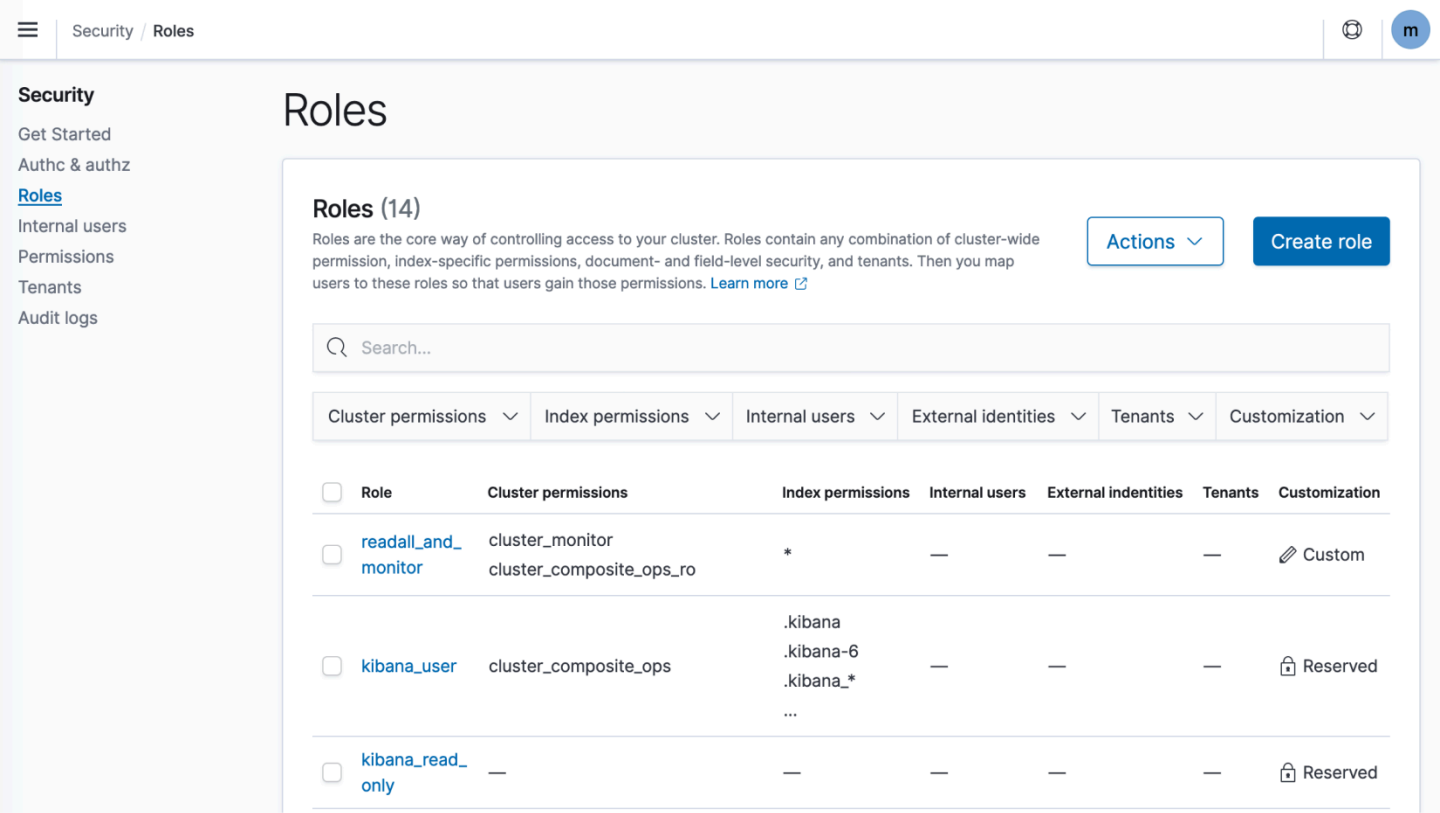
  Para obtener más información acerca de esta configuración, consulte the section called “Tutorial: Base de datos de usuarios interna con autenticación básica”.



- Si elige utilizar la autenticación SAML, puede iniciar sesión con credenciales de un proveedor de identidad externo. Para obtener más información, consulte the section called “Autenticación SAML para paneles OpenSearch”.

Administrar permisos

Tal y como se indica en the section called “Conceptos clave”, puede administrar los permisos de control de acceso detallado mediante roles, usuarios y mapeos. En esta sección, se describe cómo crear y aplicar esos recursos. Recomendamos que inicie sesión en Dashboards como usuario maestro para realizar estas operaciones.



> **Note**
>
> Los permisos que elige conceder a los usuarios varían ampliamente en función del caso de uso. No podemos cubrir de forma factible todos los escenarios de esta documentación. Al determinar qué permisos conceder a sus usuarios, asegúrese de consultar los permisos de OpenSearch clúster e indexación que se mencionan en las siguientes secciones y siga siempre el principio de privilegios mínimos.

Crear roles

Puede crear nuevas funciones para un control de acceso más detallado mediante los OpenSearch paneles de control o la `_plugins/_security` operación en la API REST. Para obtener más información, consulte Create roles (Crear roles).

El control de acceso detallado también incluye varios roles predefinidos. Los clientes como OpenSearch Dashboards y Logstash realizan una amplia variedad de solicitudes OpenSearch, lo que puede dificultar la creación manual de roles con el conjunto mínimo de permisos. Por ejemplo, el rol `opensearch_dashboards_user` incluye los permisos que un usuario necesita para trabajar con

patrones de índices, visualizaciones, paneles e inquilinos. Recomendamos mapearlo a cualquier rol de usuario o backend que tenga acceso a Dashboards, junto con los roles adicionales que permitan el acceso a otros índices.

Amazon OpenSearch Service no ofrece las siguientes OpenSearch funciones:

- `observability_full_access`
- `observability_read_access`
- `reports_read_access`
- `reports_full_access`

Amazon OpenSearch Service ofrece varios roles que no están disponibles en OpenSearch:

- `ultrawarm_manager`
- `ml_full_access`
- `cold_manager`
- `notifications_full_access`
- `notifications_read_access`

Seguridad en el nivel del clúster

Los permisos en el nivel del clúster incluyen la capacidad de ejecutar solicitudes generales como `_mget`, `_msearch` y `_bulk`, monitorear el estado o tomar instantáneas, etc. Administre estos permisos mediante la sección Permisos de clúster al crear un rol. Para obtener una lista completa de los permisos de nivel de clúster, consulte Permisos de cluster.

En lugar de permisos individuales, a menudo puede alcanzar la posición de seguridad deseada mediante una combinación de los grupos de acciones predeterminadas. Para obtener una lista de grupos de acciones en el nivel del clúster, consulte Nivel del clúster.

Seguridad en el nivel del índice

Los permisos en el nivel del índice incluyen la capacidad de crear nuevos índices, buscar índices, leer y escribir documentos, eliminar documentos, administrar alias y mucho más. Administre estos permisos con la sección Permisos de índice al crear un rol. Para obtener una lista completa de los permisos de nivel de índice, consulte Permisos de índice.

En lugar de permisos individuales, a menudo puede alcanzar la posición de seguridad deseada mediante una combinación de los grupos de acciones predeterminadas. Para obtener una lista de grupos de acciones en el nivel del índice, consulte Nivel del índice.

Seguridad en el nivel del documento

La seguridad en el nivel del documento permite restringir los documentos de un índice que puede ver un usuario. Al crear un rol, especifique un patrón de índice y una OpenSearch consulta. Cualquier usuario que se mapee a ese rol únicamente podrá ver los documentos que coincidan con la consulta. La seguridad en el nivel del documento afecta al número de aciertos que se reciben al realizar una búsqueda.

Para obtener más información, consulte Seguridad en el nivel del documento.

Seguridad en el nivel del campo

La seguridad en el nivel del campo permite controlar qué campos de documento puede ver un usuario. Al crear un rol, agregue una lista de campos para incluirlos o excluirlos. Si incluye campos, los usuarios que mapee a ese rol solo podrán ver esos campos. Si excluye campos, podrán ver todos los campos excepto los excluidos. La seguridad en el nivel del campo afecta al número de campos que se incluyen en los aciertos al realizar una búsqueda.

Para obtener más información, consulte Seguridad en el nivel del campo.

Enmascarar campos

El enmascaramiento de campos es una alternativa a la seguridad en el nivel del campo que permite anonimizar los datos de un campo en lugar de eliminarlos por completo. Al crear un rol, agregue una lista de campos que se van a enmascarar. El enmascaramiento de campos afecta a si se podrá ver el contenido de un campo al realizar búsquedas.

> **Tip**
>
> Si aplica el enmascaramiento estándar a un campo, OpenSearch Service utiliza un hash aleatorio y seguro que puede provocar resultados de agregación imprecisos. Para realizar agregaciones en campos enmascarados, utilice el enmascaramiento con base en patrones en su lugar.

Crear usuarios

Si habilitó la base de datos de usuarios interna, puede crear usuarios mediante OpenSearch paneles o la `_plugins/_security` operación de la API REST. Para obtener más información, consulte Crear usuarios.

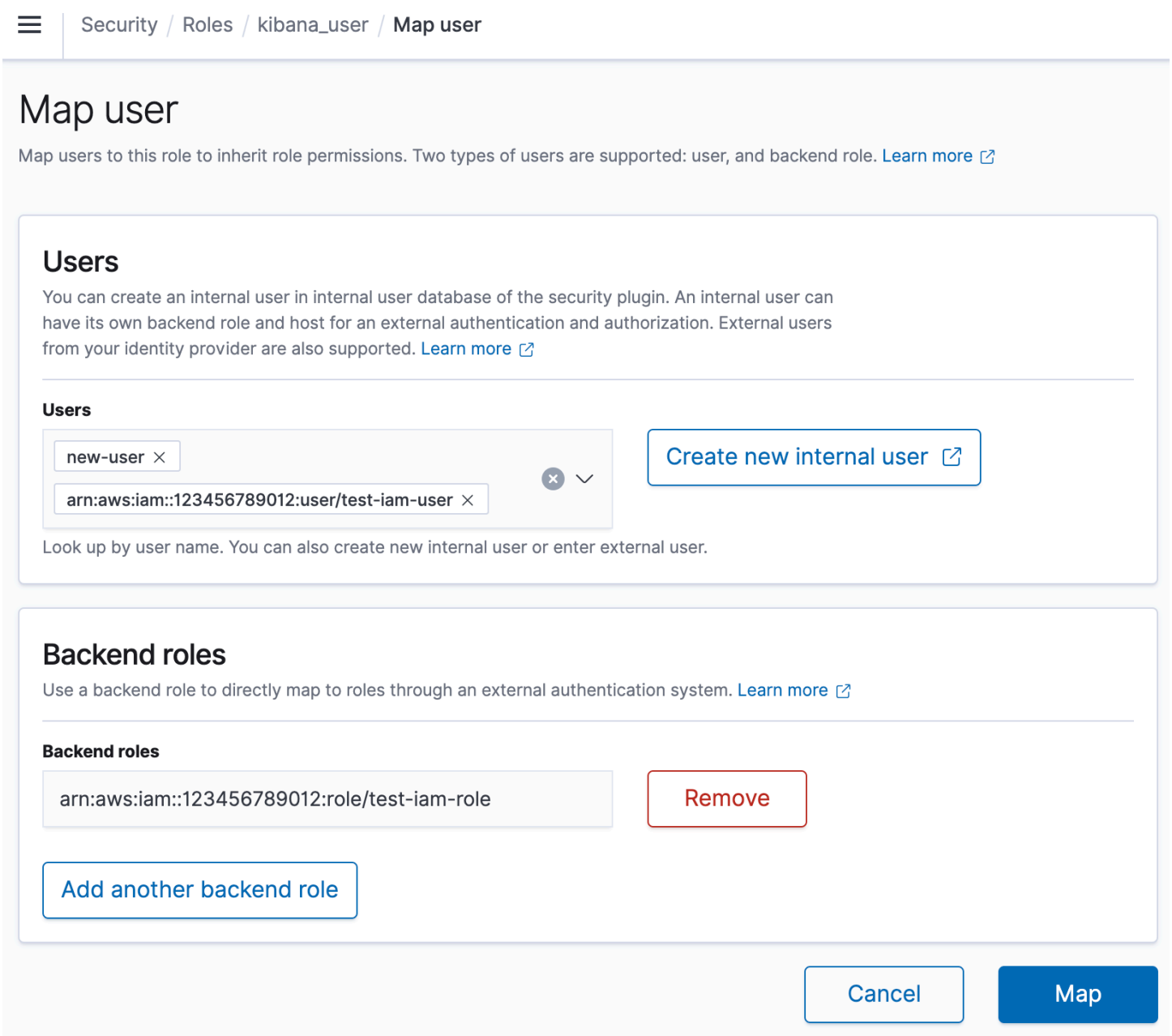
Si elige IAM para el usuario maestro, pase por alto esta parte de Dashboards. En su lugar, cree roles de IAM. Para obtener más información, consulte la Guía del usuario de IAM.

Mapear roles a usuarios

El mapeo de roles es el aspecto más crítico del control de acceso detallado. El control de acceso detallado tiene algunos roles predefinidos para ayudarlo a comenzar, pero a menos que mapee roles a los usuarios, cada solicitud al clúster dará lugar a un error de permisos.

Roles de backend pueden ayudar a simplificar el proceso de asignación de roles. En lugar de asignar el mismo rol a 100 usuarios individuales, puede asignar el rol a un único rol de backend que compartan los 100 usuarios. Los roles de backend pueden ser roles de IAM o cadenas arbitrarias.

- Especifique los usuarios ARNs, el usuario y las cadenas de usuario de Amazon Cognito en la sección Usuarios. Las cadenas de usuario de Cognito toman la forma de `Cognito/`*`user-pool-id`*`/`*`username`*.
- Especifique las funciones de backend y la función de IAM ARNs en la sección de funciones de backend.



Puede asignar funciones a los usuarios mediante los OpenSearch paneles de control o la `_plugins/_security` operación de la API REST. Para obtener más información, consulte Mapear usuarios a roles.

Crear grupos de acciones

Los grupos de acciones son conjuntos de permisos que puede reutilizar en diferentes recursos. Puede crear nuevos grupos de acciones mediante los OpenSearch paneles de control o la `_plugins/_security` operación de la API REST, aunque los grupos de acciones predeterminados

son suficientes para la mayoría de los casos de uso. Para obtener más información acerca de los grupos de acciones predeterminados, consulte Grupos de acciones predeterminados.

OpenSearch Tableros de mandos multiusuario

Los inquilinos son espacios para guardar patrones de índices, visualizaciones, paneles y otros objetos de Dashboards. El uso de varios inquilinos en Dashboards permite compartir el trabajo de forma segura con otros usuarios de Dashboards (o mantenerlo en privado) y configurar inquilinos de forma dinámica. Puede controlar qué roles tienen acceso a un inquilino y si esos roles tienen acceso de lectura o de escritura. El inquilino global es el predeterminado. Para obtener más información, consulte OpenSearch Dashboards sobre tenencia múltiple.

Para ver al inquilino actual o cambiar de inquilino

1. Ve a los OpenSearch paneles de control e inicia sesión.
2. Seleccione el icono de usuario en la parte superior derecha y elija Cambiar inquilinos.
3. Compruebe el inquilino antes de crear visualizaciones o paneles. Si desea compartir su trabajo con todos los demás usuarios de Dashboards, seleccione Global. Para compartir el trabajo con un subconjunto de usuarios de Dashboards, elija otro inquilino compartido. De lo contrario, seleccione Privado.

> **Note**
>
> OpenSearch Dashboards mantiene un índice independiente para cada inquilino y crea una plantilla de índice llamada. `tenant_template` No elimine ni modifique el `tenant_template` índice, ya que podría provocar un mal funcionamiento de los OpenSearch paneles si el mapeo del índice de inquilinos está mal configurado.

Configuraciones recomendadas

Debido a la forma en que el control de acceso detallado interactúa con otras características de seguridad, recomendamos varias configuraciones de control de acceso detallado que funcionan bien en la mayoría de los casos de uso.

Description (Descripción)	Usuario maestro	Política de acceso al dominio

Utilice las credenciales de IAM para las llamadas a los paneles y utilice la OpenSearch APIs autenticación SAML para acceder a los paneles. Administre los roles de control de acceso detallado mediante Dashboards o la API de REST.	Rol o usuario de 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* /*"
		}
		]
		}

Utilice las credenciales de IAM o la autentica ción básica para las llamadas al. OpenSearc h APIs Administre los roles de control de acceso detallado mediante Dashboards o la API de REST.
 	
Esta configuración ofrece mucha flexibilidad, especialmente si tiene OpenSearch clientes que solo admiten la autentica ción básica.
	Nombre de usuario y contraseñ a	JSON
		{
		"Version":"2012-10-17",
		"Statement": [
		{
		"Effect": "Allow",
		"Principal": {
		"AWS": "*"
		},
		"Action": "es:ESHttp*",
		"Resource": "arn:aws:es: *us-east-1* :*111122223333* :domain/*domain-name* /*"
		}
		]
		}

Description (Descripción)	Usuario maestro	Política de acceso al dominio
Si tiene un proveedor de identidad existente, utilice autenticación SAML para acceder a Dashboard s. En caso contrario, administre los usuarios de Dashboards en la base de datos de usuarios interna.		

Utilice las credenciales de IAM para las llamadas a y Amazon Cognito para acceder a los paneles de control. OpenSearc h APIs Administre los roles de control de acceso detallado mediante Dashboards o la API de REST.	Rol o usuario de 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* /*"
		}
		]
		}

Description (Descripción)	Usuario maestro	Política de acceso al dominio

Utilice las credenciales de IAM para las llamadas a los paneles y bloquee la OpenSearch APIs mayoría de los accesos a ellos. Administre los roles de control de acceso detallado mediante la API de REST.	Rol o usuario de 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*"
		}
		]
		}

Limitaciones

El control de acceso detallado tiene varias limitaciones importantes:

- El aspecto `hosts` de los mapeos de roles, que mapea roles a nombres de anfitrión o direcciones IP, no funciona si el dominio está dentro de una VPC. Aun así, puede mapear roles a usuarios y roles de backend.

- Si elige IAM para el usuario maestro y no habilita la autenticación de Amazon Cognito o SAML, Dashboards muestra una página de inicio de sesión no funcional.
- Si elige IAM para el usuario maestro, podrá crear usuarios en la base de datos de usuarios interna. Sin embargo, debido a que la autenticación básica de HTTP no está habilitada en esta configuración, todas las solicitudes firmadas con esas credenciales de usuario se rechazarán.
- Si utiliza SQL para consultar un índice al que no tiene acceso, recibirá un error debido a que no tiene los permisos necesarios. Si el índice no existe, recibirá un error debido a que ese índice no existe. Esta diferencia en los mensajes de error significa que puede confirmar si un índice existe o no, en el caso de que trate de adivinar su nombre.

  Para minimizar el problema, no incluya información confidencial en los nombres de índice. Para denegar por completo el acceso a SQL, agregue el siguiente elemento a la política de acceso al dominio:

```
{
  "Effect": "Deny",
  "Principal": {
    "AWS": [
      "*"
    ]
  },
  "Action": [
    "es:*"
  ],
  "Resource": "arn:aws:es:us-east-1:123456789012:domain/my-domain/_plugins/_sql"
}
```

- Si la versión de su dominio es 2.3 o superior y tiene activado el control de acceso detallado, configurar `max_clause_count` en 1 provocará problemas con su dominio. Le recomendamos configurar esta cuenta con un número superior.
- Si está habilitando un control de acceso detallado en un dominio en el que no está configurado un control de acceso detallado, en el caso de los orígenes de datos creados para consultas directas, tendrá que configurar las funciones de control de acceso detalladas. Para obtener más información sobre cómo configurar funciones de acceso detalladas, consulte Creación de integraciones de fuentes de datos de Amazon OpenSearch Service con Amazon S3.

Modificar el usuario maestro

Si olvida los detalles del usuario maestro, puede reconfigurarlo mediante la consola, la AWS CLI o la API de configuración.

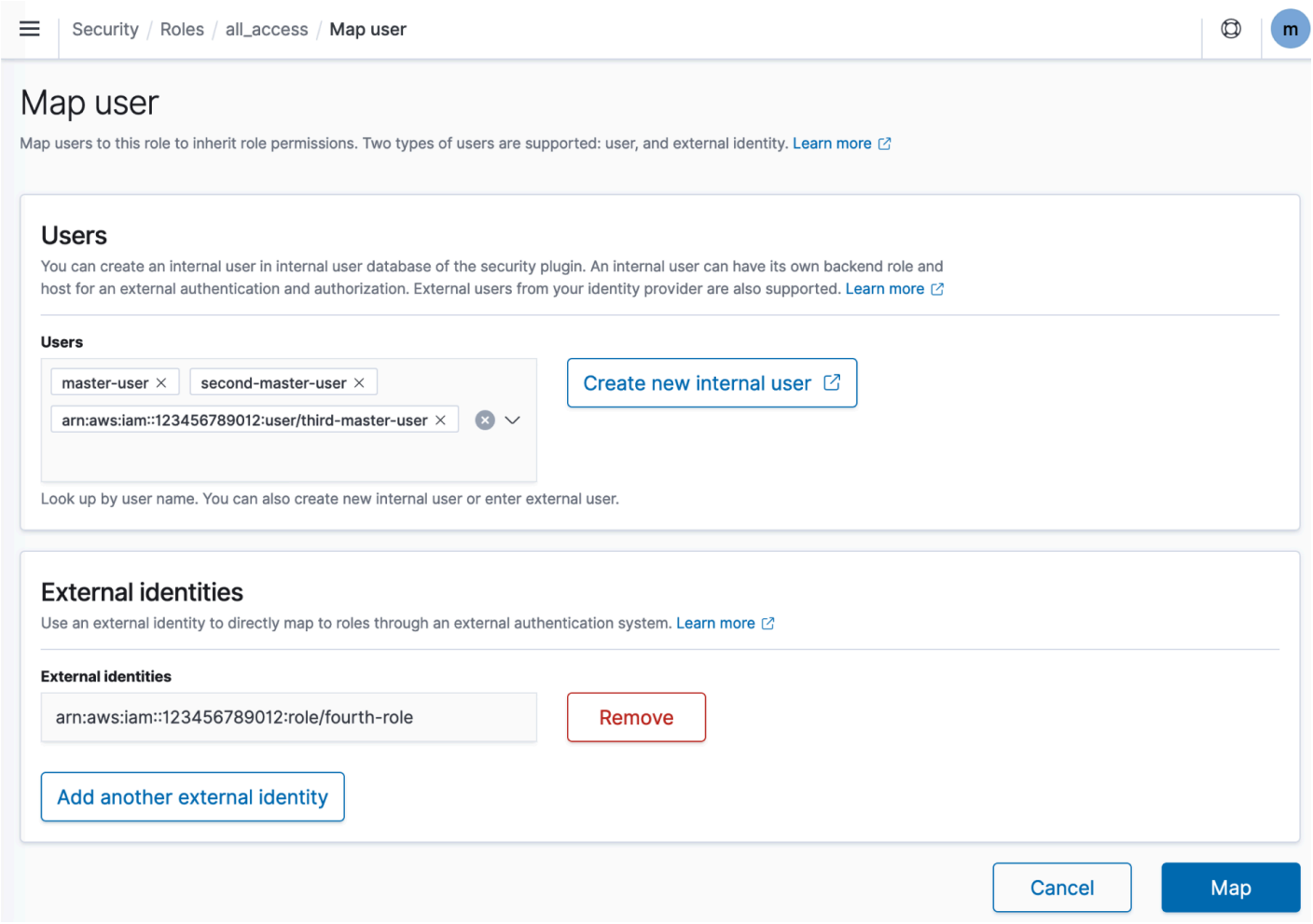
Para modificar el usuario maestro (consola)

1. Navega hasta la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa/.
2. Seleccione el dominio y seleccione Acciones, Editar la configuración de seguridad.
3. Elika Establecer ARN de IAM como usuario maestro o bien Crear usuario maestro.
   - Si ha utilizado anteriormente un usuario maestro de IAM, el control de acceso detallado volverá a mapear el rol `all_access` al nuevo ARN de IAM que especifique.
   - Si ha utilizado previamente la base de datos de usuarios interna, el control de acceso detallado creará un usuario maestro. Puede utilizar el nuevo usuario maestro para eliminar el anterior.
   - Cambiar la base de datos de usuario interna a un usuario maestro de IAM no elimina ningún usuario de la base de datos de usuarios interna. En su lugar, simplemente desactiva la autenticación básica de HTTP. Elimine manualmente usuarios de la base de datos interna de usuarios o guárdelos en caso de que necesite volver a habilitar la autenticación básica de HTTP.
4. Seleccione Guardar cambios.

Usuarios maestros adicionales

Se designa un usuario maestro al crear un dominio. Sin embargo, si lo desea, puede utilizar este usuario maestro para crear usuarios maestros adicionales. Tienes dos opciones: OpenSearch paneles de control o la API REST.

- En Dashboards, seleccione Seguridad, Roles y, a continuación, mapee el nuevo usuario maestro a los roles `all_access` y `security_manager`.



- Para utilizar la API de REST, envíe las siguientes solicitudes:

```
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"
  ]
}
```

Estas solicitudes reemplazan los mapeos de roles actuales. Por consiguiente, deber realizar las solicitudes `GET` primero para que pueda incluir todos los roles actuales en las solicitudes `PUT`. La API REST resulta especialmente útil si no puede obtener acceso a Dashboards y desea mapear un rol de IAM desde Amazon Cognito al rol `all_access`.

Instantáneas manuales

El control de acceso detallado presenta algunas complicaciones adicionales con la toma de instantáneas manuales. Para registrar un repositorio de instantáneas, aunque utilice la autenticación básica de HTTP para todos los demás fines, debe mapear el rol `manage_snapshots` a un rol de IAM que tenga permisos `iam:PassRole` para asumir `TheSnapshotRole`, tal como se define en the section called “Requisitos previos”.

A continuación, utilice ese rol de IAM para enviar una solicitud firmada al dominio, como se describe en the section called “Registrar un repositorio de instantáneas manuales”.

Integraciones

Si utiliza otros AWS servicios con el OpenSearch Servicio, debe proporcionar las funciones de IAM para esos servicios con los permisos adecuados. Por ejemplo, los flujos de entrega de Firehose suelen utilizar un rol de IAM denominado `firehose_delivery_role`. En Dashboards, cree un rol para el control de acceso detallado y mapee el rol de IAM a ese rol. En este caso, el nuevo rol necesita los siguientes permisos:

```
{
  "cluster_permissions": [
    "cluster_composite_ops",
    "cluster_monitor"
  ],
```

```
  "index_permissions": [{
    "index_patterns": [
      "firehose-index*"
    ],
    "allowed_actions": [
      "create_index",
      "manage",
      "crud"
    ]
  }]
}
```

Los permisos varían en función de las acciones que realice cada servicio. Es probable que una AWS IoT regla o AWS Lambda función que indexe datos necesite permisos similares a los de Firehose, mientras que una función de Lambda que solo realiza búsquedas puede utilizar un conjunto más limitado.

Diferencias en la API de REST

La API REST de control de acceso preciso presenta ligeras diferencias en función de la versión de OpenSearch/Elasticsearch . Antes de realizar una solicitud PUT, realice una solicitud GET para saber cómo es el cuerpo de la solicitud esperado. Por ejemplo, una solicitud GET para `_plugins/_security/api/user` devuelve todos los usuarios, que luego puede modificar y utilizar para realizar solicitudes PUT válidas.

En Elasticsearch 6.x, las solicitudes para crear usuarios tienen este aspecto:

```
PUT _opendistro/_security/api/user/new-user
{
  "password": "some-password",
  "roles": ["new-backend-role"]
}
```

En OpenSearch Elasticsearch 7.x, las solicitudes tienen este aspecto (`_pluginscámbiese` a si usa Elasticsearch): `_opendistro`

```
PUT _plugins/_security/api/user/new-user
{
  "password": "some-password",
  "backend_roles": ["new-backend-role"]
}
```

Además, los inquilinos son propiedades de los roles en Elasticsearch 6.x:

```
GET _opendistro/_security/api/roles/all_access

{
  "all_access": {
    "cluster": ["UNLIMITED"],
    "tenants": {
      "admin_tenant": "RW"
    },
    "indices": {
      "*": {
        "*": ["UNLIMITED"]
      }
    },
    "readonly": "true"
  }
}
```

En OpenSearch Elasticsearch 7.x, son objetos con su propio URI (cámbielo si usa Elasticsearch):: `_plugins` `_opendistro`

```
GET _plugins/_security/api/tenants

{
  "global_tenant": {
    "reserved": true,
    "hidden": false,
    "description": "Global tenant",
    "static": false
  }
}
```

Para obtener documentación sobre la API OpenSearch REST, consulta la referencia sobre la API del complemento de seguridad.

Tip

Si utiliza la base de datos de usuarios interna, puede utilizar curl para realizar solicitudes y probar el dominio. Pruebe los siguientes comandos de muestra:

```
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: Configurar un dominio con un usuario maestro de IAM y la autenticación de Amazon Cognito

En este tutorial se describe un caso de uso popular de Amazon OpenSearch Service para un control de acceso detallado: un usuario maestro de IAM con autenticación de Amazon Cognito para paneles de control. OpenSearch

En el tutorial, configuraremos un rol de IAM maestro y un rol de IAM limitado, que luego asociaremos a los usuarios de Amazon Cognito. A continuación, el usuario maestro puede iniciar sesión en OpenSearch Dashboards, asignar al usuario limitado a un rol y usar un control de acceso detallado para limitar los permisos del usuario.



Aunque estos pasos utilizan el grupo de usuarios de Amazon Cognito para la autenticación, este mismo proceso básico funciona para cualquier proveedor de autenticación de Cognito que permita asignar diferentes roles de IAM a diferentes usuarios.

En este tutorial, deberá completar los siguientes pasos:

1. Crear roles de IAM maestros y limitados
2. Crear un dominio con la autenticación de Cognito
3. Configurar un grupo de usuarios y grupo de identidades de Cognito
4. Asigne funciones en los paneles OpenSearch
5. Evaluar los permisos

Paso 1: Crear roles de IAM maestros y limitados

Diríjase a la consola AWS Identity and Access Management (IAM) y cree dos funciones distintas:

- `MasterUserRole`: es el usuario maestro, que tiene permisos completos para el clúster y, además, administra los roles y las asignaciones de los roles.
- `LimitedUserRole`: es un rol más restringido, al que concederá acceso limitado como usuario maestro.

Para obtener instrucciones acerca de cómo crear roles, consulte Creación de un rol mediante políticas de confianza personalizadas en la Guía del usuario de IAM.

Ambos roles deben tener la siguiente política de confianza, la cual permite que el grupo de identidades de Cognito asuma los roles:

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**
>
> Reemplace `identity-pool-id` por el identificador único de su grupo de identidades de Amazon Cognito. Por ejemplo, `us-east-1:0c6cdba7-3c3c-443b-a958-fb9feb207aa6`.

Paso 2: Crear un dominio con la autenticación de Cognito

Ve a la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/casa/ y crea un dominio con los siguientes ajustes:

- OpenSearch 1.0 o posterior, o Elasticsearch 7.8 o posterior
- Acceso público
- Control de acceso detallado activado con `MasterUserRole` como el usuario maestro (creado en el paso anterior)
- Autenticación de Amazon Cognito habilitada para los paneles OpenSearch . Para obtener instrucciones sobre cómo habilitar la autenticación de Cognito y seleccionar un grupo de usuarios e identidades, consulte the section called “Configuración de un dominio para utilizar la autenticación de Amazon Cognito”.
- La siguiente política de acceso 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}/*"
    }
  ]
}
```

- Se requiere HTTPS para todo el tráfico al dominio
- Node-to-node cifrado
- Cifrado de datos en reposo

Paso 3: configurar usuarios de Cognito

Mientras se crea su dominio, configure los usuarios maestros y limitados en Amazon Cognito siguiendo Crear un grupo de usuarios en la Guía para desarrolladores de Amazon Cognito. Por último, configure su grupo de identidades siguiendo los pasos de Crear un grupo de identidades en Amazon Cognito. El grupo de usuarios y el grupo de identidades deben estar en la misma Región de AWS.

Paso 4: mapear las funciones en los OpenSearch paneles

Ahora que sus usuarios están configurados, puede iniciar sesión en OpenSearch Dashboards como usuario maestro y asignar los usuarios a los roles.

1. Vuelva a la consola OpenSearch de servicio y navegue hasta la URL de los OpenSearch paneles del dominio que creó. La URL tiene este formato: *domain-endpoint*`/_dashboards/`.
2. Inicie sesión con las credenciales de `master-user`.
3. Seleccione Agregar datos de muestra y agregue los datos de vuelos de muestra.
4. En el panel de navegación izquierdo, seleccione Seguridad, Roles, Crear rol.
5. Llame al rol `new-role`.
6. En Índice, especifique `opensearch_dashboards_sample_data_fli*` (`kibana_sample_data_fli*` en los dominios de Elasticsearch).
7. Para obtener Permisos de índice, seleccione leer.
8. Para Seguridad a nivel del documento, especifique la siguiente consulta:

```
{
  "match": {
    "FlightDelay": true
  }
}
```

9. Para la seguridad a nivel del campo, seleccione Excluir y especifique `FlightNum`.
10. Para Anonimización, especifique `Dest`.

11. Seleccione Crear.
12. Seleccione Usuarios asignados, Administrar mapeo. A continuación, agregue el Nombre de recurso de Amazon (ARN) para `LimitedUserRole` como una identidad externa y elija Asignar.
13. Vuelva a la lista de roles y elija opensearch_dashboards_user. Seleccione Usuarios asignados, Administrar mapeo. Agregue el ARN para `LimitedUserRole` como un rol de backend y elija Mapear.

Paso 5: Probar los permisos

Cuando los roles están asignados de manera correcta, puede iniciar sesión como el usuario limitado y probar los permisos.

1. En una nueva ventana privada del navegador, navegue hasta la URL de OpenSearch Dashboards del dominio, inicie sesión con las `limited-user` credenciales y seleccione Explorar por mi cuenta.
2. Vaya a Herramientas para desarrolladores y ejecute la búsqueda predeterminada:

   ```
   GET _search
   {
     "query": {
       "match_all": {}
     }
   }
   ```

   Observe el error de permisos. `limited-user` no tiene permisos para ejecutar búsquedas en todo el clúster.
3. Ejecute otra búsqueda:

   ```
   GET opensearch_dashboards_sample_data_flights/_search
   {
     "query": {
       "match_all": {}
     }
   }
   ```

   Tenga en cuenta que todos los documentos coincidentes tienen un campo `FlightDelay` de `true`, un campo `Dest` anonimizado y ningún campo `FlightNum`.

4. En la ventana original del navegador, tras haber iniciado sesión como `master-user`, seleccione Herramientas para desarrolladores y, a continuación, realice las mismas búsquedas. Observe la diferencia en permisos, número de aciertos, documentos coincidentes y campos incluidos.

Tutorial: Configurar un dominio con la base de datos de usuarios interna y la autenticación básica de HTTP

En este tutorial se describe otro caso de uso muy detallado del control de acceso: un usuario maestro en la base de datos de usuarios interna y la autenticación básica mediante HTTP para los paneles de control. OpenSearch A continuación, el usuario maestro puede iniciar sesión en OpenSearch Dashboards, crear un usuario interno, asignar al usuario a un rol y usar un control de acceso detallado para limitar los permisos del usuario.

En este tutorial, deberá completar los siguientes pasos:

1. Cree un dominio con un usuario maestro
2. Configure un usuario interno en los paneles OpenSearch
3. Asigne funciones en los paneles OpenSearch
4. Evaluar los permisos

Paso 1: crear un dominio

Ve a la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/casa/ y crea un dominio con los siguientes ajustes:

- OpenSearch 1.0 o posterior, o Elasticsearch 7.9 o posterior
- Acceso público
- Control de acceso detallado con un usuario maestro en la base de datos de usuarios interna (`TheMasterUser` en el resto de este tutorial)
- Autenticación de Amazon Cognito para Dashboards deshabilitada
- La siguiente política de acceso:

  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}/*"
    }
  ]
}
```

- Se requiere HTTPS para todo el tráfico al dominio
- Node-to-node cifrado
- Cifrado de datos en reposo

Paso 2: Crear un usuario interno en los OpenSearch paneles

Ahora que tiene un dominio, puede iniciar sesión en OpenSearch Dashboards y crear un usuario interno.

1. Vuelva a la consola OpenSearch de servicio y navegue hasta la URL de los OpenSearch paneles del dominio que creó. La URL tiene este formato: *domain-endpoint*`/_dashboards/`.
2. Inicie sesión con el `TheMasterUser`.
3. Seleccione Agregar datos de muestra y agregue los datos de vuelos de muestra.
4. En el panel de navegación izquierdo, seleccione Seguridad, Usuarios internos y Crear usuario interno.
5. Llame al usuario `new-user` y especifique una contraseña. A continuación, seleccione Crear.

Paso 3: Asigne funciones en los paneles OpenSearch

Ahora que el usuario está configurado, puede asignarlo a un rol.

1. Permanezca en la sección Seguridad de los OpenSearch paneles de control y elija Roles y Crear rol.
2. Llame al rol `new-role`.

3. Para Índice, especifique `opensearch_dashboards_sample_data_fli*` (`kibana_sample_data_fli*` en dominios Elasticsearch) para el patrón de índice.
4. Para el grupo de acciones, seleccione leer.
5. Para Seguridad a nivel del documento, especifique la siguiente consulta:

```
{
  "match": {
    "FlightDelay": true
  }
}
```

6. Para la seguridad a nivel del campo, seleccione Excluir y especifique `FlightNum`.
7. Para Anonimización, especifique `Dest`.
8. Seleccione Crear.
9. Seleccione Usuarios asignados, Administrar mapeo. A continuación, agregue `new-user` a Usuarios y elija Mapear.
10. Vuelva a la lista de roles y elija opensearch_dashboards_user. Seleccione Usuarios asignados, Administrar mapeo. A continuación, agregue `new-user` a Usuarios y elija Mapear.

Paso 4: probar los permisos

Cuando los roles están asignados de manera correcta, puede iniciar sesión como el usuario limitado y probar los permisos.

1. En una nueva ventana privada del navegador, vaya a la URL de OpenSearch Dashboards del dominio, inicie sesión con las `new-user` credenciales y elija Explorar por mi cuenta.
2. Vaya a Herramientas para desarrolladores y ejecute la búsqueda predeterminada:

```
GET _search
{
  "query": {
    "match_all": {}
  }
}
```

   Observe el error de permisos. `new-user` no tiene permisos para ejecutar búsquedas en todo el clúster.

3. Ejecute otra búsqueda:

```
GET dashboards_sample_data_flights/_search
{
  "query": {
    "match_all": {}
  }
}
```

   Tenga en cuenta que todos los documentos coincidentes tienen un campo `FlightDelay` de `true`, un campo `Dest` anonimizado y ningún campo `FlightNum`.

4. En la ventana original del navegador, inicie sesión como `TheMasterUser`, seleccione Herramientas para desarrolladores y realice las mismas búsquedas. Observe la diferencia en permisos, número de aciertos, documentos coincidentes y campos incluidos.

Validación de conformidad para Amazon OpenSearch Service

Los auditores externos evalúan la seguridad y el cumplimiento de Amazon OpenSearch Service como parte de varios programas de AWS cumplimiento. Entre estos, se incluyen SOC, PCI e HIPAA.

Si tienes requisitos de conformidad, considera usar cualquier versión de OpenSearch Elasticsearch 6.0 o posterior. Las versiones anteriores de Elasticsearch no ofrecen una combinación de cifrado de datos en reposo y node-to-node cifrado y es poco probable que satisfagan tus necesidades. También puedes considerar usar cualquier versión de Elasticsearch 6.7 OpenSearch o posterior si es importante tener un control de acceso detallado para tu caso de uso. En cualquier caso, elegir una versión concreta OpenSearch o de Elasticsearch al crear un dominio no garantiza la conformidad.

Para saber si uno Servicio de AWS está dentro del ámbito de aplicación de programas de cumplimiento específicos, consulte Servicios de AWS Alcance por programa de cumplimiento Servicios de AWS y elija el programa de cumplimiento que le interese. Para obtener información general, consulte Programas de AWS cumplimiento > Programas AWS .

Puede descargar informes de auditoría de terceros utilizando AWS Artifact. Para obtener más información, consulte Descarga de informes en AWS Artifact .

Su responsabilidad de cumplimiento al Servicios de AWS utilizarlos viene determinada por la confidencialidad de sus datos, los objetivos de cumplimiento de su empresa y las leyes y reglamentos aplicables. Para obtener más información sobre su responsabilidad de conformidad al utilizarlos Servicios de AWS, consulte AWS la documentación de seguridad.

Resiliencia en Amazon OpenSearch Service

La infraestructura AWS global se basa en distintas zonas Regiones de AWS de disponibilidad. Regiones de AWS proporcionan varias zonas de disponibilidad aisladas y separadas físicamente, que están conectadas mediante redes de baja latencia, alto rendimiento y alta redundancia. Con las zonas de disponibilidad, puede diseñar y utilizar aplicaciones y bases de datos que realizan una conmutación por error automática entre zonas de disponibilidad sin interrupciones. Las zonas de disponibilidad tienen una mayor disponibilidad, tolerancia a errores y escalabilidad que las infraestructuras tradicionales de centros de datos únicos o múltiples.

Para obtener más información sobre las zonas de disponibilidad Regiones de AWS y las zonas de disponibilidad, consulte Infraestructura global.AWS

Además de la infraestructura AWS global, el OpenSearch servicio ofrece varias funciones para ayudarlo a satisfacer sus necesidades de respaldo y resiliencia de datos:

- Dominios Multi-AZ y fragmentos replicados
- Instantáneas automatizadas y manuales

Autenticación y autorización JWT para Amazon Service OpenSearch

Amazon OpenSearch Service ahora te permite usar JSON Web Tokens (JWTs) para la autenticación y la autorización. JWTs son tokens de acceso basados en JSON que se utilizan para conceder el acceso de inicio de sesión único (SSO). Puedes usar JWTs In OpenSearch Service para crear tokens de inicio de sesión único para validar las solicitudes enviadas a tu dominio de Service. OpenSearch Para usarlos JWTs, debes tener habilitado un control de acceso detallado y debes proporcionar una clave pública válida con formato PEM de RSA o ECDSA. Para obtener más información sobre el control de acceso detallado, consulta Control de acceso detallado en Amazon Service. OpenSearch

Puede configurar los tokens web JSON mediante la consola de OpenSearch servicio, el () o el AWS Command Line Interface .AWS CLI AWS SDKs

Consideraciones

Antes de usarlo JWTs con Amazon OpenSearch Service, debes tener en cuenta lo siguiente:

- Debido al tamaño de las claves públicas de RSA en formato PEM, se recomienda utilizar la consola de AWS para configurar la autenticación y la autorización de JWT.
- Debe proporcionar usuarios y roles válidos al especificar los campos de asuntos y roles para sus campos; de lo contrario JWTs, se denegarán las solicitudes.
- OpenSearch La 2.11 es la primera versión compatible que se puede utilizar para la autenticación JWT.

Modificación de la política de acceso al dominio

Antes de poder configurar el dominio para usar la autenticación y la autorización de JWT, debe actualizar la política de acceso al dominio para permitir a los usuarios de JWT acceder al dominio. De lo contrario, se deniegan todas las solicitudes entrantes autorizadas de JWT. La política de acceso al dominio recomendada para proporcionar acceso total a los subrecursos (/*) es:

JSON

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

Configuración de la autenticación y la autorización de JWT

Puede habilitar la autenticación y la autorización de JWT durante el proceso de creación del dominio o mediante la actualización de un dominio existente. Los pasos de configuración varían ligeramente según la opción que seleccione.

Los siguientes pasos explican cómo configurar un dominio existente para la autenticación y autorización de JWT en la OpenSearch consola de servicio:

1. En Configuración del dominio, vaya a Autenticación y autorización de JWT y seleccione Habilitar la autenticación y autorización de JWT. OpenSearch
2. Configure la clave pública que se usará en el dominio. Para ello, puede cargar un archivo PEM que contenga una clave pública o ingresarla manualmente.

   > **Note**
   >
   > Si la clave cargada o ingresada no es válida, aparecerá una advertencia sobre el cuadro de texto que especifica el problema.

3. (Opcional) En Configuración adicional, puede configurar los siguientes campos opcionales:

   - Clave de asunto: puedes dejar este campo vacío para usar la `sub` clave predeterminada para tu. JWTs
   - Clave de roles: puede dejar este campo vacío para usar la `roles` clave predeterminada para su JWTs.

   Una vez hechos los cambios, guarde el dominio.

Uso de un JWT para enviar una solicitud de prueba

Tras crear un nuevo JWT con un par específico de sujeto y rol, puede enviar una solicitud de prueba. Para ello, utiliza la clave privada para firmar tu solicitud a través de la herramienta que creó el JWT. OpenSearch El servicio puede validar la solicitud entrante al verificar esta firma.

> **Note**
>
> Si especificó una clave de asunto o de roles personalizada para el JWT, debe utilizar los nombres de reclamaciones correctos para el JWT.

El siguiente es un ejemplo de cómo usar un token JWT para acceder al OpenSearch Servicio a través del punto final de búsqueda de su dominio:

```
curl -XGET "$search_endpoint" -H "Authorization: Bearer <JWT>"
```

Configuración de la autenticación y la autorización de JWT (AWS CLI)

El siguiente AWS CLI comando habilita la autenticación y la autorización de JWT OpenSearch siempre que el dominio exista:

```
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>"}}'
```

Configuración de la autenticación y la autorización de JWT (configuración a través de la API)

La siguiente solicitud a la API de configuración habilita la autenticación y la autorización de JWT OpenSearch en un dominio existente:

```
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"
   }
  }
}
```

Generación de un par de claves

JWTs Para configurar su OpenSearch dominio, necesitará proporcionar una clave pública en formato de correo mejorado con privacidad (PEM). Actualmente, Amazon OpenSearch Service admite dos algoritmos de cifrado asimétrico cuando se utiliza JWTs: RSA y ECDSA.

Para crear un par de claves de RSA mediante la biblioteca openssl común, siga estos pasos:

1. `openssl genrsa -out privatekey.pem 2048`
2. `openssl rsa -in privatekey.pem -pubout -out publickey.pem`

En este ejemplo, el `publickey.pem` archivo contiene la clave pública para usar con Amazon OpenSearch Service, mientras que `privatekey.pem` contiene la privada para firmar lo JWTs enviado al servicio. Además, tienes la opción de convertir la clave privada al `pkcs8` formato más utilizado si lo necesitas para generar la tuya JWTs.

Si utiliza el botón de carga para agregar un archivo PEM directamente a la consola, el archivo debe tener una extensión `.pem`; otras extensiones de archivo, como `.crt`, `.cert`, o `.key` no se admiten actualmente.

Seguridad de la infraestructura en Amazon OpenSearch Service

Como servicio gestionado, está protegido por la seguridad de la red AWS global. Para obtener información sobre los servicios AWS de seguridad y cómo se AWS protege la infraestructura, consulte Seguridad AWS en la nube. Para diseñar su AWS entorno utilizando las mejores prácticas de seguridad de la infraestructura, consulte Protección de infraestructuras en un marco de buena AWS arquitectura basado en el pilar de la seguridad.

Utiliza las llamadas a la API AWS publicadas para acceder a través de la red. Los clientes deben admitir lo siguiente:

- Seguridad de la capa de transporte (TLS). Exigimos TLS 1.2 y recomendamos TLS 1.3.
- Conjuntos de cifrado con confidencialidad directa total (PFS) como DHE (Ephemeral Diffie-Hellman) o ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). La mayoría de los sistemas modernos como Java 7 y posteriores son compatibles con estos modos.

Las llamadas a la API AWS publicadas se utilizan para acceder a la API OpenSearch de configuración del servicio a través de la red. A fin de configurar la versión mínima requerida de TLS para aceptar, especifique el valor `TLSSecurityPolicy` en las opciones de punto de conexión del dominio:

```
aws opensearch update-domain-config --domain-name my-domain --domain-endpoint-options
 '{"TLSSecurityPolicy": "Policy-Min-TLS-1-2-2019-07"}'
```

Para obtener más información, consulte la Referencia de comandos de la AWS CLI.

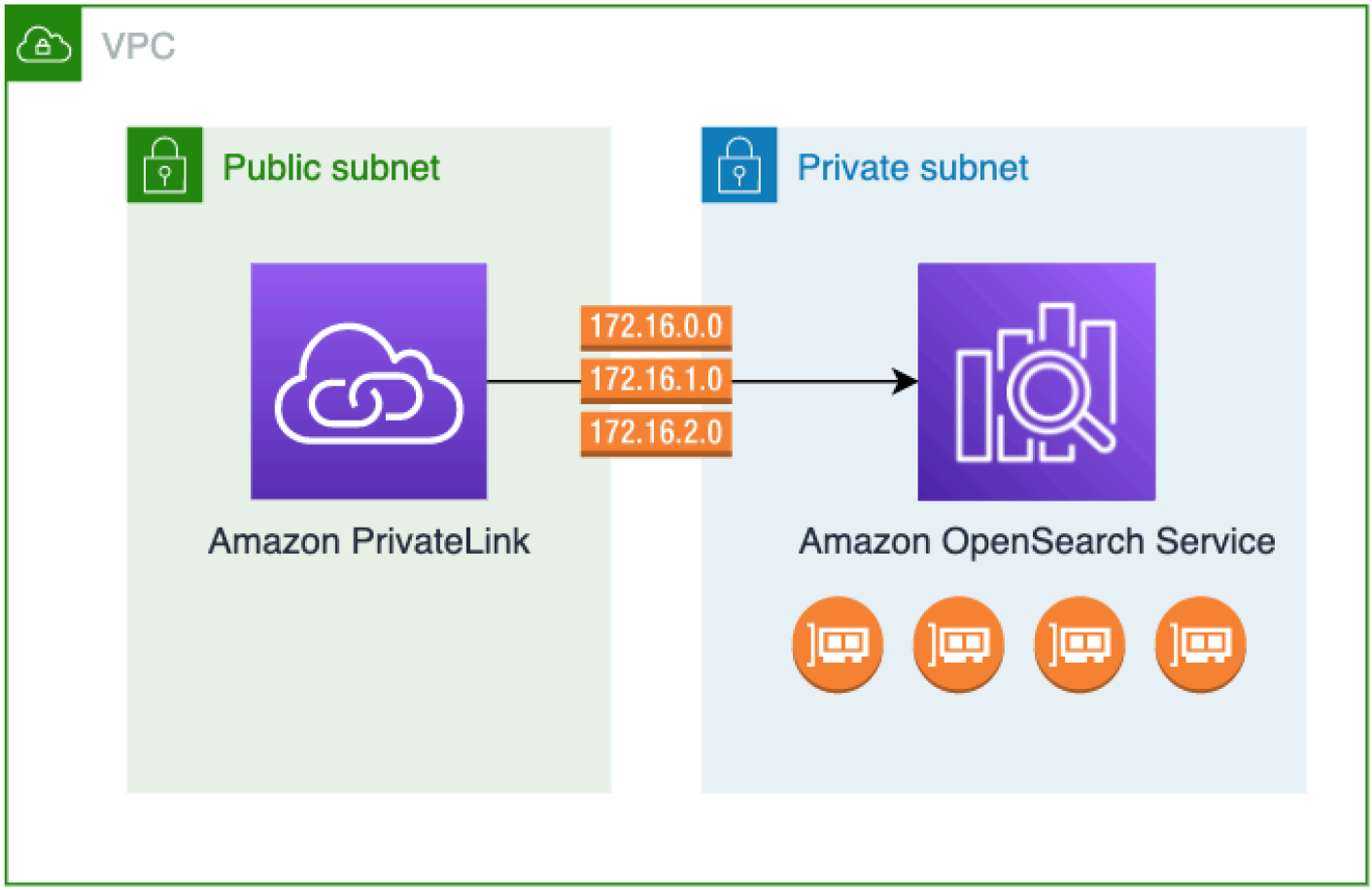
En función de la configuración del dominio, es posible que también tengas que firmar las solicitudes dirigidas al OpenSearch APIs. Para obtener más información, consulte the section called “Realizar y firmar solicitudes OpenSearch de servicio”.

OpenSearch El servicio admite dominios de acceso público, que pueden recibir solicitudes desde cualquier dispositivo conectado a Internet, y dominios de acceso de VPC, que están aislados de la Internet pública.

Acceda a Amazon OpenSearch Service mediante un punto de conexión de OpenSearch VPC gestionado por el servicio ()AWS PrivateLink

Para acceder a un dominio de Amazon OpenSearch Service, configura un punto final de OpenSearch VPC gestionado por el servicio (con tecnología de). AWS PrivateLink Estos puntos de conexión crean una conexión privada entre tu VPC y Amazon OpenSearch Service. Puede acceder a los dominios de la VPC de OpenSearch servicio como si estuvieran en su VPC, sin utilizar una puerta de enlace a Internet, un dispositivo NAT, una conexión VPN o una conexión. Direct Connect Las instancias de su VPC no necesitan direcciones IP públicas para acceder OpenSearch al Servicio.

Puede configurar los dominios de OpenSearch servicio para exponer puntos de conexión adicionales que se ejecuten en subredes públicas o privadas dentro de la misma VPC, en una VPC diferente o en una diferente. Cuentas de AWS Esto le permite agregar una capa adicional de seguridad para acceder a sus dominios sin importar dónde se ejecuten, sin necesidad de administrar una infraestructura. En el siguiente diagrama, se muestran los puntos finales OpenSearch de VPC gestionados por el servicio dentro de la misma VPC:



Esta conexión privada se establece mediante la creación de un punto final de VPC OpenSearch de interfaz gestionada por el servicio, con tecnología de. AWS PrivateLink Creamos una interfaz de red de punto de conexión en cada subred habilitada para el punto de conexión de VPC de interfaz. Se trata de interfaces de red administradas por el servicio que sirven como punto de entrada para el tráfico destinado al servicio. OpenSearch El precio estándar AWS PrivateLink de los puntos finales de interfaz se aplica a los puntos finales de VPC gestionados por el OpenSearch servicio facturados en virtud de. AWS PrivateLink

Puedes crear puntos de enlace de VPC para dominios que ejecuten todas las versiones de Elasticsearch OpenSearch y las versiones heredadas. Para obtener más información, consulte Acceso a los Servicios de AWS a través de AWS PrivateLink en la Guía de AWS PrivateLink .

Consideraciones y limitaciones del servicio OpenSearch

Antes de configurar un punto de enlace de VPC de interfaz para el OpenSearch servicio, consulte Acceder a un AWS servicio mediante un punto de enlace de VPC de interfaz en la Guía.AWS PrivateLink

Cuando utilice puntos finales de OpenSearch VPC gestionados por el servicio, tenga en cuenta lo siguiente:

- Solo puede utilizar puntos de conexión de VPC de interfaz para conectarse a los dominios de la VPC. Los dominios públicos no son compatibles.
- Los puntos de conexión de VPC solo pueden conectarse a dominios dentro de la misma Región de AWS.
- HTTPS es el único protocolo compatible para los puntos de conexión de VPC. No se permite el protocolo HTTP.
- OpenSearch El servicio permite realizar llamadas a todas las operaciones de OpenSearch API compatibles a través de un punto final de VPC de interfaz.
- Puede configurar un máximo de 50 puntos de conexión por cuenta y un máximo de 10 puntos de conexión por dominio. Un solo dominio puede tener un máximo de 10 entidades principales autorizadas.
- Actualmente no se pueden utilizar AWS CloudFormation para crear puntos finales de VPC de interfaz.
- Solo puede crear puntos finales de VPC de interfaz a través de la consola de OpenSearch servicio o mediante la OpenSearch API de servicio. No puede crear puntos de enlace de VPC de interfaz para el OpenSearch servicio mediante la consola de Amazon VPC.

- OpenSearch No se puede acceder a los puntos finales de VPC gestionados por el servicio desde Internet. Solo se puede acceder a un punto final de OpenSearch VPC administrado por el servicio dentro de la VPC en la que se aprovisiona el punto final o en cualquier enlace VPCs con la VPC en la que se aprovisiona el punto final, según lo permitan las tablas de enrutamiento y los grupos de seguridad.
- Las políticas de puntos de conexión de VPC no son compatibles con el servicio. OpenSearch Puede asociar un grupo de seguridad a las interfaces de red de los puntos finales para controlar el tráfico al OpenSearch Servicio a través del punto final de la VPC de la interfaz.
- Su función vinculada al servicio debe estar en la misma AWS cuenta que utilizó para crear el punto de enlace de la VPC.
- Para crear, actualizar y eliminar el punto de enlace de la VPC de OpenSearch servicio, debe tener los siguientes permisos de Amazon EC2 además de los permisos de Amazon OpenSearch Service:
  - `ec2:CreateVpcEndpoint`
  - `ec2:DescribeVpcEndpoints`
  - `ec2:ModifyVpcEndpoint`
  - `ec2:DeleteVpcEndpoints`
  - `ec2:CreateTags`
  - `ec2:DescribeTags`
  - `ec2:DescribeSubnets`
  - `ec2:DescribeSecurityGroups`
  - `ec2:DescribeVpcs`

> **Note**
>
> Actualmente, no puede limitar la creación de puntos finales de VPC al OpenSearch servicio. Estamos trabajando para que esto sea posible en una futura actualización.

Proporcionar el acceso a un dominio

Si la VPC a la que quieres acceder a tu dominio se encuentra en otro dominio Cuenta de AWS, debes autorizarla desde la cuenta del propietario antes de poder crear un punto final de VPC de interfaz.

Para permitir que una VPC de otra persona acceda Cuenta de AWS a tu dominio

1. Abre la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/casa/.
2. En el panel de navegación, seleccione Dominios y abra el dominio al que desea proporcionar acceso.
3. Ve a la pestaña de puntos finales de la VPC, donde se muestran las cuentas y las cuentas correspondientes VPCs que tienen acceso a tu dominio.
4. Seleccione Autorizar entidad principal.
5. Introduce el Cuenta de AWS ID de la cuenta que accederá a tu dominio. Este paso autoriza a la cuenta especificada a crear puntos de conexión de VPC en el dominio.
6. Seleccione Autorizar.

Creación de un punto de conexión de VPC de interfaz para un dominio de VPC

Puede crear un punto final de VPC de interfaz para el OpenSearch servicio mediante la consola de OpenSearch servicio o el AWS Command Line Interface ()AWS CLI.

Para crear un punto final de VPC de interfaz para un dominio de servicio OpenSearch

1. Abre la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/casa/.
2. En el panel de navegación izquierdo, seleccione Puntos de conexión de VPC.
3. Seleccione Crear punto de conexión.
4. Seleccione si desea conectar un dominio del dominio actual Cuenta de AWS o de otro Cuenta de AWS.
5. Seleccione el dominio al que se conecta con este punto de conexión. Si el dominio está en el actual Cuenta de AWS, usa el menú desplegable para elegir el dominio. Si el dominio se encuentra en otra cuenta, escriba el nombre de recurso de Amazon (ARN) del dominio para conectarse. Para elegir un dominio en una cuenta diferente, el propietario debe otorgarle acceso al dominio.
6. En el caso de la VPC, seleccione la VPC desde la que accederá al servicio. OpenSearch
7. En el caso de las subredes, seleccione una o más subredes desde las que accederá al Servicio. OpenSearch
8. En Grupos de seguridad, seleccione los grupos de seguridad que deban asociarse a las interfaces de red del punto de conexión. Este es un paso fundamental para limitar los puertos, los protocolos y los orígenes del tráfico entrante que autoriza para el punto de conexión. Las

reglas del grupo de seguridad deben permitir que los recursos que utilizarán el punto final de la VPC para comunicarse con el OpenSearch Servicio se comuniquen con la interfaz de red del punto final.

9. Seleccione Crear punto de conexión. El punto de conexión debería estar activo en un plazo de 2 a 5 minutos.

Trabajar con puntos finales de OpenSearch VPC gestionados por el servicio mediante la API de configuración

Utilice las siguientes operaciones de API para crear y administrar puntos de enlace de OpenSearch VPC gestionados por el servicio.

- CreateVpcEndpoint
- ListVpcEndpoints
- UpdateVpcEndpoint
- DeleteVpcEndpoint

Utilice las siguientes operaciones de la API para administrar el acceso de los puntos de conexión a los dominios de la VPC:

- AuthorizeVpcEndpointAccess
- ListVpcEndpointAccess
- ListVpcEndpointsForDomain
- RevokeVpcEndpointAccess

Autenticación SAML para paneles OpenSearch

La autenticación SAML para OpenSearch Dashboards te permite usar tu proveedor de identidad actual para ofrecer el inicio de sesión único (SSO) para los Dashboards de los dominios de Amazon OpenSearch Service que ejecuten OpenSearch Elasticsearch 6.7 o versiones posteriores. Para utilizar la autenticación SAML, debe habilitar el control de acceso detallado.

En lugar de autenticarse a través de Amazon Cognito o la base de datos de usuarios interna, la autenticación SAML OpenSearch para los paneles le permite utilizar proveedores de identidad de terceros para iniciar sesión en los paneles, administrar un control de acceso detallado, buscar sus

datos y crear visualizaciones. OpenSearch El servicio es compatible con los proveedores que utilizan el estándar SAML 2.0, como Okta, Keycloak, Active Directory Federation Services (ADFS), Auth0 y. AWS IAM Identity Center

La autenticación SAML para los paneles solo sirve para acceder a los paneles a través de un navegador web. OpenSearch Sus credenciales de SAML no le permiten realizar solicitudes HTTP directas a los paneles de control. OpenSearch APIs

Información general de la configuración de SAML

En esta documentación se supone que tiene un proveedor de identidad existente y que está familiarizado con él en cierta medida. No podemos proporcionar pasos de configuración detallados para tu proveedor exacto, solo para tu dominio de OpenSearch servicio.

El flujo de inicio de sesión en OpenSearch Dashboards puede adoptar una de estas dos formas:

- Proveedor de servicios (SP) iniciado: navegue al panel (por ejemplo, `https://my-domain.us-east-1.es.amazonaws.com/_dashboards`), que lo redirige a la pantalla de inicio de sesión. Después de iniciar sesión, el proveedor de identidades lo redirige al panel.
- Iniciado por el proveedor de identidad (IdP): navega hasta su proveedor de identidad, inicia sesión y elige OpenSearch Dashboards en un directorio de aplicaciones.

OpenSearch El servicio ofrece dos inicios de sesión únicos URLs, uno iniciado por SP y otro iniciado por IdP, pero solo necesitará el que coincida con el flujo de inicio de sesión deseado en los paneles de control. OpenSearch

Independientemente del tipo de autenticación que utilice, el objetivo es iniciar sesión a través de su proveedor de identidades y recibir una aserción SAML que contenga su nombre de usuario (requerido) y cualquier rol backend (opcional, pero recomendado). Esta información permite el control de acceso detallado para asignar permisos a usuarios de SAML. En los proveedores de identidad externos, los roles backend se denominan generalmente “roles” o “grupos”.

Consideraciones

Tenga en cuenta lo siguiente cuando configure la autenticación SAML:

- Debido al tamaño del archivo de metadatos del IdP, se recomienda utilizar la consola de AWS para configurar la autenticación SAML.

- Los dominios solo admiten un método de autenticación del panel a la vez. Si tiene habilitada la autenticación de Amazon Cognito para OpenSearch paneles, debe deshabilitarla antes de poder habilitar la autenticación SAML.
- Si utiliza un equilibrador de carga de red con SAML, primero debe crear un punto de conexión personalizado. Para obtener más información, consulte ???.
- Las políticas de control de servicios (SCP) no se aplicarán ni evaluarán en el caso de identidades que no sean de IAM (como SAML en OpenSearch Amazon Serverless y SAML y la autorización básica de usuario interna para Amazon Service). OpenSearch

Autenticación SAML para dominios de VPC

SAML no requiere comunicación directa entre el proveedor de identidades y el proveedor de servicios. Por lo tanto, aunque tu OpenSearch dominio esté alojado en una VPC privada, puedes seguir usando SAML siempre que tu navegador pueda comunicarse con tu OpenSearch clúster y con tu proveedor de identidad. En esencia, el navegador actúa como intermediario entre el proveedor de identidades y el proveedor de servicios. Para ver un útil diagrama que explica el flujo de autenticación SAML, consulte la documentación de Okta.

Modificación de la política de acceso al dominio

Antes de configurar la autenticación SAML, debe actualizar la política de acceso al dominio para permitir a los usuarios de SAML acceder al dominio. De lo contrario, aparecerán errores de acceso denegado.

Recomendamos la siguiente política de acceso al dominio, que proporciona acceso completo a los subrecursos (`/*`) en el 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/*"
    }
  ]
}
```

Para hacer que la política sea más restrictiva, puede agregar una condición de dirección IP a la política. Esta condición limita el acceso únicamente a la subred o rango de direcciones IP especificados. Por ejemplo, la siguiente política solo permite el acceso desde la subred 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 política de acceso a un dominio abierto requiere que se habilite un control de acceso detallado en el dominio; de lo contrario, aparece el siguiente error:

`To protect domains with public access, a restrictive policy or fine-grained access control is required.`
Si tiene un usuario maestro o un usuario interno configurado con una contraseña segura, mantener la política abierta mientras usa un control de acceso detallado podría ser aceptable desde el punto de vista de la seguridad. Para obtener más información, consulte ???.

Configuración de la autenticación iniciada por proveedor de servicios o por proveedor de identidades

En estos pasos se explica cómo habilitar la autenticación SAML con la autenticación iniciada por el SP o por el IdP para los paneles de control. OpenSearch Para conocer el paso adicional necesario para habilitar ambas, consulte Configuración de la autenticación iniciada por proveedor de servicios o por proveedor de identidades.

Paso 1: habilitar la autenticación SAML

Puede habilitar la autenticación SAML bien durante la creación del dominio o bien seleccionando Acciones, Editar la configuración de seguridad en un dominio existente. Los siguientes pasos varían ligeramente según lo que seleccione.

En la configuración del dominio, en Autenticación SAML para OpenSearch Dashboards/Kibana, selecciona Habilitar la autenticación SAML.

Paso 2: configurar el proveedor de identidades

Ejecute los siguientes pasos en función del momento de configuración de la autenticación SAML.

Si va a crear un nuevo dominio

Si estás creando un dominio nuevo, el OpenSearch servicio aún no puede generar un identificador de entidad o un SSO del proveedor de servicios. URLs El proveedor de identidades requiere estos valores para habilitar correctamente la autenticación SAML, pero solo se pueden generar después de crear el dominio. Para evitar esta interdependencia durante la creación del dominio, puede proporcionar valores temporales en la configuración del IdP con el fin de generar los metadatos necesarios, y luego actualizarlos una vez que el dominio esté activo.

Si utilizas un punto final personalizado, puedes deducir cuál URLs será. Por ejemplo, si el punto de conexión personalizado es `www.custom-endpoint.com`, el ID de entidad del proveedor de

servicios seráwww.*custom-endpoint*.com, la URL de SSO iniciado por IdP será www.*custom-endpoint*.com/_dashboards/_opendistro/_security/saml/acs/idpinitiated, y la URL de SSO iniciado por SP será www.*custom-endpoint*.com/_dashboards/_opendistro/_security/saml/acs. Puede utilizar los valores para configurar el proveedor de identidades antes de crear el dominio. Consulte la siguiente sección para ver ejemplos.

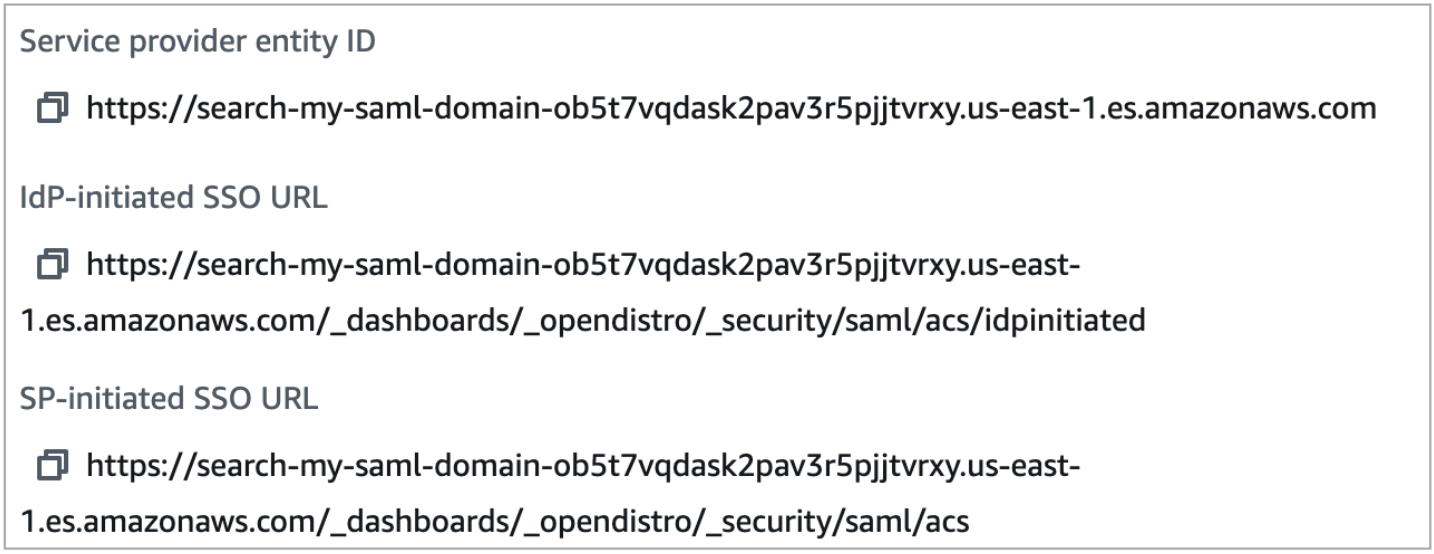
> **Note**
>
> No puede iniciar sesión con un punto de conexión de doble pila porque el FQDN de una solicitud HTTP es diferente del FQDN de una solicitud SAML. Un OpenSearch administrador tendrá que configurar un punto final personalizado y establecer el valor de CNAME en un punto final de doble pila si quieres iniciar sesión con un punto final de doble pila.

Si no utiliza un punto de conexión personalizado, puede introducir valores temporales en el IdP para generar los metadatos necesarios, y luego actualizarlos una vez que el dominio esté activo.

Por ejemplo, en Okta puede introducir https://*temp-endpoint*.amazonaws.com en los campos URL de inicio de sesión único y URI de audiencia (ID de entidad del SP), lo que permite generar los metadatos. Después, cuando el dominio esté activo, podrás recuperar los valores correctos de OpenSearch Service y actualizarlos en Okta. Para obtener instrucciones, consulte the section called “Paso 6: Actualiza tu IdP URLs”.

Si va a editar un dominio existente

Si habilitas la autenticación SAML en un dominio existente, copia el ID de la entidad del proveedor de servicios y uno de los SSO. URLs Para obtener orientación sobre qué URL debe utilizar, consulte the section called “Información general de la configuración de SAML”.



Utilice los valores para configurar el proveedor de identidades. Esta es la parte más compleja del proceso, y, desafortunadamente, la terminología y los pasos varían enormemente según el proveedor. Consulte la documentación de su proveedor.

En Okta, por ejemplo, crea una aplicación web SAML 2.0. En URL de inicio de sesión único, especifique la URL de SSO. Para URI de audiencia (ID de identidad del SP), especifique el ID de entidad del SP.

En lugar de usuarios y roles de backend, Okta tiene usuarios y grupos. En Instrucciones de atributo de grupo, se recomienda agregar `role` al campo Nombre y la expresión regular `.+` al campo Filtro. Esta instrucción indica al proveedor de identidades de Okta que incluya todos los grupos de usuarios bajo el campo `role` de la aserción SAML después de que un usuario se autentica.

En el Centro de identidades de IAM, especifique el ID de la entidad del SP como la audiencia de SAML de la aplicación. También debe especificar las siguientes asignaciones de atributos: `Subject=${user:subject}:format=unspecified` y `Role=${user:groups}:format=uri`.

En Auth0, crea una aplicación web regular y habilita el complemento SAML 2.0. En Keycloak, crea un cliente.

Paso 3: Importar metadatos del IdP

Después de configurar el proveedor de identidades, genera un archivo de metadatos de IdP. Este archivo XML contiene información sobre el proveedor, como un certificado TLS, puntos de conexión de inicio de sesión único y el ID de entidad del proveedor de identidad.

Copie el contenido del archivo de metadatos del IdP y péguelo en el campo Metadatos del IdP de la consola de servicio. OpenSearch Alternativamente, seleccione Importar desde archivo XML y cargue el archivo. El archivo de metadatos debe tener un aspecto similar al siguiente:

```
<?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>
```

Paso 4: configurar los campos SAML

Después de introducir los metadatos del IdP, configure los siguientes campos adicionales en la consola de OpenSearch servicio:

- ID de entidad del IdP: copie el valor de la propiedad `entityID` del archivo de metadatos y péguelo en este campo. Muchos proveedores de identidades también muestran este valor como parte de un resumen posterior a la configuración. Algunos proveedores lo llaman el “emisor”.
- Nombre de usuario maestro de SAML y función de backend maestro de SAML: la función de and/or backend de usuario que especifique recibe todos los permisos para el clúster, lo que equivale a un nuevo usuario maestro, pero solo puede usar esos permisos en los paneles. OpenSearch

  Por ejemplo, en Okta, es posible que tenga un usuario `jdoe` que pertenece al grupo `admins`. Si agrega `jdoe` al nombre de usuario maestro de SAML, solo ese usuario recibe permisos

completos. Si agrega `admins` Rol de backend maestro de SAML, cualquier usuario que pertenezca al grupo `admins` recibe permisos completos.

> **Note**
>
> El contenido de la aserción SAML debe coincidir exactamente con las cadenas que se utilicen para el nombre de usuario maestro de SAML y el rol maestro de SAML. Algunos proveedores de identidad añaden un prefijo antes de sus nombres de usuario, lo que puede provocar que no coincidan. hard-to-diagnose En la interfaz de usuario del proveedor de identidades, es posible que vea `jdoe`, pero la aserción SAML podría contener `auth0|jdoe`. Utilice siempre la cadena de la aserción SAML.

Muchos proveedores de identidades permiten ver una aserción de muestra durante el proceso de configuración, y herramientas como Rastreo SAML pueden ayudarlo a examinar y solucionar problemas del contenido de las aserciones reales. Las aserciones se ven algo similar a lo siguiente:

```
<?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>
```

Paso 5: (opcional) configurar ajustes adicionales

En Configuración adicional, configure los siguientes campos opcionales:

- Clave de asunto: puede dejar este campo vacío con el fin de utilizar el elemento `NameID` de la aserción SAML para el nombre de usuario. Si su aserción no utiliza este elemento estándar y, en su lugar, incluye el nombre de usuario como un atributo personalizado, especifique ese atributo aquí.
- Clave de roles: si desea utilizar roles de backend (se recomienda), especifique un atributo de la aserción en este campo; por ejemplo, `role` o `group`. Esta es otra situación en la que herramientas como Rastreo SAML pueden ayudar.
- Duración de la sesión: de forma predeterminada, OpenSearch Dashboards cierra la sesión de los usuarios después de 24 horas. Puede establecer este valor en cualquier número comprendido entre 60 y 1440 (24 horas) especificando un nuevo valor.

Cuando la configuración le parezca adecuada, guarde el dominio.

Paso 6: Actualiza tu IdP URLs

Si habilitaste la autenticación SAML al crear un dominio, tenías que especificar un valor temporal URLs en tu IdP para poder generar el archivo de metadatos XML. Una vez que el estado del dominio cambie a`Active`, podrás obtener el URLs IdP correcto y modificarlo.

Para recuperarlo URLs, selecciona el dominio y elige Acciones, editar la configuración de seguridad. En la sección Autenticación SAML para OpenSearch Dashboards/Kibana, encontrarás el identificador de entidad del proveedor de servicios y el SSO correctos. URLs Copia los valores y úsalos para configurar tu proveedor de identidad, sustituyendo el temporal URLs que proporcionaste en el paso 2.

Paso 7: Asignar usuarios de SAML a roles

Una vez que el estado de su dominio sea Activo y su IDP esté configurado correctamente, navegue hasta los OpenSearch paneles.

- Si eligió la dirección URL iniciada por el SP, diríjase a *domain-endpoint*/_dashboards. Para iniciar sesión directamente en un inquilino específico, puede agregar ?security_tenant=*tenant-name* a la URL.
- Si eligió la dirección URL iniciada por el IdP, diríjase al directorio de aplicaciones de su proveedor de identidad.

En ambos casos, inicie sesión como usuario maestro SAML o como usuario que pertenece al rol de backend maestro SAML. Para continuar con el ejemplo del paso 7, inicie sesión como jdoe o un miembro del grupo admins.

Una vez que se cargue OpenSearch el panel de mandos, selecciona Seguridad y roles. A continuación, asigne los roles para permitir que otros usuarios accedan a los OpenSearch paneles.

Por ejemplo, podría asignar a su colega de confianza jroe a los roles all_access y security_manager. También puede asignar el rol de backend analysts a los roles readall y opensearch_dashboards_user.

Si prefiere usar la API en lugar de los OpenSearch paneles, consulte el siguiente ejemplo de solicitud:

```
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"] }
  }
]
```

Configuración de la autenticación iniciada por proveedor de servicios y por proveedor de identidades

Si desea configurar la autenticación iniciada por SP e IdP, debe hacerlo a través de su proveedor de identidades. Por ejemplo, en Okta, puede seguir estos pasos:

1. Dentro de su aplicación SAML, vaya a General, Configuración SAML.
2. Para la URL de inicio de sesión único, proporcione URL SSO iniciado por IdP. Por ejemplo, `https://search-domain-hash/_dashboards/_opendistro/_security/saml/acs/idpinitiated`.
3. Activa Permitir que esta aplicación solicite otro URLs SSO.
4. En SSO solicitable URLs, agrega uno o más SSO iniciados por el SP. URLs Por ejemplo, `https://search-domain-hash/_dashboards/_opendistro/_security/saml/acs`.

Configuración de la autenticación SAML (AWS CLI)

El siguiente AWS CLI comando habilita la autenticación SAML para los OpenSearch paneles de control de un dominio existente:

```
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"}}'
```

Debe escapar de todas las comillas y caracteres de nueva línea en el XML de metadatos. Por ejemplo, utilice `<KeyDescriptor use=\"signing\">\n` en lugar de `<KeyDescriptor`

use="signing"> y un salto de línea. Para obtener información detallada sobre la utilización de la AWS CLI, consulte la Referencia de los comandos de la AWS CLI.

Configuración de la autenticación SAML (API de configuración)

La siguiente solicitud a la API de configuración habilita la autenticación SAML para los OpenSearch paneles de control de un dominio existente:

```
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"
    }
  }
}
```

Debe escapar de todas las comillas y caracteres de nueva línea en el XML de metadatos. Por ejemplo, utilice <KeyDescriptor use=\"signing\">\n en lugar de <KeyDescriptor use="signing"> y un salto de línea. Para obtener información detallada sobre el uso de la API de configuración, consulta la referencia de la API OpenSearch de servicio.

Solución de problemas de SAML

Error	Details
Tu solicitud: '*/some/path* ' no está permitida.	Compruebe que ha proporcionado la dirección URL de SSO correcta (paso 3) a su proveedor de identidad .

Error	Details
Proporcione un documento de metadatos del proveedor de identidad válido para habilitar SAML.	El archivo de metadatos del IdP no cumple con el estándar SAML 2.0. Verifique si hay errores mediante una herramienta de validación.
Las opciones de configuración de SAML no son visibles en la consola.	Actualice a la versión más reciente del software de servicio.
Error de configuración de SAML: se produjo un error al recuperar la configuración de SAML. Verifique su configuración.	Este error genérico puede producirse por muchas razones. • Compruebe que ha proporcionado a su proveedor de identidad el ID de entidad del SP y la dirección URL de SSO correctos. • Vuelva a generar el archivo de metadatos del IdP y verifique el ID de entidad de IdP. Agregue todos los metadatos actualizados en la consola de AWS . • Verifique que la política de acceso a su dominio permita el acceso a los OpenSearch paneles y `_plugins/_security/*` . En general, recomendamos una política de acceso abierto para dominios que utilicen un control de acceso detallado. • Consulte la documentación de su proveedor de identidades a fin de conocer los pasos necesarios para configurar SAML.

Error	Details
Rol faltante: no hay roles disponibles para este usuario. Póngase en contacto con el administrador del sistema.	Se ha autenticado correctamente, pero el nombre de usuario y los roles de backend de la aserción SAML no se asignan a ningún rol y, por lo tanto, no tienen permisos. Estos mapeos distinguen entre mayúsculas y minúsculas. El administrador del sistema puede verificar el contenido de su aserción SAML con una herramienta como SAML-tracer y comprobar la asignación de roles con la siguiente solicitud: `GET _plugins/_security/api/rolesmapping`
Su navegador redirige o recibe continuamente errores HTTP 500 al intentar acceder OpenSearch a los paneles.	Estos errores pueden producirse si la aserción SAML contiene un gran número de roles con un total aproximado de 1500 caracteres. Por ejemplo, si transfiere 80 roles, cuya longitud media es de 20 caracteres, es posible que supere el límite de tamaño de las cookies en su navegador web. A partir de OpenSearch la versión 2.7, la aserción SAML admite funciones de hasta 5000 caracteres.
No puede cerrar sesión en ADFS.	ADFS requiere que todas las solicitudes de cierre de sesión estén firmadas, algo que el OpenSearch servicio no admite. Elimine `<SingleLogoutService />` del archivo de metadatos del IdP para obligar al OpenSearch Servicio a utilizar su propio mecanismo de cierre de sesión interno.

Error	Details
`Could not find entity descriptor for __PATH__.`	El ID de entidad del IdP proporcionado en el XML de metadatos al OpenSearch servicio es diferente al de la respuesta de SAML. Para ello, asegúrese de que coincidan. Active los registros de errores de aplicacio nes de CW en su dominio para encontrar el mensaje de error y depurar el problema de integración con SAML.
`Signature validation failed. SAML response rejected.`	OpenSearch El servicio no puede verificar la firma de la respuesta de SAML mediante el certificado del IdP proporcionado en el XML de metadatos. Puede deberse a un error manual o a que su IdP haya rotado su certificado. Actualice el certificado más reciente de su IdP en el XML de metadatos proporcionado al OpenSearch Servicio a través del. Consola de administración de AWS
`__PATH__ is not a valid audience for this response.`	El campo de audiencia de la respuesta de SAML no coincide con el punto de conexión del dominio. Para corregir este error, actualice el campo de audiencia del SP para que coincida con el punto de conexión de su dominio. Si ha activado los puntos de conexión personalizados, el campo de audiencia debe coincidir con su punto de conexión personalizado. Active los registros de errores de aplicaciones de CW en su dominio para encontrar el mensaje de error y depurar el problema de integración con SAML.
Su navegador recibe un mensaje de error HTTP 400 con `Invalid Request Id` en la respuesta.	Este error suele producirse si ha configurado la URL iniciada por el IdP con este formato *<DashboardsURL>* `/_opendistro/_security/saml/acs`. En su lugar, configure la URL con el formato *<DashboardsURL>* `/_opendistro/_security/saml/acs/idpinitiated` .

Error	Details
La respuesta se recibió en `__PATH__` en vez de `__PATH__`.	El campo de destino de la respuesta SAML no coincide con ninguno de los siguientes formatos de URL: • *<DashboardsURL>* `/_opendistro/_security/saml/acs` • *<DashboardsURL>* `/_opendistro/_security/saml/acs/idpinitiated` . Según el flujo de inicio de sesión que utilice (iniciado por SP o iniciado por IDP), introduzca un campo de destino que coincida con uno de los. OpenSearch URLs
La respuesta tiene un atributo `InResponseTo` , pero no se esperaba `InResponseTo` .	Está utilizando la URL iniciada por IdP para un flujo de inicio de sesión iniciado por SP. En su lugar, utilice la URL iniciada por SP.

Deshabilitar la autenticación SAML

Para deshabilitar la autenticación SAML en los paneles de control (consola) OpenSearch

1. Seleccione el dominio, Acciones y Editar la configuración de seguridad.
2. Desmarque Habilitar la autenticación SAML.
3. Seleccione Guardar cambios.
4. Después de que el dominio termine de procesar, compruebe el mapeo de roles de control de acceso detallado con la siguiente solicitud:

   ```
   GET _plugins/_security/api/rolesmapping
   ```

   Al deshabilitar la autenticación SAML para los paneles de control, no se eliminan las asignaciones del nombre de usuario maestro de SAML, la función de backend del maestro de SAML. and/or Si desea eliminar estos mapeos, inicie sesión en Dashboards con la base de datos de usuario interna (si está habilitada) o utilice la API para eliminarlos:

```
PUT _plugins/_security/api/rolesmapping/all_access
{
  "users": [
    "master-user"
  ]
}
```

IAM Identity Center Trusted Identity Propagation Support para OpenSearch

Ahora puede utilizar los elementos principales del Centro de Identidad de AWS IAM (usuarios y grupos) configurados de forma centralizada mediante Trusted Identity Propagation para acceder a los dominios de Amazon OpenSearch Service a través de aplicaciones de OpenSearch servicio. Para habilitar el soporte del Centro de Identidad de IAM OpenSearch, necesitará habilitar el uso del Centro de Identidad de IAM. Para obtener más información sobre cómo hacerlo, consulte ¿Qué es el Centro de Identidad de IAM? . Consulte ¿Cómo asociar un OpenSearch dominio como fuente de datos en OpenSearch las aplicaciones? para obtener más información.

Puede configurar el Centro de identidades de IAM mediante la consola de OpenSearch servicio, el AWS Command Line Interface (AWS CLI) o el AWS SDKs.

> **Note**
>
> Dashboards (ubicado junto con el clúster) no admite las entidades principales del Centro de identidades de IAM. Solo se admiten a través de la interfaz de OpenSearch usuario centralizada (paneles de control).

Consideraciones

Antes de utilizar el Centro de identidades de IAM con Amazon OpenSearch Service, debe tener en cuenta lo siguiente:

- El Centro de identidades de IAM debe estar habilitado en la cuenta.
- El Centro de identidades de IAM debe estar disponible en su región.
- La versión del OpenSearch dominio es 1.3 o posterior.

- El Control de acceso detallado debe estar habilitado en el dominio.
- El dominio debe estar en la misma región que la instancia del Centro de identidades de IAM.
- El dominio y OpenSearch la aplicación pertenecen a la misma AWS cuenta.

Modificación de la política de acceso al dominio

Antes de configurar el Centro de identidades de IAM, debe actualizar la política de acceso al dominio o los permisos de la función de IAM configurada en OpenSearch las aplicaciones de 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/
*"
        }
    ]
}
```

Configuración de la autenticación y autorización de Centro de identidades de IAM (consola)

Puede habilitar la autenticación y la autorización del Centro de identidades de IAM durante el proceso de creación del dominio o mediante la actualización de un dominio existente. Los pasos de configuración varían ligeramente según la opción que seleccione.

En los siguientes pasos se explica cómo configurar un dominio existente para la autenticación y autorización del Centro de Identidad de IAM en la consola de Amazon OpenSearch Service:

1. En Configuración del dominio, vaya a Configuración de seguridad, seleccione Editar, vaya a la sección de autenticación del Centro de identidades de IAM y seleccione Habilitar el acceso a la API autenticado con el Centro de identidades de IAM.
2. Seleccione la clave SubjectKey y Roles de la siguiente manera.

   - Clave de asunto: elija una de las siguientes opciones UserId (predeterminada) UserName y Correo electrónico para utilizar el atributo correspondiente como principal acceso al dominio.
   - Clave de funciones: elija una de las funciones GroupId (predeterminada) y GroupName utilice los valores de atributo correspondientes como función de back-end fine-grained-access-controlpara todos los grupos asociados al centro de identidad principal de IAM.

Una vez hechos los cambios, guarde el dominio.

Configuración del control de acceso detallado

Una vez que haya activado la opción del Centro de Identidad de IAM en su OpenSearch dominio, podrá configurar el acceso a los directores del Centro de Identidad de IAM mediante la creación de una asignación de funciones a la función de back-end. El valor de la función de back-end del director se basa en la pertenencia al grupo del director del IAM Identity Center y en la configuración de o. RolesKey GroupId GroupName

> **Note**
>
> Amazon OpenSearch Service puede admitir hasta 100 grupos para un solo usuario. Si intenta utilizar más instancias de las permitidas, experimentará una incoherencia en el procesamiento de su fine-grained-access-control autorización y recibirá un mensaje de error 403.

Configuración de la autenticación y autorización del Centro de identidades de IAM (CLI)

```
  aws opensearch update-domain-config \
      --domain-name my-domain \
      --identity-center-options '{"EnabledAPIAccess": true,
 "IdentityCenterInstanceARN": "instance arn",  "SubjectKey": "UserId/UserName/
UserEmail" , "RolesKey": "GroupId/GroupName"}'
```

Deshabilitación de la autenticación del Centro de identidades de IAM en el dominio

Para inhabilitar el Centro de identidades de IAM en tu dominio, sigue estos pasos: OpenSearch

1. Seleccione el dominio, Acciones y Editar la configuración de seguridad.
2. Desmarque Habilitar el acceso a la API autenticado con el Centro de identidades de IAM.
3. Seleccione Save changes (Guardar cambios).
4. Cuando el dominio termine de procesarse, elimine las asignaciones de funciones agregadas para los directores del IAM Identity Center

Para deshabilitar el Centro de identidades de IAM a través de la CLI, puede utilizar lo siguiente:

```
aws opensearch update-domain-config \
    --domain-name my-domain \
    --identity-center-options '{"EnabledAPIAccess": false}'
```

Configuración de la autenticación de Amazon Cognito para paneles OpenSearch

Puede autenticar y proteger su instalación predeterminada de OpenSearch paneles de Amazon OpenSearch Service mediante Amazon Cognito. La autenticación de Amazon Cognito es opcional y solo está disponible para los dominios que utilizan Elasticsearch OpenSearch 5.1 o una versión posterior. Si no configura la autenticación de Amazon Cognito, puede seguir protegiendo Dashboards con una política de acceso basada en IP y un servidor proxy, una autenticación básica de HTTP o SAML.

Gran parte del proceso de autenticación se lleva a cabo en Amazon Cognito, pero en esta sección se ofrecen directrices y requisitos para configurar los recursos de Amazon Cognito para que funcionen OpenSearch con los dominios de servicio. Se aplican precios estándar a todos los recursos de Amazon Cognito.

> **Tip**
>
> La primera vez que configure un dominio para usar la autenticación de Amazon Cognito para los OpenSearch paneles, le recomendamos que utilice la consola. Los recursos de Amazon Cognito se pueden personalizar en gran medida y la consola puede ayudar a identificar y comprender las funciones importantes.

Temas

- Requisitos previos
- Configuración de un dominio para utilizar la autenticación de Amazon Cognito
- Permitir el rol autenticado
- Configuración de proveedores de identidades
- (Opcional) Configuración de acceso pormenorizado
- (Opcional) Personalización de la página de inicio de sesión
- (Opcional) Configuración de seguridad avanzada
- Testeo
- Cuotas
- Problemas habituales de configuración
- Desactivación de la autenticación de Amazon Cognito para paneles OpenSearch
- Eliminar dominios que utilizan la autenticación de Amazon Cognito para los paneles OpenSearch

Requisitos previos

Antes de poder configurar la autenticación de Amazon Cognito para los OpenSearch paneles, debe cumplir varios requisitos previos. La consola OpenSearch de servicio ayuda a agilizar la creación de estos recursos, pero comprender el propósito de cada recurso ayuda a configurar y solucionar problemas. La autenticación de Amazon Cognito para Dashboards requiere los siguientes recursos:

- Grupo de usuarios de Amazon Cognito
- Grupo de identidades de Amazon Cognito
- Rol de IAM que tiene la política de `AmazonOpenSearchServiceCognitoAccess` adjunta (`CognitoAccessForAmazonOpenSearch`)

> **Note**
>
> El grupo de usuarios y el grupo de identidades deben estar en la misma Región de AWS. Puede usar el mismo grupo de usuarios, grupo de identidades y rol de IAM para añadir la autenticación de Amazon Cognito para Dashboards a OpenSearch varios dominios de servicio. Para obtener más información, consulte the section called “Cuotas”.

Acerca del grupo de usuarios

Los grupos de usuarios tiene dos características principales: crear y administrar un directorio de usuarios y permitir que los usuarios se inscriban e inicien sesión. Para obtener instrucciones acerca de cómo crear un grupo de usuarios, consulte Introducción a los grupos de usuarios en la Guía para desarrolladores de Amazon Cognito.

Al crear un grupo de usuarios para usarlo con el OpenSearch Servicio, tenga en cuenta lo siguiente:

- Su grupo de usuarios de Amazon Cognito debe tener un nombre de dominio. OpenSearch El servicio utiliza este nombre de dominio para redirigir a los usuarios a una página de inicio de sesión para acceder a los paneles. Aparte de un nombre de dominio, el grupo de usuarios no requiere ninguna configuración no predeterminada.
- Debe especificar los atributos estándar necesarios del grupo como nombre, fecha de nacimiento, dirección de email y número de teléfono. No se pueden cambiar estos atributos después de crear el grupo de usuarios, por lo que debe elegir los que considere importantes en este momento.
- Al crear el grupo de usuarios, elija si los usuarios pueden crear sus propias cuentas, el nivel mínimo de seguridad de la contraseña para las cuentas y si se va a habilitar la autenticación multifactor. Si tiene previsto utilizar un proveedor de identidades externo, estos ajustes no tienen importancia. Técnicamente puede habilitar el grupo de usuarios como un proveedor de identidades y habilitar un proveedor de identidades externo, pero la mayoría de las personas prefieren uno u otro.

El grupo de usuarios IDs adopta la forma de`region_ID`. Si planea usar la AWS CLI o un AWS SDK para configurar el OpenSearch servicio, anote el ID.

Acerca del grupo de identidades

Los grupos de identidades permiten asignar funciones temporales con privilegios limitados a los usuarios después de que inicien sesión. Para obtener instrucciones acerca de cómo crear un grupo

de identidades, consulte Información general de la consola de grupos de identidades en la Guía para desarrolladores de Amazon Cognito. Al crear un grupo de identidades para usarlo con el OpenSearch Servicio, tenga en cuenta lo siguiente:

- Si utiliza la consola de Amazon Cognito, debe seleccionar la casilla de verificación Habilitar el acceso a identidades sin autenticar para crear el grupo de identidades. Tras crear el grupo de identidades y configurar el dominio del OpenSearch servicio, Amazon Cognito inhabilita este ajuste.
- No es necesario agregar proveedores de identidades externos al grupo de identidades. Cuando configura el OpenSearch Servicio para usar la autenticación de Amazon Cognito, configura el grupo de identidades para que use el grupo de usuarios que acaba de crear.
- Después de crear el grupo de identidades, debe elegir funciones de IAM autenticadas y sin autenticar. Estos roles especifican las políticas de acceso que los usuarios tienen antes y después de que inicien sesión. Si utiliza la consola de Amazon Cognito, puede crear estas funciones para usted. Una vez creado el rol autenticado, anote el ARN, que adopta la forma de `arn:aws:iam::123456789012:role/Cognito_identitypoolnameAuth_Role`.

El grupo de identidades IDs adopta la forma de. *region*`:`*ID*`-`*ID*`-`*ID*`-`*ID*`-`*ID* Si planea usar la AWS CLI o un AWS SDK para configurar el OpenSearch servicio, anote el ID.

Acerca del CognitoAccessForAmazonOpenSearch rol

OpenSearch El servicio necesita permisos para configurar los grupos de usuarios e identidades de Amazon Cognito y usarlos para la autenticación. Para ello`AmazonOpenSearchServiceCognitoAccess`, puede utilizar una política AWS gestionada. `AmazonESCognitoAccess`es una política antigua que se sustituyó por la que `AmazonOpenSearchServiceCognitoAccess` se cambió el nombre del servicio a Amazon OpenSearch Service. Ambas políticas proporcionan los permisos mínimos de Amazon Cognito necesarios para habilitar la autenticación de Amazon Cognito. Para obtener más información sobre la política, consulta AmazonOpenSearchServiceCognitoAccessla Guía de referencia de políticas AWS gestionadas.

Si utilizas la consola para crear o configurar tu dominio de OpenSearch servicio, esta crea un rol de IAM para ti y adjunta la `AmazonOpenSearchServiceCognitoAccess` política (o la `AmazonESCognitoAccess` política si es un dominio de Elasticsearch) al rol. El nombre predeterminado del rol es `CognitoAccessForAmazonOpenSearch`.

`AmazonESCognitoAccess`Tanto la función como las políticas `AmazonOpenSearchServiceCognitoAccess` de permisos permiten al OpenSearch Servicio realizar las siguientes acciones en todos los grupos de identidades y usuarios:

- Acción: `cognito-idp:DescribeUserPool`
- Acción: `cognito-idp:CreateUserPoolClient`
- Acción: `cognito-idp:DeleteUserPoolClient`
- Acción: `cognito-idp:UpdateUserPoolClient`
- Acción: `cognito-idp:DescribeUserPoolClient`
- Acción: `cognito-idp:AdminInitiateAuth`
- Acción: `cognito-idp:AdminUserGlobalSignOut`
- Acción: `cognito-idp:ListUserPoolClients`
- Acción: `cognito-identity:DescribeIdentityPool`
- Acción: `cognito-identity:SetIdentityPoolRoles`
- Acción: `cognito-identity:GetIdentityPoolRoles`

Si usa uno de ellos AWS SDKs, debe crear su propio rol, adjuntar la política y especificar el ARN de este rol al configurar su dominio de OpenSearch servicio. AWS CLI El rol debe tener la siguiente relación de confianza:

JSON

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

Para obtener instrucciones, consulte Crear un rol para delegar permisos a un AWS servicio y Agregar y eliminar permisos de identidad de IAM en la Guía del usuario de IAM.

Configuración de un dominio para utilizar la autenticación de Amazon Cognito

Tras completar los requisitos previos, puede configurar un dominio de OpenSearch servicio para usar Amazon Cognito for Dashboards.

> **Note**
>
> Amazon Cognito no está disponible en todos los dispositivos. Regiones de AWS Para ver una lista de las regiones admitidas, consulte Puntos de conexión del servicio de Amazon Cognito. No necesita usar la misma región para Amazon Cognito que usa para OpenSearch el Servicio.

Configuración de la autenticación de Amazon Cognito (consola)

Dado que se crea el rol `CognitoAccessForAmazonOpenSearch` para que usted no tenga que hacerlo, la consola ofrece la experiencia de configuración más sencilla. Además de los permisos de OpenSearch servicio estándar, necesita el siguiente conjunto de permisos para usar la consola y crear un dominio que utilice la autenticación de Amazon Cognito para OpenSearch Dashboards.

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"
    }
  ]
}
```

Para obtener instrucciones sobre cómo agregar permisos a una identidad (usuario, grupo de usuarios o rol), consulteAdición de permisos de identidad de IAM (consola).

Si el `CognitoAccessForAmazonOpenSearch` ya existe, necesita menos permisos:

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"
    }
  ]
}
```

Para configurar la autenticación de Amazon Cognito para Dashboards (consola)

1. Abre la consola de Amazon OpenSearch Service en https://console.aws.amazon.com/aos/casa/.
2. En Dominios, seleccione el dominio que desea configurar.
3. Seleccione Acciones y Editar la configuración de seguridad.
4. Seleccione Habilitar la autenticación de Amazon Cognito.
5. En Región, seleccione la Región de AWS que contenga su grupo de usuarios y su grupo de identidades de Amazon Cognito.
6. En Grupo de usuarios de Cognito, seleccione un grupo de usuarios o cree uno. Para obtener más información, consulte the section called “Acerca del grupo de usuarios”.
7. En Grupo de identidades de Cognito, seleccione un grupo de identidades o cree uno. Para obtener más información, consulte the section called “Acerca del grupo de identidades”.

   > **Note**
   >
   > Los enlaces Crear grupo de usuarios y Crear grupo de identidades dirigen a la consola de Amazon Cognito y requieren crear estos recursos de forma manual. El proceso no es automático. Para obtener más información, consulte the section called “Requisitos previos”.

8. Para Nombre de rol de IAM, utilice el valor de `CognitoAccessForAmazonOpenSearch` predeterminado (recomendado) o ingrese un nombre nuevo. Para obtener más información, consulte the section called “Acerca del CognitoAccessForAmazonOpenSearch rol”.
9. Elija Guardar cambios.

Después de que el dominio termine de procesarse, consulte the section called “Permitir el rol autenticado” y the section called “Configuración de proveedores de identidades” para conocer los pasos adicionales de configuración.

Configuración de la autenticación de Amazon Cognito (AWS CLI)

Utilice el `--cognito-options` parámetro para configurar su dominio OpenSearch de servicio. Los comandos `create-domain` y `update-domain-config` emplean la siguiente sintaxis:

```
--cognito-options Enabled=true,UserPoolId="user-pool-id",IdentityPoolId="identity-pool-id",RoleArn="arn:aws:iam::123456789012:role/CognitoAccessForAmazonOpenSearch"
```

Ejemplo

En el siguiente ejemplo se crea un dominio en la región `us-east-1` que habilita la autenticación de Amazon Cognito para Dashboards mediante el rol `CognitoAccessForAmazonOpenSearch` y proporciona acceso al dominio a `Cognito_Auth_Role`:

```
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"
```

Después de que el dominio termine de procesarse, consulte the section called “Permitir el rol autenticado” y the section called “Configuración de proveedores de identidades” para conocer los pasos adicionales de configuración.

Configuración de la autenticación de Amazon Cognito ()AWS SDKs

Las AWS SDKs (excepto Android e iOS SDKs) admiten todas las operaciones que se definen en la referencia de la API de Amazon OpenSearch Service, incluido el `CognitoOptions` parámetro de las `UpdateDomainConfig` operaciones `CreateDomain` y. Para obtener más información sobre la instalación y el uso de AWS SDKs, consulte los kits de desarrollo de AWS software.

Después de que el dominio termine de procesarse, consulte the section called “Permitir el rol autenticado” y the section called “Configuración de proveedores de identidades” para conocer los pasos adicionales de configuración.

Permitir el rol autenticado

De forma predeterminada, el rol de IAM autenticado que configuró siguiendo las instrucciones que se indican en él the section called “Acerca del grupo de identidades” no tiene los privilegios necesarios para acceder OpenSearch a los paneles. Debe proporcionar permisos adicionales al rol.

> **Note**
>
> Si configuró el control de acceso detallado y utiliza una política de acceso abierta o basada en IP, puede omitir este paso.

Puede incluir estos permisos en una política basada en la identidad, pero a menos que desee que los usuarios autenticados tengan acceso a todos los dominios del OpenSearch servicio, lo mejor es aplicar una política basada en recursos asociada a un único dominio.

Para el `Principal`, especifique el ARN del rol autenticado de Cognito que configuró con las pautas que figuran en the section called “Acerca del grupo de identidades”.

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/*"
      }
   ]
}
```

Para obtener instrucciones sobre cómo agregar una política basada en recursos a un dominio de servicio, consulte. OpenSearch the section called “Configurar políticas de acceso”

Configuración de proveedores de identidades

Al configurar un dominio para usar la autenticación de Amazon Cognito para los paneles, OpenSearch Service agrega un cliente de aplicación al grupo de usuarios y agrega el grupo de usuarios al grupo de identidades como proveedor de autenticación.

> **Warning**
>
> No cambie el nombre ni elimine el cliente de aplicación.

En función de cómo configure el grupo de usuarios, puede que necesite crear cuentas de usuarios de forma manual, o bien los usuarios podrían crear sus propias cuentas. Si estos ajustes son aceptables, no es necesario realizar más acciones. Sin embargo, muchas personas prefieren usar proveedores de identidades externos.

Para habilitar un proveedor de identidades SAML 2.0, debe proporcionar un documento de metadatos de SAML. Para habilitar proveedores de identidades sociales como Login with Amazon, Facebook y Google, debe disponer de un ID de aplicación y de la clave secreta de la aplicación de esos proveedores. Puede habilitar cualquier combinación de proveedores de identidades.

La forma más sencilla de configurar el grupo de usuarios consiste en utilizar la consola de Amazon Cognito. Para obtener instrucciones, consulte Inicio de sesión en grupos de usuarios con proveedores de identidad de terceros y Ajustes específicos de una aplicación en los clientes de aplicación en la Guía para desarrolladores de Amazon Cognito.

(Opcional) Configuración de acceso pormenorizado

Es posible que haya observado que la configuración predeterminada del grupo de identidades asigna a todos los usuarios que inician sesión en el mismo rol de IAM (`Cognito_identitypoolAuth_Role`), lo que significa que todos los usuarios pueden acceder a los mismos recursos. AWS Si desea utilizar el control de acceso detallado con Amazon Cognito, por ejemplo, si desea que los analistas de la organización tengan acceso de solo lectura a varios índices, pero que los desarrolladores tengan acceso de escritura a todos los índices, tiene dos opciones:

- Cree grupos de usuarios y configure su proveedor de identidades de tal forma que elija el rol de IAM en función del token de autenticación del usuario (recomendado).
- Configure su proveedor de identidades para que elija el rol de IAM en función de una o varias reglas.

Para obtener un tutorial que incluya el control de acceso detallado, consulte the section called "Tutorial: Control de acceso detallado con la autenticación de Cognito".

> **Important**
>
> Al igual que sucede con el rol predeterminado, Amazon Cognito debe formar parte de la relación de confianza de cada rol adicional. Para obtener más información, consulte Creación de roles para la asignación de roles en la Guía para desarrolladores de Amazon Cognito.

Grupos de usuarios y tokens

Al crear un grupo de usuarios, elige un rol de IAM para los miembros del grupo. Para obtener información acerca de cómo crear grupos, consulte Agregar grupos a un grupo de usuarios en la Guía para desarrolladores de Amazon Cognito.

Después de crear uno o más grupos de usuarios, puede configurar su proveedor de autenticación para asignar a los usuarios las funciones de sus grupos en lugar del rol predeterminado del grupo de identidades. Seleccione Elegir rol a partir de un token, luego elija Utilizar rol autenticado predeterminado o DENEGAR para especificar de qué manera el grupo de identidades se encargará de los usuarios que no forman parte del grupo.

Reglas

Las reglas son básicamente una serie de instrucciones `if` que Amazon Cognito evalúa de forma secuencial. Por ejemplo, si una dirección de email del usuario contiene `@corporate`, Amazon Cognito asigna `Role_A` a ese usuario. Si la dirección de correo electrónico de un usuario contiene `@subsidiary`, asigna a ese usuario `Role_B`. De lo contrario, asigna al usuario el rol autenticado predeterminado.

Para obtener más información, consulte Uso de la asignación basada en reglas para la asignación de roles a los usuarios en la Guía para desarrolladores de Amazon Cognito.

(Opcional) Personalización de la página de inicio de sesión

Puede usar la consola de Amazon Cognito para cargar un logotipo personalizado y realizar cambios de CSS en la página de inicio de sesión. Para obtener instrucciones y una lista completa de las propiedades de CSS, consulte Personalización de la marca de la interfaz de usuario alojada (clásica) en la Guía para desarrolladores de Amazon Cognito.

(Opcional) Configuración de seguridad avanzada

Los grupos de usuarios de Amazon Cognito admiten características de seguridad avanzadas como la autenticación multifactor, la verificación de credenciales comprometidas y la autenticación flexible. Para obtener más información, consulte Uso de las características de seguridad de los grupos de usuarios de Amazon Cognito en la Guía para desarrolladores de Amazon Cognito.

Testeo

Una vez que esté satisfecho con la configuración, verifique que la experiencia del usuario cumpla sus expectativas.

Para acceder a los paneles OpenSearch

1. Vaya a `https://`*`opensearch-domain`*`/_dashboards` en un navegador Web. Para iniciar sesión directamente en un inquilino específico, agregue `?security_tenant=`*`tenant-name`* a la URL.
2. Inicie sesión con las credenciales que prefiera.
3. Una vez que se carguen los OpenSearch paneles, configure al menos un patrón de índice. Dashboards utiliza estos patrones para identificar los índices que desea analizar. Escriba `*`, seleccione Siguiente paso y, a continuación, seleccione Crear patrón de índice.
4. Para buscar datos o explorar en ellos, seleccione Detección.

Si cualquier paso de este proceso da error, consulte the section called “Problemas habituales de configuración” para obtener información sobre resolución de problemas.

Cuotas

Amazon Cognito cuenta con límites flexibles en muchos de los recursos. Si desea habilitar la autenticación de Dashboards para un gran número de dominios de OpenSearch servicio, revise las cuotas en Amazon Cognito y solicite los aumentos del límite según sea necesario.

Cada dominio OpenSearch de servicio agrega un cliente de aplicación al grupo de usuarios, que agrega un proveedor de autenticación al grupo de identidades. Si habilita la autenticación de OpenSearch Dashboards para más de 10 dominios, es posible que se encuentre con el límite «máximo de proveedores de grupos de usuarios de Amazon Cognito por grupo de identidades». Si supera un límite, cualquier dominio de OpenSearch servicio que intente configurar para usar

la autenticación de Amazon Cognito para los paneles puede quedar atrapado en un estado de configuración de Procesamiento.

Problemas habituales de configuración

En las tablas siguientes se muestran los problemas habituales de configuración y las soluciones.

Configuración del servicio OpenSearch

Problema	Solución
`OpenSearch Service can't create the role` (consola)	No dispone de los permisos de IAM correctos. Agregue los permisos especificados en the section called "Configuración de la autenticación de Amazon Cognito (consola)".
`User is not authorized to perform: iam:PassRole on resource CognitoAccessForAmazonOpenSearch` (consola)	No tienes `iam:PassRole` permisos para el CognitoAccessForAmazonOpenSearchrol. Asocie la siguiente política a su cuenta: JSON `{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Action": [ "iam:PassRole" ], "Resource": "arn:aws:iam:: 123456789012:role/service-role/CognitoAccessForAmazonOpenSearch " } ] }` Como alternativa, puede adjuntar la política `IAMFullAccess` .

Problema	Solución
`User is not authorize d to perform: cognito- identity:ListIdenti tyPools on resource`	No tiene permisos de lectura para Amazon Cognito. Adjunte la política `AmazonCognitoReadOnly` a su cuenta.
`An error occurred (Validati onException) when calling the CreateDomain operation : OpenSearch Service must be allowed to use the passed role`	OpenSearch El servicio no está especificado en la relación de confianza del `CognitoAccessForAm azonOpenSearch` rol. Verifique si el rol utiliza la relación de confianza que se especifica en the section called “Acerca del CognitoAccessForAmazonOpenS earch rol”. De manera alternativa, utilice la consola para configurar la autenticación de Amazon Cognito. La consola crea un rol para usted.
`An error occurred (Validati onException) when calling the CreateDomain operation : User is not authorize d to perform: cognito-i dp: action on resource: user pool`	El rol especificado en `--cognito-options` no cuenta con permisos para acceder a Amazon Cognito. Compruebe que el rol tenga la `AmazonOpenSearchSe rviceCognitoAccess` política AWS gestionada adjunta. De manera alternativa, utilice la consola para configurar la autenticación de Amazon Cognito. La consola crea un rol para usted.
`An error occurred (Validati onException) when calling the CreateDomain operation : User pool does not exist`	OpenSearch El servicio no encuentra el grupo de usuarios. Verifique que creó uno y que tiene el ID correcto. Para encontrar el ID, puede utilizar la consola de Amazon Cognito o el siguiente comando: AWS CLI `aws cognito-idp list-user-pools --max-results 60 --region region`

Problema	Solución
`An error occurred (ValidationException) when calling the CreateDomain operation: IdentityPool not found`	OpenSearch El servicio no encuentra el grupo de identidades. Verifique que creó uno y que tiene el ID correcto. Para encontrar el ID, puede utilizar la consola de Amazon Cognito o el siguiente comando: AWS CLI `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`	El grupo de usuarios no tiene un nombre de dominio. Puede configurar uno con la consola de Amazon Cognito o el siguiente comando de la AWS CLI : `aws cognito-idp create-user-pool-domain --domain name --user-pool-id id`

Acceder a OpenSearch los paneles

Problema	Solución
La página de inicio de sesión no muestra mis proveedores de identidades preferidos.	Compruebe que ha activado el proveedor de identidad para el cliente de la aplicación de OpenSearch servicio, tal y como se especifica enthe section called “Configuración de proveedores de identidades”.
La página de inicio de sesión no parece estar asociada a mi organización.	Consulte the section called “(Opcional) Personalización de la página de inicio de sesión”.
Mis credenciales de inicio de sesión no funcionan.	Verifique que configuró el proveedor de identidades tal y como se especifica en the section called “Configuración de proveedores de identidades”. Si utiliza el grupo de usuarios como proveedor de identidades, verifique que la cuenta exista en la consola de Amazon Cognito.

Problema	Solución
OpenSearch Los paneles no se cargan en absoluto o no funcionan correctamente.	El rol autenticado de Amazon Cognito necesita permisos `es:ESHttp*` para el dominio (`/*`) a fin de obtener acceso y utilizar Dashboards. Verifique que agregó una política de acceso tal y como se especifica en the section called “Permitir el rol autenticado”.
Al cerrar sesión en los OpenSearch paneles de control desde una pestaña, en las pestañas restantes aparece un mensaje que indica que se ha revocado el token de actualización.	Cuando cierra sesión en una sesión de OpenSearch Dashboards mientras utiliza la autenticación de Amazon Cognito OpenSearch , Service ejecuta AdminUserGlobalSignOut una operación que cierra la sesión de todas las sesiones de Dashboards OpenSearch activas.

Problema	Solución
`Invalid identity pool configuration. Check assigned IAM roles for this pool.`	Amazon Cognito no cuenta con permisos para asumir el rol de IAM en nombre del usuario autenticado. Modifique la relación de confianza para el rol de tal forma que incluya: 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" } } }] }`
`Token is not from a supported provider of this identity pool.`	Este error poco habitual se puede producir al eliminar el cliente de aplicación del grupo de usuarios. Intente abrir Dashboards en una nueva sesión del navegador.

Desactivación de la autenticación de Amazon Cognito para paneles OpenSearch

Utilice el siguiente procedimiento para deshabilitar la autenticación de Amazon Cognito para Dashboards.

Para deshabilitar la autenticación de Amazon Cognito para Dashboards (consola)

1. Abre la consola OpenSearch de Amazon Service.
2. En Dominios, seleccione el dominio que desea configurar.
3. Seleccione Acciones y Editar la configuración de seguridad.
4. Desactive Habilitar la autenticación de Amazon Cognito.
5. Seleccione Guardar cambios.

> **Important**
>
> Si ya no necesita el grupo de usuarios y de identidades de Amazon Cognito, elimínelos. De lo contrario, seguirá incurriendo en costos.

Eliminar dominios que utilizan la autenticación de Amazon Cognito para los paneles OpenSearch

Para evitar que los dominios que utilizan la autenticación de Amazon Cognito para los paneles se atasquen en un estado de configuración de Procesamiento, elimine los dominios de OpenSearch servicio antes de eliminar sus grupos de usuarios e identidades de Amazon Cognito asociados.

Uso de roles vinculados a servicios para Amazon Service OpenSearch

Amazon OpenSearch Service utiliza funciones AWS Identity and Access Management vinculadas al servicio (IAM). Un rol vinculado a un servicio es un tipo único de rol de IAM que está vinculado directamente al Servicio. OpenSearch Los roles vinculados al servicio están predefinidos por el OpenSearch Servicio e incluyen todos los permisos que el servicio requiere para llamar a otros AWS servicios en su nombre.

Un rol vinculado a un servicio facilita la configuración del OpenSearch Servicio, ya que no es necesario añadir manualmente los permisos necesarios. OpenSearch Service define los permisos de sus funciones vinculadas al servicio y, a menos que se defina lo contrario, solo OpenSearch Service puede asumir sus funciones. Los permisos definidos incluyen las políticas de confianza y de permisos, y que la política de permisos no se pueda asociar a ninguna otra entidad de IAM. Para ver las actualizaciones de las políticas de permisos y funciones vinculadas a servicios, consulta el historial de documentos de Amazon OpenSearch Service.

Para obtener información sobre otros servicios que admiten funciones vinculadas a servicios, consulte los AWS servicios que funcionan con IAM y busque los servicios con la palabra Sí en la columna Funciones vinculadas a servicios. Elija una opción Sí con un enlace para ver la documentación acerca del rol vinculado al servicio en cuestión.

Temas

- Uso de roles vinculados al servicio para crear dominios de VPC y orígenes de datos de consulta directa
- Uso de roles vinculados a servicios para crear colecciones sin servidor OpenSearch
- Uso de roles vinculados a servicios para crear canalizaciones de ingestión OpenSearch

Uso de roles vinculados al servicio para crear dominios de VPC y orígenes de datos de consulta directa

Amazon OpenSearch Service utiliza funciones AWS Identity and Access Management vinculadas al servicio (IAM). Un rol vinculado a un servicio es un tipo único de rol de IAM que está vinculado directamente al Servicio. OpenSearch Los roles vinculados al servicio están predefinidos por el OpenSearch Servicio e incluyen todos los permisos que el servicio requiere para llamar a otros AWS servicios en su nombre.

OpenSearch El servicio usa el rol vinculado al servicio denominado AWSServiceRoleForAmazonOpenSearchService, que proporciona los permisos mínimos de Amazon EC2 y Elastic Load Balancing necesarios para que el rol permita el acceso de la VPC a un dominio o a una fuente de datos de consulta directa.

Función heredada de Elasticsearch

Amazon OpenSearch Service utiliza un rol vinculado a un servicio denominado. `AWSServiceRoleForAmazonOpenSearchService` Es posible que las cuentas también contengan

un rol vinculado a servicio denominado `AWSServiceRoleForAmazonElasticsearchService`, que funciona con los puntos de conexión de la API de Elasticsearch obsoletos.

Si el rol heredado de Elasticsearch no existe en tu cuenta, OpenSearch Service crea automáticamente un nuevo rol OpenSearch vinculado al servicio la primera vez que crees un dominio. OpenSearch En caso contrario, la cuenta seguirá utilizando el rol de Elasticsearch. Para que esta creación automática se realice correctamente, es necesario disponer de permisos para la acción `iam:CreateServiceLinkedRole`.

Permisos

El rol vinculado al servicio `AWSServiceRoleForAmazonOpenSearchService` depende de los siguientes servicios para asumir el rol:

- `opensearchservice.amazonaws.com`

La política de permisos de roles denominada `AmazonOpenSearchServiceRolePolicy`permite a OpenSearch Service completar las siguientes acciones en los recursos especificados:

- Acción: `acm:DescribeCertificate` en *
- Acción: `cloudwatch:PutMetricData` en *
- Acción: `ec2:CreateNetworkInterface` en *
- Acción: `ec2:DeleteNetworkInterface` en *
- Acción: `ec2:DescribeNetworkInterfaces` en *
- Acción: `ec2:ModifyNetworkInterfaceAttribute` en *
- Acción: `ec2:DescribeSecurityGroups` en *
- Acción: `ec2:DescribeSubnets` en *
- Acción: `ec2:DescribeVpcs` en *
- Acción: `ec2:CreateTags` en todas las interfaces de red y puntos de conexión de VPC
- Acción: `ec2:DescribeTags` en *
- Acción: `ec2:CreateVpcEndpoint` en todos los grupos de seguridad VPCs, subredes y tablas de enrutamiento, así como en todos los puntos finales de VPC cuando la solicitud contiene la etiqueta `OpenSearchManaged=true`

- Acción: `ec2:ModifyVpcEndpoint` en todos los grupos de seguridad VPCs, subredes y tablas de enrutamiento, así como en todos los puntos finales de VPC cuando la solicitud contiene la etiqueta `OpenSearchManaged=true`
- Acción: `ec2:DeleteVpcEndpoints` en todos los extremos cuando la solicitud contiene la etiqueta `OpenSearchManaged=true`
- Acción: `ec2:AssignIpv6Addresses` en *
- Acción: `ec2:UnAssignIpv6Addresses` en *
- Acción: `elasticloadbalancing:AddListenerCertificates` en *
- Acción: `elasticloadbalancing:RemoveListenerCertificates` en *

Debe configurar permisos para permitir a una entidad de IAM (como un usuario, grupo o rol) crear, editar o eliminar un rol vinculado a servicios. Para obtener más información, consulte Permisos de roles vinculados a servicios en la Guía del usuario de IAM.

Creación del rol vinculado a servicio

No necesita crear manualmente un rol vinculado a servicios. Cuando crea un dominio habilitado para VPC o una fuente de datos de consulta directa mediante el Consola de administración de AWS, el OpenSearch Servicio crea el rol vinculado al servicio para usted. Para que esta creación automática se realice correctamente, es necesario disponer de permisos para la acción `iam:CreateServiceLinkedRole`.

También puede utilizar la consola de IAM, la CLI de IAM o la API de IAM para crear un rol vinculado a servicios manualmente. Para obtener más información, consulte Creating a service-linked role en la Guía del usuario de IAM.

Edición del rol vinculado al servicio

OpenSearch El servicio no le permite editar el rol vinculado al servicio. `AWSServiceRoleForAmazonOpenSearchService` Después de crear un rol vinculado a un servicio, no puede cambiarle el nombre, ya que varias entidades pueden hacer referencia a él. Sin embargo, puede editar la descripción del rol mediante IAM. Para obtener más información, consulte Editar un rol vinculado a servicios en la Guía del usuario de IAM.

Eliminación del rol vinculado a un servicio

Si ya no necesita usar una característica o servicio que requieran un rol vinculado a un servicio, le recomendamos que elimine dicho rol. De esta forma, no tiene una entidad no utilizada que no se

monitoree ni mantenga de forma activa. Sin embargo, debe limpiar el rol vinculado a servicios antes de eliminarlo manualmente.

Limpieza del rol vinculado al servicio

Antes de poder utilizar IAM para eliminar un rol vinculado a un servicio, primero debe confirmar que dicho rol no tiene sesiones activas y eliminar los recursos que utiliza.

Para comprobar si el rol vinculado a un servicio tiene una sesión activa en la consola de IAM

1. Inicie sesión en la consola de IAM Consola de administración de AWS y ábrala en. https://console.aws.amazon.com/iam/
2. En el panel de navegación de la consola de IAM, elija Roles (Roles). A continuación, seleccione el nombre (no la casilla de verificación) del rol de `AWSServiceRoleForAmazonOpenSearchService`.
3. En la página Resumen del rol seleccionado, seleccione la pestaña Asesor de acceso.
4. En la pestaña Asesor de acceso, revise la actividad reciente del rol vinculado a servicios.

> **Note**
>
> Si no está seguro de si el OpenSearch Servicio está utilizando el `AWSServiceRoleForAmazonOpenSearchService` rol, puede intentar eliminarlo. Si el servicio utiliza el rol, este no podrá eliminarse y se podrán ver los recursos que lo utilizan. Si se está utilizando el rol, debe esperar a que finalice la sesión para poder eliminar el rol (eliminar los recursos que utilizan el rol). and/or No se puede revocar la sesión de un rol vinculado a servicios.

Eliminar manualmente un rol vinculado a servicios

Elimine las funciones vinculadas al servicio de la consola, la API o la CLI de IAM. AWS Para obtener más información, consulte Deleting a service-linked role en la Guía del usuario de IAM.

Uso de roles vinculados a servicios para crear colecciones sin servidor OpenSearch

OpenSearch Serverless usa roles vinculados al AWS Identity and Access Management servicio (IAM). Un rol vinculado a un servicio es un tipo único de rol de IAM que está vinculado directamente

al servicio. OpenSearch Los roles vinculados al servicio están predefinidos por el OpenSearch Servicio e incluyen todos los permisos que el servicio requiere para llamar a otros AWS servicios en su nombre.

OpenSearch Serverless usa el rol vinculado al servicio denominado AWSServiceRoleForAmazonOpenSearchServerless, que proporciona los permisos necesarios para que el rol publique métricas relacionadas con el servicio sin servidor CloudWatch en su cuenta. Se denomina a la política de permisos del rol asociada. AWSService RoleForAmazonOpenSearchServerless `AmazonOpenSearchServerlessServiceRolePolicy` Para obtener más información sobre la política, consulte AmazonOpenSearchServerlessServiceRolePolicyla Guía de referencia de políticas AWS administradas.

Permisos de rol vinculados al servicio para Serverless OpenSearch

OpenSearch Serverless usa el rol vinculado al servicio denominado AWSServiceRoleForAmazonOpenSearchServerless, que permite a OpenSearch Serverless llamar a los servicios en tu nombre. AWS

El rol AWSService RoleForAmazonOpenSearchServerless vinculado al servicio confía en los siguientes servicios para que asuman el rol:

- `observability.aoss.amazonaws.com`

La política de permisos de roles denominada `AmazonOpenSearchServerlessServiceRolePolicy` permite a OpenSearch Serverless realizar las siguientes acciones en los recursos especificados:

- Acción: `cloudwatch:PutMetricData` en todos los recursos de AWS .

> **Note**
>
> La política incluye la clave de condición`{"StringEquals": {"cloudwatch:namespace": "AWS/AOSS"}}`, lo que significa que el rol vinculado al servicio solo puede enviar datos de métricas al espacio de nombres. `AWS/AOSS` CloudWatch

Debe configurar permisos para permitir a una entidad de IAM (como un usuario, grupo o rol) crear, editar o eliminar un rol vinculado a servicios. Para obtener más información, consulte Permisos de roles vinculados a servicios en la Guía del usuario de IAM.

Crear el rol vinculado al servicio para Serverless OpenSearch

No necesita crear manualmente un rol vinculado a servicios. Al crear una colección OpenSearch Serverless en la Consola de administración de AWS, la o la AWS API AWS CLI, OpenSearch Serverless crea automáticamente el rol vinculado al servicio.

> **Note**
>
> La primera vez que cree una colección, se le debe asignar la `iam:CreateServiceLinkedRole` en una política basada en identidades.

Si elimina este rol vinculado a servicios y necesita crearlo de nuevo, puede utilizar el mismo proceso para volver a crear el rol en su cuenta. Al crear una colección Serverless, OpenSearch OpenSearch Serverless vuelve a crear el rol vinculado al servicio automáticamente.

También puede usar la consola de IAM para crear un rol vinculado a un servicio con el caso de uso de Amazon OpenSearch Serverless. En la API AWS CLI o en la AWS API, cree una función vinculada a un servicio con el nombre del servicio: `observability.aoss.amazonaws.com`

```
aws iam create-service-linked-role --aws-service-name
 "observability.aoss.amazonaws.com"
```

Para obtener más información, consulte Crear un rol vinculado a un servicio en la Guía del usuario de IAM. Si elimina este rol vinculado al servicio, puede utilizar este mismo proceso para volver a crear el rol.

Edición del rol vinculado al servicio para Serverless OpenSearch

OpenSearch Serverless no permite editar el rol vinculado al servicio. AWSService RoleForAmazonOpenSearchServerless Después de crear un rol vinculado a un servicio, no puede cambiarle el nombre, ya que varias entidades pueden hacer referencia a él. Sin embargo, puede editar la descripción del rol mediante IAM. Para obtener más información, consulte Editar un rol vinculado a servicios en la Guía del usuario de IAM.

Eliminar el rol vinculado al servicio para Serverless OpenSearch

Si ya no necesita usar una característica o servicio que requieran un rol vinculado a un servicio, le recomendamos que elimine dicho rol. Esto evita tener una entidad sin uso que no se supervisa ni mantiene activamente. Sin embargo, debe limpiar los recursos de su rol vinculado al servicio antes de eliminarlo manualmente.

Para eliminarlo AWSServiceRoleForAmazonOpenSearchServerless, primero debe eliminar todas las colecciones de OpenSearch Serverless de su. Cuenta de AWS

> **Note**
>
> Si OpenSearch Serverless utiliza el rol al intentar eliminar los recursos, es posible que la eliminación no se realice correctamente. En tal caso, espere unos minutos e intente de nuevo la operación.

Para eliminar manualmente el rol vinculado a servicios mediante IAM

Utilice la consola de IAM AWS CLI, la o la AWS API para eliminar la función vinculada al AWSService RoleForAmazonOpenSearchServerless servicio. Para obtener más información, consulte Eliminación de un rol vinculado a servicios en la Guía del usuario de IAM.

Regiones compatibles para OpenSearch funciones vinculadas a servicios sin servidor

OpenSearch Serverless admite el uso del rol AWSService RoleForAmazonOpenSearchServerless vinculado al servicio en todas las regiones en las que Serverless esté disponible. OpenSearch Para obtener una lista de las regiones compatibles, consulte los puntos de enlace y las cuotas de Amazon OpenSearch Serverless en. Referencia general de AWS

Uso de roles vinculados a servicios para crear canalizaciones de ingestión OpenSearch

Amazon OpenSearch Ingestion utiliza funciones vinculadas a servicios AWS Identity and Access Management (IAM). Un rol vinculado a un servicio es un tipo único de rol de IAM que está vinculado directamente a Ingestion. OpenSearch Los roles vinculados al servicio están predefinidos por OpenSearch Ingestion e incluyen todos los permisos que el servicio requiere para llamar a otros servicios en tu nombre. AWS

OpenSearch La ingestión usa el rol vinculado al servicio denominado AWSServiceRoleForAmazonOpenSearchIngestionService, excepto cuando se usa una VPC autogestionada, en cuyo caso usa el rol vinculado al servicio denominado. AWSServiceRoleForOpensearchIngestionSelfManagedVpce La política adjunta proporciona los permisos necesarios para que el rol cree una nube privada virtual (VPC) entre su cuenta e OpenSearch Ingestion y publique CloudWatch métricas en su cuenta.

Permisos

El rol vinculado al servicio `AWSServiceRoleForAmazonOpenSearchIngestionService` depende de los siguientes servicios para asumir el rol:

- `osis.amazonaws.com`

La política de permisos de roles denominada `AmazonOpenSearchIngestionServiceRolePolicy` permite a OpenSearch Ingestion realizar las siguientes acciones en los recursos especificados:

- Acción: `cloudwatch:PutMetricData` en `cloudwatch:namespace": "AWS/OSIS"`
- Acción: `ec2:CreateTags` en `arn:aws:ec2:*:*:network-interface/*`
- Acción: `ec2:CreateVpcEndpoint` en `*`
- Acción: `ec2:DeleteVpcEndpoints` en `*`
- Acción: `ec2:DescribeSecurityGroups` en `*`
- Acción: `ec2:DescribeSubnets` en `*`
- Acción: `ec2:DescribeVpcEndpoints` en `*`
- Acción: `ec2:ModifyVpcEndpoint` en `*`

El rol vinculado al servicio `AWSServiceRoleForOpensearchIngestionSelfManagedVpce` depende de los siguientes servicios para asumir el rol:

- `self-managed-vpce.osis.amazonaws.com`

La política de permisos de roles denominada `OpenSearchIngestionSelfManagedVpcePolicy` permite a OpenSearch Ingestion completar las siguientes acciones en los recursos especificados:

- Acción: `ec2:DescribeSubnets` en `*`

- Acción: `ec2:DescribeSecurityGroups` en *
- Acción: `ec2:DescribeVpcEndpoints` en *
- Acción: `cloudwatch:PutMetricData` en `cloudwatch:namespace": "AWS/OSIS"`

Debe configurar permisos para permitir a una entidad de IAM (como un usuario, grupo o rol) crear, editar o eliminar un rol vinculado a servicios. Para obtener más información, consulte Permisos de roles vinculados a servicios en la Guía del usuario de IAM.

Crear el rol vinculado al servicio para Ingestion OpenSearch

No necesita crear manualmente un rol vinculado a servicios. Al crear una canalización OpenSearch de Ingestión en la Consola de administración de AWS, la o la AWS API AWS CLI, OpenSearch Ingestión crea automáticamente la función vinculada al servicio.

Si elimina este rol vinculado a servicios y necesita crearlo de nuevo, puede utilizar el mismo proceso para volver a crear el rol en su cuenta. Al crear una canalización de Ingestión, OpenSearch OpenSearch Ingestión vuelve a crear el rol vinculado al servicio automáticamente.

Edición del rol vinculado al servicio para Ingestion OpenSearch

OpenSearch La ingestión no permite editar el rol vinculado al servicio. `AWSServiceRoleForAmazonOpenSearchIngestionService` Después de crear un rol vinculado al servicio, no podrá cambiar el nombre del rol, ya que varias entidades podrían hacer referencia al rol. Sin embargo, sí puede editar la descripción del rol con IAM. Para obtener más información, consulte Editar un rol vinculado a servicios en la Guía del usuario de IAM.

Eliminar el rol vinculado al servicio para Ingestion OpenSearch

Si ya no necesita utilizar una característica o servicio que requiere un rol vinculado a un servicio, le recomendamos que elimine dicho rol. De esta forma no tiene una entidad no utilizada que no se monitoree ni mantenga de forma activa. Sin embargo, debe limpiar los recursos del rol vinculado al servicio antes de eliminarlo manualmente.

Saneamiento de un rol vinculado a servicios

Antes de que pueda utilizar IAM para eliminar un rol vinculado a servicios, primero debe eliminar los recursos que utiliza el rol.

> **Note**
>
> Si OpenSearch Ingestion utiliza el rol al intentar eliminar los recursos, es posible que la eliminación no se realice correctamente. En tal caso, espere unos minutos e intente de nuevo la operación.

Para eliminar los recursos OpenSearch de ingestión utilizados por el **`AWSServiceRoleForAmazonOpenSearchIngestionService`** rol o **`AWSServiceRoleForOpensearchIngestionSelfManagedVpce`**

1. Ve a la consola OpenSearch de Amazon Service y selecciona Ingestión.
2. Elimine todas las canalizaciones. Para obtener instrucciones, consulte the section called "Eliminación de canalizaciones".

Elimine el rol vinculado al servicio para Ingestion OpenSearch

Puede usar la consola de OpenSearch ingestión para eliminar un rol vinculado a un servicio.

Para eliminar un rol vinculado a un servicio (consola)

1. Vaya a la consola de IAM.
2. Elija Roles y busque el AWSServiceRoleForAmazonOpenSearchIngestionServicerol o. AWSServiceRoleForOpensearchIngestionSelfManagedVpce
3. Seleccione el rol y elija Eliminar.

Código de ejemplo para Amazon OpenSearch Service

En este capítulo se incluyen ejemplos de código habituales para trabajar con Amazon OpenSearch Service: firma de solicitudes HTTP en varios lenguajes de programación, comprime los cuerpos de las solicitudes HTTP y utiliza AWS SDKs para crear dominios.

Temas

- Compatibilidad con clientes de Elasticsearch
- Comprimir las solicitudes HTTP en Amazon OpenSearch Service
- Uso del OpenSearch Servicio AWS SDKs para interactuar con Amazon

Compatibilidad con clientes de Elasticsearch

Las versiones más recientes de los clientes de Elasticsearch pueden incluir verificaciones de licencia o versión que rompen artificialmente la compatibilidad. La siguiente tabla incluye recomendaciones sobre qué versiones de esos clientes se deben utilizar para lograr una mejor compatibilidad con el OpenSearch Servicio.

> **Important**
>
> Estas versiones de cliente están desactualizadas y no se actualizan con las dependencias más recientes, incluida Log4j. Recomendamos encarecidamente utilizar las OpenSearch versiones de los clientes siempre que sea posible.

Cliente	Versión recomendada
Cliente REST de bajo nivel de Java	7.13.4
Cliente REST de alto nivel de Java	7,13.4
Cliente de Elasticsearch de Python	7,13.4
Cliente de Elasticsearch de Ruby	7.13.3
Cliente de Elasticsearch de Node.js	7,13.0

Comprimir las solicitudes HTTP en Amazon OpenSearch Service

Puedes comprimir las solicitudes y respuestas HTTP en los dominios de Amazon OpenSearch Service mediante la compresión gzip. La compresión gzip puede ayudar a reducir el tamaño de los documentos y a reducir la utilización y la latencia de la banda ancha, lo que permite mejorar las velocidades de transferencia.

La compresión Gzip es compatible con todos los dominios que ejecutan Elasticsearch OpenSearch 6.0 o una versión posterior. Algunos OpenSearch clientes tienen soporte integrado para la compresión gzip, y muchos lenguajes de programación tienen bibliotecas que simplifican el proceso.

Habilitar la compresión gzip

No debe confundirse con una OpenSearch configuración similar, `http_compression.enabled` es específica del OpenSearch Servicio y habilita o deshabilita la compresión gzip en un dominio. Dominios que ejecutan OpenSearch Elasticsearch 7. x tienen la compresión gzip habilitada de forma predeterminada, mientras que los dominios que ejecutan Elasticsearch 6. x la tienen deshabilitada de forma predeterminada.

Para habilitar la compresión gzip, envíe la siguiente solicitud:

```
PUT _cluster/settings
{
  "persistent" : {
    "http_compression.enabled": true
  }
}
```

Solicitudes a `_cluster/settings` deben estar descomprimidas, por lo que es posible que necesite utilizar un cliente independiente o una solicitud HTTP estándar para actualizar la configuración del clúster.

Para confirmar que ha habilitado correctamente la compresión gzip, envíe la siguiente solicitud:

```
GET _cluster/settings?include_defaults=true
```

Asegúrese de ver la siguiente configuración en la respuesta:

```
...
```

```
"http_compression": {
  "enabled": "true"
}
...
```

Encabezados obligatorios

Al incluir un cuerpo de la solicitud comprimido por gzip, mantenga el encabezado estándar `Content-Type: application/json` y agregue el encabezado `Content-Encoding: gzip`. Para aceptar una respuesta comprimida con gzip, agregue también el encabezado `Accept-Encoding: gzip`. Si un OpenSearch cliente admite la compresión gzip, es probable que incluya estos encabezados automáticamente.

Código de muestra (Python 3)

En el siguiente ejemplo se utiliza opensearch-py para realizar la compresión y enviar la solicitud. Este código firma la solicitud mediante sus credenciales de 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))
```

Como alternativa, puede especificar los encabezados adecuados, comprimir el cuerpo de la solicitud usted mismo y utilizar una biblioteca HTTP estándar como Solicitudes. Este código firma la solicitud con credenciales básicas HTTP, que podrían ser compatibles con el dominio si utiliza el control de acceso detallado.

```
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)
```

Uso del OpenSearch Servicio AWS SDKs para interactuar con Amazon

En esta sección se incluyen ejemplos de cómo utilizarla AWS SDKs para interactuar con la API de configuración OpenSearch de Amazon Service. Estos ejemplos de código muestran cómo crear, actualizar y eliminar dominios OpenSearch de servicio.

Java

En esta sección se incluyen ejemplos de las versiones 1 y 2 de AWS SDK para Java.

Version 2

En este ejemplo, se utiliza el OpenSearchClientBuilderconstructor de la versión 2 AWS SDK para Java para crear un OpenSearch dominio, actualizar su configuración y eliminarlo. Quite los comentarios de las llamadas a `waitForDomainProcessing` (y el comentario de la llamada a `deleteDomain`) para permitir que se active el dominio y que se pueda utilizar.

```
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

En este ejemplo, se usa el AWSElasticsearchClientBuilderconstructor de la versión 1 AWS SDK para Java para crear un dominio de Elasticsearch heredado, actualizar su configuración y eliminarlo. Quite los comentarios de las llamadas a `waitForDomainProcessing` (y el comentario de la llamada a `deleteDomain`) para permitir que se active el dominio y que se pueda utilizar.

```
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

En este ejemplo, se utiliza el cliente Python de OpenSearchServicebajo nivel del AWS SDK para Python (Boto) para crear un dominio, actualizar su configuración y 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

En este ejemplo, se utiliza la versión 3 del SDK del OpenSearch cliente Node.js para crear un dominio, actualizar su configuración y eliminarlo. JavaScript

```
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.');
            }
    };
}
```

Indexación de datos en Amazon Service OpenSearch

Como Amazon OpenSearch Service usa una API REST, existen numerosos métodos para indexar documentos. Puede utilizar clientes estándar como curl o cualquier lenguaje de programación que pueda enviar solicitudes HTTP. Para simplificar aún más el proceso de interacción con ella, OpenSearch Service tiene clientes para muchos lenguajes de programación. Los usuarios avanzados pueden pasar directamente a the section called “Cargando datos de streaming en el OpenSearch servicio”.

Le recomendamos encarecidamente que utilice Amazon OpenSearch Ingestion para ingerir datos, que es un recopilador de datos totalmente gestionado integrado en OpenSearch Service. Para obtener más información, consulta Amazon OpenSearch Ingestion.

Para obtener una introducción a la indexación, consulte la OpenSearch documentación.

Restricciones de nomenclatura de los índices

OpenSearch Los índices de servicios tienen las siguientes restricciones de denominación:

- Todas las letras deben estar en minúsculas.
- Los nombres de los índices no pueden empezar por `_` ni `-`.
- Los nombres de los índices no pueden contener espacios, comas, `:`, `"`, `*`, `+`, `/`, `\`, `|`, `?`, `#`, `>` o `<`.

No incluya información confidencial en los nombres de los índices, tipos o identificadores de documentos. OpenSearch El servicio usa estos nombres en sus identificadores uniformes de recursos (URIs). Los servidores y las aplicaciones suelen registrar las solicitudes HTTP, lo que puede provocar una exposición innecesaria de los datos si URIs contienen información confidencial:

```
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"
```

Aunque no tenga permisos para ver el documento JSON asociado, puede inferir de esta línea de registro falsa que uno de los pacientes del Dr. Doe, cuyo número de teléfono es el 202-555-0100, tuvo gripe en 2018.

Si el OpenSearch servicio detecta una dirección IP real o percibida en un nombre de índice (por ejemplo,`my-index-12.34.56.78.91`), oculta la dirección IP. Una llamada a `_cat/indices` produce la siguiente respuesta:

```
green open my-index-x.x.x.x.91    soY19tBERoKo71WcEScidw 5 1 0 0   2kb  1kb
```

Para evitar confusiones innecesarias, evite incluir direcciones IP en los nombres de los índices.

Reducción del tamaño de la respuesta

Las respuestas de `_index` y `_bulk` APIs contienen bastante información. Esta información puede resultar útil para resolver problemas de las solicitudes o para implementar la lógica de reintentos, aunque puede utilizar una banda ancha considerable. En este ejemplo, la indexación de un documento de 32 bytes se traduce en una respuesta de 339 bytes (incluidos los encabezados):

```
PUT opensearch-domain/more-movies/_doc/1
{"title": "Back to the Future"}
```

Respuesta

```
{
  "_index": "more-movies",
  "_type": "_doc",
  "_id": "1",
  "_version": 4,
  "result": "updated",
  "_shards": {
    "total": 2,
    "successful": 2,
    "failed": 0
  },
  "_seq_no": 3,
  "_primary_term": 1
}
```

Este tamaño de respuesta puede parecer mínimo, pero si indexa 1 000 000 de documentos al día (aproximadamente 11,5 documentos por segundo), 339 bytes por respuesta equivalen a 10,17 GB de tráfico de descarga al mes.

Si le preocupan los costes de transferencia de datos, utilice el `filter_path` parámetro para reducir el tamaño de la respuesta del OpenSearch Servicio, pero tenga cuidado de no filtrar los campos que necesite para identificar o reintentar las solicitudes fallidas. Estos campos varían por cliente. El `filter_path` parámetro funciona para todos los REST del OpenSearch servicio APIs, pero resulta especialmente útil cuando se llama con APIs frecuencia, por ejemplo, `_index` y `_bulk` APIs:

```
PUT opensearch-domain/more-movies/_doc/1?filter_path=result,_shards.total
{"title": "Back to the Future"}
```

Respuesta

```
{
  "result": "updated",
  "_shards": {
    "total": 2
  }
}
```

En lugar de incluir campos, puede excluir campos con un prefijo `-`. `filter_path` también es compatible con comodines:

```
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"}
```

Respuesta

```
{
  "errors": false,
  "items": [
    {
      "index": {
        "result": "updated",
        "status": 200
      }
    },
    {
      "index": {
        "result": "updated",
```

```
        "status": 200
      }
    }
  ]
}
```

Códecs de índice

Los códecs de índice determinan cómo se comprimen y almacenan en el disco los campos almacenados en un índice. El códec de índice se controla mediante la configuración `index.codec` estática, que especifica el algoritmo de compresión. Esta configuración afecta el tamaño de la partición del índice y el rendimiento de la operación.

Para obtener una lista de los códecs compatibles y sus características de rendimiento, consulte Códecs compatibles en la documentación. OpenSearch

Cuando elija un códec de índice, tenga en cuenta lo siguiente:

- Para evitar los problemas que supone cambiar la configuración del códec de un índice existente, pruebe una carga de trabajo representativa en un entorno que no sea de producción antes de utilizar una nueva configuración de códec. Para más información, consulte Cambio de un códec de índice.
- No puede usar los códecs de compresión Zstandard (`"index.codec": "zstd"` o `"index.codec": "zstd_no_dict"`) para los índices k-NN o Security Analytics.

Carga de datos de streaming en Amazon OpenSearch Service

Puedes usar OpenSearch Ingestion para cargar directamente los datos de streaming en tu dominio de Amazon OpenSearch Service, sin necesidad de utilizar soluciones de terceros. Para enviar datos a OpenSearch Ingestion, debes configurar tus generadores de datos y el servicio entrega automáticamente los datos al dominio o la colección que especifiques. Para empezar a usar OpenSearch Ingestion, consulte. the section called “Tutorial: incorporar datos en una colección”

Puedes seguir utilizando otras fuentes para cargar datos de streaming, como Amazon Data Firehose y Amazon CloudWatch Logs, que tienen soporte integrado para OpenSearch Service. Otros, como Amazon S3, Amazon Kinesis Data Streams y Amazon DynamoDB, utilizan funciones AWS Lambda como controladores de eventos. Las funciones de Lambda responden a los nuevos datos procesándolos y transmitiéndolos al dominio.

> **Note**
>
> Lambda es compatible con varios lenguajes de programación populares y está disponible en la mayoría de las Regiones de AWS. Para más información, consulte Introducción a Lambda en la Guía para desarrolladores de AWS Lambda y Puntos de conexión de servicio de AWS en Referencia general de AWS.

Cargando datos de streaming desde Ingestion OpenSearch

Puede usar Amazon OpenSearch Ingestion para cargar datos en un dominio OpenSearch de servicio. Usted configura sus generadores de datos para que envíen datos a OpenSearch Ingestion y esta entrega automáticamente los datos a la colección que usted especifique. También puede configurar OpenSearch Ingestion para transformar los datos antes de entregarlos. Para obtener más información, consulte OpenSearch Ingestión de Amazon.

Carga de datos de streaming desde Amazon S3

Puede usar Lambda para enviar datos a su dominio de OpenSearch servicio desde Amazon S3. Cuando llegan datos nuevos a un bucket de S3, activan una notificación de eventos en Lambda que, a su vez, ejecuta el código personalizado para realizar la indexación.

Este método de streaming de datos es extremadamente flexible. Es posible indexar los metadatos del objeto o, si el objeto contiene texto sin formato, analizar e indexar algunos elementos del cuerpo del objeto. En esta sección, se incluye código de muestra sencillo de Python que utiliza expresiones regulares para analizar un archivo de registros e indexar los resultados obtenidos.

Requisitos previos

Antes de continuar, debe contar con los siguientes recursos.

Requisito previo	Description (Descripción)
Bucket de Amazon S3	Para más información, consulte Crear su primer bucket de S3 en la Guía del usuario de Amazon Simple Storage Service. El bucket debe residir en la misma región que su dominio OpenSearch de servicio.

Requisito previo	Description (Descripción)
OpenSearch Dominio de servicio	Es el destino de los datos después de que la función de Lambda los procesa. Para más información, consulte the section called “Creación de dominios OpenSearch de servicio”.

Crear el paquete de implementación de Lambda

Los paquetes de implementación son archivos ZIP o JAR que contienen el código y sus dependencias. En esta sección, se incluye un código de muestra de Python. Para otros lenguajes de programación, consulte Paquetes de implementación de Lambda en la Guía para desarrolladores de AWS Lambda .

1. Cree un directorio. En este ejemplo, utilizamos el nombre `s3-to-opensearch`.
2. Cree un archivo en el directorio denominado `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)
```

Edite las variables para `region` y `host`.

3. Instale pip, si todavía no lo hizo, luego instale las dependencias en un directorio `package` nuevo:

```
cd s3-to-opensearch

pip install --target ./package requests
pip install --target ./package requests_aws4auth
```

Todos los entornos de ejecución de Lambda tienen instalado Boto3, por lo que no es necesario incluirlo en el paquete de implementación.

4. Cree el paquete con el código de la aplicación y las dependencias:

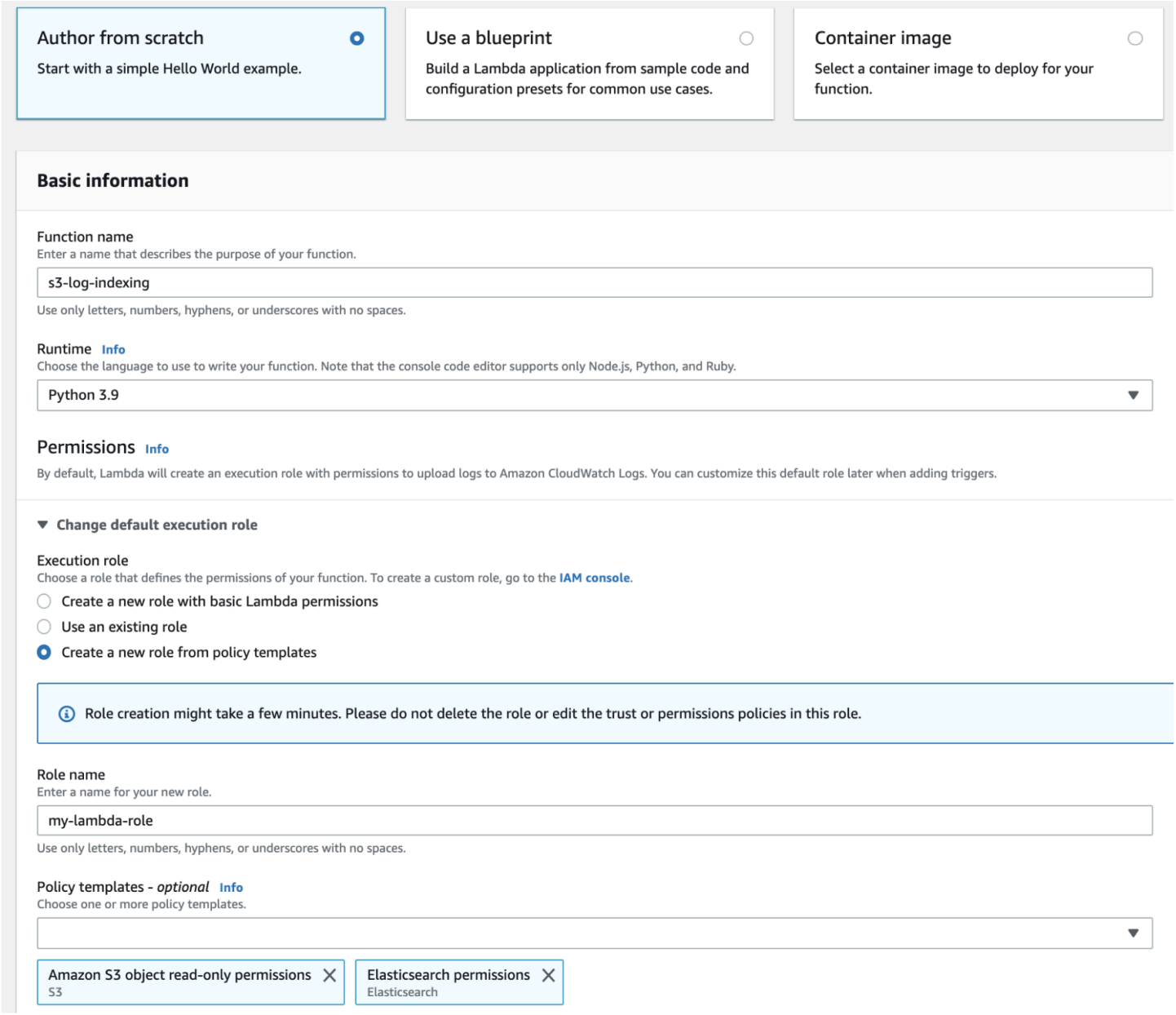
```
cd package
zip -r ../lambda.zip .
```

```
cd ..
zip -g lambda.zip sample.py
```

Crear la función de Lambda

Después de crear el paquete de implementación, puede crear la función de Lambda. Al crear una función, elija un nombre, un tiempo de ejecución (por ejemplo, Python 3.8) y un rol de IAM. El rol de IAM define los permisos para la función. Para obtener instrucciones detalladas, consulte Create a Lambda function with the consol en la Guía para desarrolladores de AWS Lambda .

En este ejemplo, se supone que utiliza la consola. Elija Python 3.9 y un rol que tenga permisos de lectura en S3 y permisos de escritura en el OpenSearch servicio, como se muestra en la siguiente captura de pantalla:



Después de crear la función, debe agregar un desencadenador. En este ejemplo, queremos que el código se ejecute cada vez que llega un archivo de registros a un bucket de S3:

1. Seleccione Agregar desencadenador y Seleccione S3.
2. Seleccione el bucket.
3. Para el Tipo de evento, elija PUT.
4. Para el Prefijo, escriba `logs/`.
5. Para el Sufijo, escriba `.log`.
6. Reconozca la advertencia de invocación recursiva y elija Agregar.

Por último, puede cargar el paquete de implementación:

1. Seleccione Cargar desde y archivo .zip, luego siga las instrucciones para cargar el paquete de implementación.
2. Una vez finalizada la carga, edite la Configuración de tiempo de ejecución y cambie el Controlador a `sample.handler`. Esta configuración indica a Lambda el archivo (`sample.py`) y el método (`handler`) que debe ejecutar cuando se produzca un desencadenador.

En este punto, dispone de un conjunto completo de recursos: un depósito para los archivos de registro, una función que se ejecuta cada vez que se añade un archivo de registro al depósito, un código que realiza el análisis y la indexación y un dominio de OpenSearch servicio para la búsqueda y la visualización.

Prueba de la función de Lambda

Después de crear la función, puede probarla mediante la carga de un archivo en el bucket de Amazon S3. Cree un archivo denominado `sample.log` que contenga los siguientes ejemplos de líneas de registro:

```
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"
```

Cargue el archivo en la carpeta `logs` del bucket de S3. Para obtener instrucciones, consulte Carga de un objeto en el bucket en la Guía del usuario de Amazon Simple Storage Service.

A continuación, utilice la consola OpenSearch de servicio o los OpenSearch paneles de control para comprobar que el `lambda-s3-index` índice contiene dos documentos. También puede realizar una petición de búsqueda estándar:

```
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"
        }
      }
    ]
  }
}
```

Cargar datos de streaming desde Amazon Kinesis Data Streams

Puede cargar datos de streaming desde Kinesis Data Streams OpenSearch a Service. Cuando llegan datos nuevos al flujo de datos, activan una notificación de eventos en Lambda que, a su vez, ejecuta el código personalizado para realizar la indexación. En esta sección, se incluye un código de muestra simple de Python.

Requisitos previos

Antes de continuar, debe contar con los siguientes recursos.

Requisito previo	Description (Descripción)
Amazon Kinesis Data Streams	Fuente de eventos de la función Lambda Para más información, consulte Kinesis Data Streams.
OpenSearch Dominio de servicio	Es el destino de los datos después de que la función Lambda los procesa. Para más información, consulte the section called “Creación de dominios OpenSearch de servicio”.

Requisito previo	Description (Descripción)

Rol de IAM	Esta función debe tener permisos básicos OpenSearch de Servicio, Kinesis y Lambda, como los siguientes:
	
	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": "*" } ] }`
	
	El rol debe tener la siguiente relación de confianza:
	
	JSON
	
	`{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Principal": { "Service": "lambda.amazonaws.com" },`

Requisito previo	Description (Descripción)
	"Action": "sts:AssumeRole" }] } Para más información, consulte Creación de roles de IAM en la Guía del usuario de IAM.

Crear la función de Lambda

Siga las instrucciones de the section called “Crear el paquete de implementación de Lambda”, pero cree un directorio denominado `kinesis-to-opensearch` y utilice el siguiente código para `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.'
```

Edite las variables para `region` y `host`.

Instale pip, si todavía no lo hizo, luego utilice los siguientes comandos para instalar las dependencias:

```
cd kinesis-to-opensearch

pip install --target ./package requests
pip install --target ./package requests_aws4auth
```

A continuación, siga las instrucciones de the section called “Crear la función de Lambda”, pero especifique el rol de IAM de the section called “Requisitos previos” y la configuración siguiente para el desencadenador:

- Flujo de Kinesis: su flujo de Kinesis
- Tamaño del lote: 100
- Posición inicial: Trim horizon

Para más información, consulte ¿Qué es Amazon Kinesis Data Streams? en la Guía para desarrolladores de Amazon Kinesis Data Streams.

En este punto, dispone de un conjunto completo de recursos: una transmisión de datos de Kinesis, una función que se ejecuta después de que la transmisión reciba nuevos datos y los indexe, y un dominio de OpenSearch servicio para la búsqueda y la visualización.

Prueba de la función de Lambda

Después de crear la función, puede probarla al agregar un registro al flujo de datos mediante la AWS CLI:

```
aws kinesis put-record --stream-name test --data "My test data." --partition-key
 partitionKey1 --region us-west-1
```

A continuación, utilice la consola OpenSearch de servicio o los OpenSearch paneles de control para comprobar que `lambda-kine-index` contiene un documento. También puede utilizar la solicitud siguiente:

```
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"
      }
    }
  ]
}
```

Carga de datos de streaming desde Amazon DynamoDB

Puede utilizarlos AWS Lambda para enviar datos a su dominio de OpenSearch servicio desde Amazon DynamoDB. Cuando llegan datos nuevos a la tabla de base de datos, activan una notificación de eventos en Lambda que, a su vez, ejecuta el código personalizado para realizar la indexación.

Requisitos previos

Antes de continuar, debe contar con los siguientes recursos.

Requisito previo	Description (Descripción)
Tabla de DynamoDB	La tabla contiene los datos de origen. Para más información, consulte Operaciones básicas en tablas de DynamoDB en la Guía para desarrolladores de Amazon DynamoDB. La tabla debe residir en la misma región que su dominio de OpenSearch servicio y tener una transmisión configurada como Nueva imagen. Para más información, consulte Habilitación de un flujo.
OpenSearch Dominio de servicio	Es el destino de los datos después de que la función de Lambda los procesa. Para más información, consulte the section called “Creación de dominios OpenSearch de servicio”.

rol de IAM	Esta función debe tener permisos básicos OpenSearch de ejecución de Servicios, DynamoDB y Lambda, como los siguientes:
	
	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": "*"
	}
	]
	}

Requisito previo	Description (Descripción)
	El rol debe tener la siguiente relación de confianza: JSON `{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Principal": { "Service": "lambda.amazonaws.com" }, "Action": "sts:AssumeRole" } ] }` Para más información, consulte Creación de roles de IAM en la Guía del usuario de IAM.

Crear la función de Lambda

Siga las instrucciones de the section called “Crear el paquete de implementación de Lambda”, pero cree un directorio denominado `ddb-to-opensearch` y utilice el siguiente código para `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.'
```

Edite las variables para `region` y `host`.

Instale pip, si todavía no lo hizo, luego utilice los siguientes comandos para instalar las dependencias:

```
cd ddb-to-opensearch

pip install --target ./package requests
pip install --target ./package requests_aws4auth
```

A continuación, siga las instrucciones de the section called “Crear la función de Lambda”, pero especifique el rol de IAM de the section called “Requisitos previos” y la configuración siguiente para el desencadenador:

- Tabla: tabla de DynamoDB
- Tamaño del lote: 100
- Posición inicial: Trim horizon

Para más información, consulte Process New Items with DynamoDB Streams and Lambda en la Guía para desarrolladores de Amazon DynamoDB.

En este punto, dispone de un conjunto completo de recursos: una tabla de DynamoDB para los datos de origen, un flujo de cambios de DynamoDB en la tabla, una función que se ejecuta después de que los datos de origen cambien e indexa esos cambios, y un dominio de servicio para la búsqueda y la visualización. OpenSearch

Prueba de la función de Lambda

Después de crear la función, puede probarla al agregar un elemento a la tabla de DynamoDB mediante 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
```

A continuación, utilice la consola de OpenSearch servicio o los OpenSearch paneles de control para comprobar que contiene un documento. `lambda-index` También puede utilizar la solicitud siguiente:

```
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"
        }
    }
}
```

Carga de datos de streaming desde Amazon Data Firehose

Firehose admite el OpenSearch servicio como destino de entrega. Para obtener instrucciones sobre cómo cargar datos de streaming en el OpenSearch servicio, consulte Crear una transmisión de

entrega de Kinesis Data Firehose OpenSearch y elegir un servicio para su destino en la Guía para desarrolladores de Amazon Data Firehose.

Antes de cargar datos en OpenSearch Service, es posible que deba realizar transformaciones en los datos. Para obtener más información acerca de cómo utilizar las funciones de Lambda para realizar esta tarea, consulte Amazon Kinesis Data Firehose Data Transformation en la misma guía.

Al configurar una transmisión de entrega, Firehose incluye una función de IAM con un solo clic que le proporciona el acceso a los recursos que necesita para enviar datos al OpenSearch Servicio, hacer copias de seguridad de los datos en Amazon S3 y transformar los datos con Lambda. Debido a la complejidad que supone la creación de ese rol manualmente, recomendamos utilizar el rol proporcionado.

Carga de datos de streaming desde Amazon CloudWatch

Puedes cargar datos de streaming desde CloudWatch Logs a tu dominio OpenSearch de Service mediante una suscripción a CloudWatch Logs. Para obtener información sobre CloudWatch las suscripciones de Amazon, consulta Procesamiento en tiempo real de datos de registro con suscripciones. Para obtener información sobre la configuración, consulte Transmisión de datos de CloudWatch registros a Amazon OpenSearch Service en la Guía para CloudWatch desarrolladores de Amazon.

Cargando datos de streaming desde AWS IoT

Puede enviar datos AWS IoT mediante reglas. Para obtener más información, consulta la OpenSearchacción en la Guía para AWS IoT desarrolladores.

Carga de datos en Amazon OpenSearch Service con Logstash

La versión de código abierto de Logstash (Logstash OSS) proporciona una forma cómoda de utilizar la API masiva para cargar datos en tu dominio de Amazon Service. OpenSearch El servicio es compatible con todos los complementos de entrada estándar de Logstash, incluido el complemento de entrada Amazon S3. OpenSearch El servicio es compatible con el complemento logstash-output-opensearchde salida, que admite tanto la autenticación básica como las credenciales de IAM. El complemento funciona con la versión 8.1 y versiones inferiores de Logstash OSS.

Configuración

La configuración de Logstash varía en función del tipo de autenticación que utiliza su dominio.

Independientemente del método de autenticación que utilice, debe configurar `ecs_compatibility` en `disabled` en la sección de salida del archivo de configuración. Logstash 8.0 introdujo un cambio revolucionario en el que se ejecutan todos los complementos en Modo de compatibilidad ECS de forma predeterminada. Debe anular el valor predeterminado para mantener el comportamiento heredado.

Configuración de control de acceso detallado

Si su dominio OpenSearch de servicio utiliza un control de acceso detallado con autenticación básica HTTP, la configuración es similar a la de cualquier otro clúster. OpenSearch Este archivo de configuración de ejemplo toma su entrada de la versión de código abierto de 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 configuración varía según la aplicación de Beats y el caso de uso, pero la configuración de Filebeat OSS podría verse así:

```
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"]
```

Configuración de IAM

Si tu dominio usa una política de acceso al dominio basada en IAM o un control de acceso detallado con un usuario maestro, debes firmar todas las solicitudes al Servicio con credenciales de IAM. OpenSearch La siguiente política basada en identidades concede todas las solicitudes HTTP a los subrecursos de su dominio.

JSON

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

Para configurar su configuración de Logstash, cambie su archivo de configuración para usar el complemento en su salida. Este ejemplo de archivo de configuración toma su entrada de archivos en un bucket de 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
  }
}
```

Si no desea proporcionar sus credenciales de IAM dentro del archivo de configuración, puede exportarlas (o ejecutarlas `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"
```

Si su dominio de OpenSearch servicio está en una VPC, la máquina Logstash OSS debe poder conectarse a la VPC y tener acceso al dominio a través de los grupos de seguridad de la VPC. Para obtener más información, consulte the section called “Acerca de las políticas de acceso en los dominios de VPC”.

Búsqueda de datos en Amazon OpenSearch Service

Existen varios métodos comunes para buscar documentos en Amazon OpenSearch Service, incluidas las búsquedas de URI y los registros corporales de las solicitudes. OpenSearch El servicio ofrece funciones adicionales que mejoran la experiencia de búsqueda, como paquetes personalizados, compatibilidad con SQL y búsqueda asíncrona. Para obtener una referencia completa sobre la API de OpenSearch búsqueda, consulta la documentación. OpenSearch

> **Note**
>
> Los siguientes ejemplos de solicitudes funcionan con OpenSearch APIs. Algunas solicitudes podrían no funcionar con versiones anteriores de Elasticsearch.

Temas

- Búsquedas de URI
- Búsquedas de cuerpo de la solicitud
- Paginar los resultados de búsqueda
- Lenguaje de consulta de paneles
- Importación y gestión de paquetes en Amazon OpenSearch Service
- Consulta tus datos de Amazon OpenSearch Service con SQL
- Búsqueda entre clústeres en Amazon Service OpenSearch
- Aprender a clasificar para Amazon OpenSearch Service
- Búsqueda asíncrona en Amazon Service OpenSearch
- Búsqueda puntual en Amazon OpenSearch Service
- Búsqueda de agentes en Amazon Service OpenSearch
- Búsqueda semántica en Amazon Service OpenSearch
- Búsqueda simultánea de segmentos en Amazon Service OpenSearch
- Generación de consultas en lenguaje natural en Amazon OpenSearch Service

Búsquedas de URI

Las búsquedas de identificador de recursos universal (URI) son la forma más sencilla de búsqueda. En una búsqueda de URI, hay que especificar la consulta como un parámetro de solicitud HTTP:

```
GET https://search-my-domain.us-west-1.es.amazonaws.com/_search?q=house
```

Una respuesta de muestra tendría el siguiente aspecto:

```
{
  "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"
        }
      },
      ...
    ]
  }
}
```

De forma predeterminada, esta consulta realiza búsquedas en todos los campos de todos los índices para encontrar el término house. Para aportar más precisión a la búsqueda, especifique un índice (`movies`) y un campo de documento (`title`) en el URI:

```
GET https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search?q=title:house
```

Puede incluir parámetros adicionales en la solicitud, pero los parámetros admitidos proporcionan solo un pequeño subconjunto de las opciones de OpenSearch búsqueda. La siguiente solicitud devuelve 20 resultados (en vez de los 10 predeterminados) y los clasifica por año (en vez de clasificarlos por `_score`):

```
GET https://search-my-domain.us-west-1.es.amazonaws.com/movies/_search?
q=title:house&size=20&sort=year:desc
```

Búsquedas de cuerpo de la solicitud

Para realizar búsquedas más complejas, utilice el cuerpo de la solicitud HTTP y el lenguaje específico de dominio (DSL) de OpenSearch para las consultas. El DSL de la consulta le permite especificar la gama completa de opciones de búsqueda de OpenSearch .

> **Note**
>
> No puede incluir caracteres especiales Unicode en el valor de un campo de texto, o el valor se analizará como varios valores separados por el carácter especial. Este análisis incorrecto puede provocar un filtrado involuntario de los documentos y comprometer potencialmente el control sobre su acceso. Para obtener más información, consulte una nota sobre los caracteres especiales de Unicode en los campos de texto en la OpenSearch documentación.

La siguiente consulta de `match` es similar al ejemplo de búsqueda de URI final:

```
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**
>
> La API de `_search` acepta `GET` y `POST` HTTP para búsquedas del cuerpo de la solicitud, pero no todos los clientes HTTP admiten la adición de un cuerpo de la solicitud a una solicitud `GET`. `POST` es la opción más universal.

En muchos casos, es posible que desee buscar varios campos, pero no todos los campos. Utilice la consulta `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"]
    }
  }
}
```

Potenciar campos

Puede mejorar la relevancia de la búsqueda mediante la “potenciación” de ciertos campos. Los potenciadores son multiplicadores que dan mayor peso a las coincidencias en un campo que a las de otros campos. En el siguiente ejemplo, una coincidencia de john en el campo `title` tiene el doble de peso en `_score` que una coincidencia en el campo `plot` y cuatro veces más que otra en los campos `actors` o `directors`. El resultado es que películas como John Wick y John Carter están cerca de los primeros resultados de búsqueda y las películas protagonizadas por John Travolta están casi al final.

```
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"]
    }
  }
}
```

Resaltado de resultados de búsqueda

La `highlight` opción indica OpenSearch que hay que devolver un objeto adicional dentro de la `hits` matriz si la consulta coincide con uno o más campos:

```
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": {}
    }
  }
}
```

Si la consulta coincide con el contenido del campo `plot`, un acierto podría tener el siguiente aspecto:

```
{
  "_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>."
    ]
  }
}
```

De forma predeterminada, OpenSearch agrupa la cadena coincidente en etiquetas, proporciona hasta 100 caracteres de contexto en torno a la coincidencia y divide el contenido en oraciones identificando los signos de puntuación, los espacios, las tabulaciones y los saltos de línea. Todas estas configuraciones son personalizables:

```
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 de recuento

Si no está interesado en el contenido de sus documentos y solo quiere saber el número de coincidencias, puede utilizar la API de `_count` en lugar de la API de `_search`. En la siguiente solicitud, se utiliza la consulta `query_string` para identificar comedias románticas:

```
POST https://search-my-domain.us-west-1.es.amazonaws.com/movies/_count
{
  "query": {
    "query_string": {
      "default_field": "genres",
```

```
      "query": "romance AND comedy"
    }
  }
}
```

Una respuesta de muestra tendría el siguiente aspecto:

```
{
  "count": 564,
  "_shards": {
    "total": 5,
    "successful": 5,
    "skipped": 0,
    "failed": 0
  }
}
```

Paginar los resultados de búsqueda

Si necesita mostrar un gran número de resultados de búsqueda, puede implementar la paginación usando varios métodos diferentes.

Punto en el tiempo

La característica de punto en el tiempo (PIT) es un tipo de búsqueda que permite ejecutar diferentes consultas en un conjunto de datos fijo en el tiempo. Este es el método de paginación preferido OpenSearch, especialmente para la paginación profunda. Puede usar PIT con la versión 2.5 y posteriores del OpenSearch servicio. Para más información sobre PIT, consulte ???.

Los parámetros **`from`** y **`size`**

La forma más sencilla de paginar es con los parámetros `from` y `size`. La siguiente solicitud devuelve los resultados 20-39 de la lista de resultados de búsqueda indexada a cero:

```
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"]
    }
  }
}
```

Para obtener más información sobre la paginación de búsquedas, consulte Paginar los resultados en la documentación. OpenSearch

Lenguaje de consulta de paneles

Puede usar el lenguaje de consulta de paneles (DQL) para buscar datos y visualizaciones en los paneles. OpenSearch DQL utiliza cuatro tipos de consultas principales: términos, booleano, fecha e intervalo y campo anidado.

Consulta de términos

Una consulta de términos requiere que se especifique el término que se busca.

Para llevar a cabo una consulta de términos, ingrese lo siguiente:

```
host:www.example.com
```

Consulta booleana

Puede utilizar los operadores booleanos `AND`, `or` y `not` para combinar varias consultas.

Para llevar a cabo una consulta booleana, pegue lo siguiente:

```
host.keyword:www.example.com and response.keyword:200
```

Consulta de fecha e intervalo

Puede utilizar una consulta de fecha e intervalo para buscar una fecha anterior o posterior a la consulta.

- > indica una búsqueda de una fecha posterior a la fecha especificada.
- < indica una búsqueda de una fecha anterior a la fecha especificada.

```
@timestamp > "2020-12-14T09:35:33"
```

Consulta de campo anidada

Si tiene un documento con campos anidados, debe especificar qué partes del documento desea recuperar. A continuación, se muestra un documento de muestra con campos anidados:

```
{"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"
    }
  ]
}
```

Para recuperar un campo específico con DQL, pegue lo siguiente:

```
NBA players: {player-name: Lebron James}
```

Para recuperar varios objetos del documento anidado, pegue lo siguiente:

```
NBA players: {player-name: Lebron James} and NBA players: {player-name: Giannis
 Antetokounmpo}
```

Para buscar dentro de un intervalo, pegue lo siguiente:

```
NBA players: {player-name: Lebron James} and NBA players: {player-name: Giannis
 Antetokounmpo and < 30}
```

Si el documento tiene un objeto anidado dentro de otro objeto, aún puede recuperar datos especificando todos los niveles. Para ello, pegue lo siguiente:

```
Top-Power-forwards.NBA players: {player-name:Lebron James}
```

Importación y gestión de paquetes en Amazon OpenSearch Service

Amazon OpenSearch Service te permite subir archivos de diccionarios personalizados, como palabras vacías y sinónimos, y asociar complementos a tu dominio. Estos complementos pueden ser preempaquetados, personalizados o de terceros, lo que le brinda flexibilidad para ampliar la funcionalidad de su dominio. El término genérico para referirse a todos estos tipos de archivos es paquetes.

- Los archivos de diccionario ayudan a refinar los resultados de búsqueda, ya OpenSearch que indican que hay que ignorar las palabras más frecuentes o tratar términos similares, como «crema congelada», «helado» o «helado», como equivalentes. También pueden mejorar la derivación, como se observa con el complemento de análisis japonés (kuromoji).
- Los complementos preempaquetados ofrecen funciones integradas, como el complemento Amazon Personalize para obtener resultados de búsqueda personalizados. Estos complementos utilizan el tipo de paquete `ZIP-PLUGIN`. Para obtener más información, consulte the section called “Complementos por versión de motor”.
- Los complementos personalizados y de terceros le permiten agregar características personalizadas o integrarlos con sistemas externos, lo que ofrece aún más flexibilidad para su dominio. Al igual que los complementos preempaquetados, sube los complementos personalizados como paquetes `ZIP-PLUGIN`. En el caso de los complementos de terceros, también debe importar los archivos de licencia y configuración del complemento como paquetes independientes y, a continuación, asociarlos todos al dominio.

  Para obtener más información, consulte los temas siguientes:

  - the section called “Plugins personalizados”
  - the section called “Complementos de terceros”

> **Note**
>
> Puede asociar un máximo de 20 complementos con un único dominio. Este límite incluye todos los tipos de complementos: opcionales, de terceros y personalizados.

Temas

- Permisos necesarios
- Carga de paquetes en Amazon S3
- Importación y asociación de paquetes
- Uso de paquetes con OpenSearch
- Actualización de paquetes
- Actualización manual de los índices con un nuevo diccionario
- Disociación y eliminación de paquetes
- Gestión de complementos personalizados en Amazon OpenSearch Service
- Instalación de complementos de terceros en Amazon OpenSearch Service

Permisos necesarios

Los usuarios sin acceso de administrador requieren ciertas acciones AWS Identity and Access Management (IAM) para poder administrar los paquetes:

- `es:CreatePackage`: crear un paquete
- `es:DeletePackage`: eliminar un paquete
- `es:AssociatePackage`: asociar un paquete a un dominio
- `es:DissociatePackage`: disociar un paquete de un dominio

También necesita permisos en la ruta del bucket de Amazon S3 u objeto donde reside el paquete personalizado.

Otorgue todos los permisos en IAM, no en la política de acceso al dominio. Para más información, consulte the section called “Gestión de identidad y acceso”.

Carga de paquetes en Amazon S3

Esta sección explica cómo cargar paquetes de diccionarios personalizados, dado que los paquetes de complementos prediseñados ya están instalados. Antes de poder asociar un diccionario personalizado a un dominio, debe cargarlo en un bucket de Amazon S3. Para ver las instrucciones, consulte Carga de objetos en la Guía del usuario de Amazon Simple Storage Service. No es necesario cargar los complementos admitidos.

Si su diccionario contiene información confidencial, especifique el cifrado del lado del servidor con claves administradas por S3 cuando lo cargue. OpenSearch El servicio no puede acceder a los archivos de S3 que protejas con una clave. AWS KMS

Después de cargar el archivo, tome nota de su ruta de S3. El formato de la ruta es `s3://amzn-s3-demo-bucket/file-path/file-name`.

Puede utilizar el siguiente archivo de sinónimos para fines de prueba. Guárdelo como `synonyms.txt`.

```
danish, croissant, pastry
ice cream, gelato, frozen custard
sneaker, tennis shoe, running shoe
basketball shoe, hightop
```

Algunos diccionarios, como los diccionarios Hunspell, utilizan varios archivos y requieren sus propios directorios en el sistema de archivos. En este momento, el OpenSearch servicio solo admite diccionarios de un solo archivo.

Importación y asociación de paquetes

La consola es la forma más sencilla de importar un diccionario personalizado a OpenSearch Service. Al importar un diccionario de Amazon S3, OpenSearch Service almacena su propia copia del paquete y la cifra automáticamente mediante AES-256 con OpenSearch claves administradas por el servicio.

Los complementos opcionales ya vienen preinstalados en el OpenSearch Servicio, por lo que no es necesario que los cargue usted mismo, pero sí que tiene que asociar un complemento a un dominio. Los complementos disponibles aparecen en la pantalla Paquetes de la consola.

Para importar y asociar un paquete a un dominio

1. En la consola OpenSearch de Amazon Service, selecciona Paquetes.
2. Elija Import package (Importar paquete).
3. Asigne al paquete un nombre descriptivo.
4. Proporcione la ruta de S3 al archivo y, a continuación, elija Importar.
5. Vuelva a la pantalla Paquetes.
6. Cuando el estado del paquete sea Disponible, elíjalo.

7. Seleccione Asociar a un dominio.
8. Seleccione un dominio y, a continuación, elija Siguiente. Revise los paquetes y seleccione Asociar.
9. En el panel de navegación, seleccione el dominio y, a continuación, vaya a la pestaña Paquetes.
10. Si el paquete es un diccionario personalizado, anote el ID cuando el paquete esté disponible. `analyzers/`*`id`*Utilízala como ruta del archivo en las solicitudes para OpenSearch.

Uso de paquetes con OpenSearch

Esta sección explica cómo usar ambos tipos de paquetes: diccionarios personalizados y complementos prediseñados.

Uso de diccionarios personalizados

Después de asociar un archivo a un dominio, puede usarlo en parámetros como `synonyms_path`, `stopwords_path` y `user_dictionary` al crear generadores de tokens y filtros de token. El parámetro exacto varía según el objeto. Varios objetos admiten `synonyms_path` y `stopwords_path`, pero `user_dictionary` es exclusivo del complemento kuromoji.

Para el complemento IK (chino) Analysis, puede cargar un archivo de diccionario personalizado en forma de paquete personalizado y asociarlo a un dominio, y el complemento lo recogerá automáticamente sin necesidad de parámetro `user_dictionary`. Si se trata de un archivo de sinónimos, utilice el parámetro `synonyms_path`.

El siguiente ejemplo agrega un archivo de sinónimos a un nuevo índice:

```
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"
      }
    }
  }
}
```

Esta solicitud crea un analizador personalizado para el índice que utiliza el generador de tokens estándar y un filtro de token de sinónimos.

- Los generadores de tokens dividen las secuencias de caracteres en tokens (normalmente palabras) en función de algún conjunto de reglas. El ejemplo más simple es el generador de tokens de espacios en blanco, que divide los caracteres anteriores en un token cada vez que encuentra un carácter de espacio en blanco. Un ejemplo más complejo es el generador de tokens estándar, que utiliza un conjunto de reglas gramaticales para funcionar en muchos idiomas.
- Los filtros de tokens agregan, modifican o eliminan tokens. Por ejemplo, un filtro de token de sinónimos agrega tokens cuando encuentra una palabra en la lista de sinónimos. El filtro de tokens de palabras vacías elimina los tokens cuando encuentra una palabra en la lista de palabras vacías.

Esta solicitud también agrega un campo de texto (`description`) al mapeo e indica OpenSearch que se use el nuevo analizador como analizador de búsqueda. Puede ver que todavía utiliza el analizador estándar como analizador de índices.

Por último, tenga en cuenta la línea `"updateable": true` en el filtro de token. Este campo solo se aplica a los analizadores de búsqueda, no a los analizadores de índice, y es crítico si luego desea actualizar el analizador de búsqueda de forma automática.

Con fines de prueba, agregue algunos documentos al índice:

```
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" }
```

A continuación, use un sinónimo para buscarlos:

```
GET my-index/_search
{
  "query": {
    "match": {
      "description": "gelato"
    }
  }
}
```

En este caso, OpenSearch devuelve la siguiente respuesta:

```
{
  "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**
>
> Los archivos de diccionario utilizan espacio de montón de Java proporcional a su tamaño. Por ejemplo, un archivo de diccionario de 2 GiB podría consumir 2 GiB de espacio de montón en un nodo. Si utiliza archivos de gran tamaño, asegúrese de que los nodos tengan suficiente espacio en el montón para acomodarlos. Monitoree la métrica `JVMMemoryPressure` y escale el clúster según sea necesario.

Uso de complementos preempaquetados

OpenSearch El servicio le permite asociar OpenSearch complementos opcionales preinstalados para usarlos con su dominio. Un paquete de complementos preempaquetado es compatible con una OpenSearch versión específica y solo se puede asociar a los dominios con esa versión. La lista de paquetes disponibles para su dominio incluye todos los complementos compatibles con la versión de su dominio. Después de asociar un complemento a un dominio, comienza un proceso de instalación en el dominio. Luego, puedes hacer referencia al complemento y usarlo cuando realices solicitudes al OpenSearch Servicio.

La asociación y disociación de un complemento requiere una implementación. blue/green Para obtener más información, consulte the section called “Cambios que suelen provocar blue/green despliegues”.

Los complementos opcionales incluyen analizadores de idioma y resultados de búsqueda personalizados. Por ejemplo, el complemento Amazon Personalize Search Ranking utiliza machine learning para personalizar los resultados de búsqueda para sus clientes. Para obtener más información sobre este complemento, consulte Personalización de los resultados de búsqueda desde. OpenSearch Para ver una lista de todos los complementos admitidos, consulte the section called “Complementos por versión de motor”.

Complemento Sudachi

En el caso del complemento Sudachi, al volver a asociar un archivo de diccionario, no se refleja inmediatamente en el dominio. El diccionario se actualiza cuando se ejecuta la siguiente blue/green implementación en el dominio como parte de un cambio de configuración u otra actualización. Como alternativa, puede crear un nuevo paquete con los datos actualizados, crear un nuevo índice con un

nuevo paquete, volver a indexar el índice existente en el nuevo índice y, a continuación, eliminar el índice anterior. Si prefiere utilizar el enfoque de reindexación, utilice un alias de índice para que no se interrumpa el tráfico.

Además, el complemento Sudachi solo admite los diccionarios binarios de Sudachi, que se pueden cargar con la operación de la CreatePackageAPI. Para obtener información sobre el diccionario del sistema prediseñado y el proceso de compilación de diccionarios de usuario, consulte Documentación de Sudachi.

> **Note**
>
> Al cargar archivos de diccionarios binarios a Amazon S3, debe establecer el tipo de contenido del objeto S3 en. `binary/octet-stream` Si `application/octet-stream` se utiliza, se producirá un error en la importación del paquete.

En el siguiente ejemplo, se muestra cómo utilizar los diccionarios de sistema y de usuario con el tokenizador de Sudachi. Debe cargar estos diccionarios como paquetes personalizados con el tipo `TXT-DICTIONARY` y proporcionar su paquete IDs en la configuración adicional.

```
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"
          }
        }
      }
    }
  }
}
```

Actualización de paquetes

En esta sección solo se explica cómo actualizar un paquete de diccionarios personalizados, ya que los paquetes de complementos preempaquetados ya están actualizados. Al cargar una nueva versión de un diccionario en Amazon S3, el paquete no se actualiza automáticamente en Amazon OpenSearch Service. OpenSearch El servicio almacena su propia copia del archivo, por lo que si subes una nueva versión a S3, tendrás que actualizarla manualmente.

Cada uno de los dominios asociados también almacena su propia copia del archivo. Para mantener el comportamiento de búsqueda predecible, los dominios siguen utilizando la versión actual del paquete hasta que los actualice explícitamente. Para actualizar un paquete personalizado, modifique el archivo Amazon S3 Control, actualice el paquete en OpenSearch Service y, a continuación, aplique la actualización.

Consola

1. En la consola OpenSearch de servicio, selecciona Paquetes.
2. Seleccione un paquete y luego Actualizar.
3. Proporcione una nueva ruta de S3 al archivo y, a continuación, seleccione Actualizar paquete.
4. Vuelva a la pantalla Paquetes.
5. Cuando el estado del paquete cambie a Disponible, selecciónelo. A continuación, seleccione uno o más dominios asociados, Aplicar actualización y confirme. Espere a que el estado de asociación cambie a Activo.
6. Los siguientes pasos variarán en función de cómo haya configurado los índices:

   - Si tu dominio ejecuta OpenSearch Elasticsearch 7.8 o una versión posterior y solo usa analizadores de búsqueda con el campo actualizable establecido en true, no necesitas realizar ninguna otra acción. OpenSearch El servicio actualiza automáticamente tus índices mediante la API _plugins/_refresh_search_analyzers.

- Si el dominio ejecuta Elasticsearch 7.7 o una versión anterior, utiliza analizadores de índice o no utiliza el campo `updateable`, consulte the section called “Actualización manual de los índices con un nuevo diccionario”.

Si bien la consola es el método más sencillo, también puedes usar la API, o la API de configuración para actualizar los paquetes de servicio. AWS CLI SDKs OpenSearch Para obtener más información, consulta la Referencia de AWS CLI comandos y la Referencia de la API de Amazon OpenSearch Service.

AWS SDK

En lugar de actualizar manualmente un paquete en la consola, puede utilizar el SDKs para automatizar el proceso de actualización. El siguiente script de Python de ejemplo carga un nuevo archivo de paquete en Amazon S3, actualiza el paquete en OpenSearch Service y aplica el nuevo paquete al dominio especificado. Tras confirmar que la actualización se ha realizado correctamente, realiza un ejemplo de llamada para OpenSearch demostrar que se han aplicado los nuevos sinónimos.

Debe proporcionar valores para `host`, `region`, `file_name`, `bucket_name`, `s3_key`, `package_id`, `domain_name` y `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**
>
> Si recibes un error que indica que no se ha encontrado el paquete al ejecutar el script AWS CLI, es probable que Boto3 esté utilizando la región especificada en ~/.aws/config, que no es la región en la que se encuentra tu bucket de S3. Puede ejecutar `aws configure` y especificar la región correcta o agregar explícitamente la región al cliente:
>
> ```
> client = boto3.client('opensearch', region_name='us-east-1')
> ```

Actualización manual de los índices con un nuevo diccionario

Las actualizaciones manuales de índices solo se aplican a los diccionarios personalizados, no a los complementos preempaquetados. Para utilizar un diccionario actualizado, debe actualizar de forma manual los índices, si cumple alguna de las siguientes condiciones:

- El dominio ejecuta Elasticsearch 7.7 o anterior.
- Utiliza paquetes personalizados como analizadores de índices.
- Utiliza paquetes personalizados como analizadores de búsqueda, pero no incluye el campo actualizable.

Para actualizar los analizadores con los nuevos archivos de paquete, dispone de dos opciones:

- Cierre y abra los índices que desee actualizar:

```
POST my-index/_close
POST my-index/_open
```

- Vuelva a indexar los índices. Primero, cree un índice que utilice el archivo de sinónimos actualizado (o un archivo completamente nuevo). Tenga en cuenta que solo se admite 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"
      }
    }
  }
}
```

Luego vuelva a indexar el índice anterior con ese nuevo índice:

```
POST _reindex
{
  "source": {
    "index": "my-index"
  },
  "dest": {
    "index": "my-new-index"
  }
}
```

Si actualiza con frecuencia los analizadores de índices, utilice alias de índice para mantener una ruta coherente al índice más reciente:

```
POST _aliases
{
  "actions": [
    {
      "remove": {
        "index": "my-index",
        "alias": "latest-index"
      }
    },
    {
      "add": {
        "index": "my-new-index",
        "alias": "latest-index"
      }
    }
  ]
}
```

Si no necesita el índice anterior, elimínelo:

```
DELETE my-index
```

Disociación y eliminación de paquetes

Si disocia un paquete de un dominio, ya sea un diccionario personalizado o un complemento preempaquetado, ya no podrá usar ese paquete cuando cree nuevos índices. Una vez disociado

un paquete, los índices existentes que utilizaban el paquete ya no pueden usarlo. Debe eliminar el paquete de cualquier índice antes de poder disociarlo; de lo contrario, la disociación fallará.

La consola es la forma más sencilla de disociar un paquete de un dominio y eliminarlo del Servicio. OpenSearch Al eliminar un paquete del OpenSearch Servicio, no se elimina de su ubicación original en Amazon S3.

Disociar un paquete de un dominio

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación, elija Dominios.
3. Elija el dominio y, a continuación, vaya a la pestaña Paquetes.
4. Seleccione un paquete, Acciones y, a continuación, seleccione Disociar. Confirme su elección.
5. Espere a que el paquete desaparezca de la lista. Puede que tenga que actualizar el navegador.
6. Si desea utilizar el paquete con otros dominios, deténgase aquí. Para continuar con la eliminación del paquete (si es un diccionario personalizado), seleccione Paquetes en el panel de navegación.
7. Seleccione el paquete y elija Eliminar.

También puede utilizar la API de configuración o la AWS CLI API de configuración para disociar y eliminar paquetes. SDKs Para obtener más información, consulta la Referencia de AWS CLI comandos y la Referencia de la API de Amazon OpenSearch Service.

Gestión de complementos personalizados en Amazon OpenSearch Service

Al usar los complementos personalizados para OpenSearch Service, puedes ampliar la OpenSearch funcionalidad en áreas como el análisis del idioma, el filtrado personalizado, la clasificación y más, lo que te permite crear experiencias de búsqueda personalizadas. OpenSearch Se pueden desarrollar complementos personalizados ampliando la `org.opensearch.plugins.Plugin` clase y luego empaquetándola en un `.zip` archivo.

Amazon OpenSearch Service admite actualmente las siguientes extensiones de complementos:

- AnalysisPlugin— Amplía la funcionalidad de análisis añadiendo, por ejemplo, analizadores personalizados, tokenizadores de caracteres o filtros para el procesamiento de texto.

- SearchPlugin— Mejora las capacidades de búsqueda con tipos de consultas personalizados, algoritmos de similitud, opciones de sugerencias y agregaciones.
- MapperPlugin— Le permite crear tipos de campos personalizados y sus configuraciones de mapeo OpenSearch, lo que le permite definir cómo se deben almacenar e indexar los diferentes tipos de datos.
- ScriptPlugin— Permite añadir funciones de creación de secuencias de comandos personalizadas OpenSearch, por ejemplo, a las secuencias de comandos personalizadas para operaciones como la puntuación, la clasificación y la transformación de los valores de los campos durante la búsqueda o la indexación.

Puedes usar la consola de OpenSearch servicio o los comandos de API existentes para paquetes personalizados para cargar y asociar el complemento al clúster de Amazon OpenSearch Service. También puedes usar el DescribePackagescomando para describir todos los paquetes de tu cuenta y ver detalles como la OpenSearch versión y los detalles del error. OpenSearch El servicio valida el paquete de complementos para comprobar la compatibilidad de versiones, las vulnerabilidades de seguridad y las operaciones de complementos permitidas. Para obtener más información sobre los paquetes personalizados, consulte the section called “Paquetes”.

OpenSearch versión y soporte Región de AWS

Los siguientes Regiones de AWS dominios de OpenSearch servicio que ejecutan la OpenSearch versión 2.15 admiten complementos personalizados:

- Este de EE. UU. (Ohio) (us-east-2)
- Este de EE. UU. (Norte de Virginia) (us-east-1)
- Oeste de EE. UU. (Oregón) (us-west-2)
- Asia-Pacífico (Mumbai) (ap-south-1)
- Asia-Pacífico (Seúl) (ap-northeast-2)
- Asia-Pacífico (Singapur) (ap-southeast-1)
- Asia-Pacífico (Sídney) (ap-southeast-2)
- Asia-Pacífico (Tokio) (ap-northeast-1)
- Canadá (centro) (ca-central-1)
- Europa (Fráncfort) (eu-central-1)
- Europa (Irlanda) (eu-west-1)
- Europa (Londres) (eu-west-2)

- Europa (París) (eu-west-3)
- América del Sur (São Paulo) (sa-east-1)

> **Note**
>
> Los complementos personalizados contienen código desarrollado por el usuario. Cualquier problema, incluido el incumplimiento del SLA, causado por un código desarrollado por el usuario no da derecho a los créditos del SLA. Para obtener más información, consulta Amazon OpenSearch Service: Acuerdo de nivel de servicio.

Temas

- Cuotas de complementos
- Requisitos previos
- Resolución de problemas
- Instalación de un complemento personalizado mediante la consola
- Administrar complementos personalizados mediante el AWS CLI
- AWS KMS Integración de paquetes personalizados de Amazon OpenSearch Service

Cuotas de complementos

- Puede crear hasta 25 complementos personalizados por región para cada cuenta.
- El tamaño máximo sin comprimir de un complemento es 1 GB.
- El número máximo de complementos que se pueden asociar con un solo dominio es 20. Esta cuota se aplica a todos los tipos de complementos combinados: opcionales, de terceros y personalizados.
- Los complementos personalizados son compatibles con los dominios que ejecutan OpenSearch la versión 2.15 o posterior.
- El archivo `descriptor.properties` para su complemento debe ser compatible con una versión de motor similar a 2.15.0 o cualquier versión 2.x.x, donde la versión de parche esté establecida en cero.

Requisitos previos

Antes de instalar un complemento personalizado y asociarlo a un dominio, asegúrese de que cumple los siguientes requisitos:

- La versión de motor compatible con el complemento del archivo `descriptor.properties` debe ser similar a `2.15.0` o `2.x.0`. Es decir, la versión del parche debe ser cero.
- Las siguientes características deben estar habilitadas en su dominio:
  - Node-to-node cifrado
  - Cifrado en reposo
  - `EnforceHTTPS` está establecido en “verdadero”

    Consulte también opensearch-https-requireden la Guía para AWS Config desarrolladores.
  - Los clientes deben admitir Policy-Min-TLS-1-2-PFS-2023-10. Puede especificar esta compatibilidad usando el siguiente comando. Sustitúyalo por su propia información: *`placeholder value`*

    ```
    aws opensearch update-domain-config \
        --domain-name domain-name \
        --domain-endpoint-options '{"TLSSecurityPolicy":"Policy-Min-TLS-1-2-
    PFS-2023-10" }'
    ```

    Para obtener más información, consulta DomainEndpointOptionsla referencia de la API OpenSearch de Amazon Service.

Resolución de problemas

Si el sistema devuelve el error `PluginValidationFailureReason : The provided plugin could not be loaded`, consulte the section called “La instalación del complemento personalizado falla debido a la compatibilidad de las versiones” para obtener información de solución de problemas.

Instalación de un complemento personalizado mediante la consola

Para asociar un complemento de terceros a un dominio, primero importe la licencia y la configuración del complemento como paquetes.

Para instalar un complemento personalizado

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, elija Paquetes.
3. Seleccione Importar paquete.
4. En Nombre, ingrese un nombre único y fácil de identificar para el complemento.
5. (Opcional) En la descripción, proporciona cualquier detalle útil sobre el paquete o su propósito.
6. En Tipo de paquete, elija Complemento.
7. En Origen del paquete, ingrese la ruta o busque el archivo ZIP del complemento en Amazon S3.
8. Para la versión OpenSearch del motor, elige la OpenSearch versión compatible con el plugin.
9. Para el cifrado de paquetes, elija si desea personalizar la clave de cifrado del paquete. De forma predeterminada, OpenSearch Service cifra el paquete de complementos con un Clave propiedad de AWS. En su lugar, puede usar una clave administrada por el cliente.
10. Seleccione Importar.

Después de importar el paquete de complementos, asócielo a un dominio. Para obtener instrucciones, consulte the section called “Para importar y asociar un paquete a un dominio”.

Administrar complementos personalizados mediante el AWS CLI

Puedes utilizar el AWS CLI para gestionar una serie de tareas de complementos personalizados.

Tareas

- Instalación de un complemento personalizado mediante el AWS CLI
- Actualización de un complemento personalizado mediante el AWS CLI
- Crea o actualiza un complemento personalizado con una AWS KMS clave de seguridad
- Actualizar un dominio de OpenSearch servicio con complementos personalizados a una versión posterior OpenSearch mediante el AWS CLI
- Desinstalar y ver el estado de disociación de un complemento personalizado

Instalación de un complemento personalizado mediante el AWS CLI

Antes de empezar

Antes de poder asociar un complemento personalizado con su dominio, debe cargarlo en un bucket de Amazon Simple Storage Service (Amazon S3). El depósito debe estar ubicado en el mismo Región de AWS lugar en el que pretendes utilizar el complemento. Para obtener información sobre cómo agregar un objeto a un bucket de S3, consulte Cargar objetos en la Guía del usuario de Amazon Simple Storage Service.

Si su complemento contiene información sensible, especifique el cifrado del lado del servidor con claves administradas por S3 al cargarlo. Después de cargar el archivo, tome nota de su ruta de S3. El formato de la ruta es `s3://amzn-s3-demo-bucket/file-path/file-name`.

> **Note**
>
> Si lo desea, puede proteger un complemento personalizado al crear el complemento especificando una clave AWS Key Management Service (AWS KMS). Para obtener información, consulte the section called “Crea o actualiza un complemento personalizado con una AWS KMS clave de seguridad”.

Para instalar un complemento personalizado mediante el AWS CLI

1. Cree un nuevo paquete para su complemento personalizado ejecutando el siguiente comando create-package, asegurándose de cumplir con los siguientes requisitos:

   - La ubicación del bucket y la clave debe apuntar al archivo `.zip` del complemento en un bucket de S3 de la cuenta en la que está ejecutando los comandos.
   - El bucket de S3 debe estar en la misma región en la que se está creando el paquete.
   - Solo se admiten archivos `.zip` para los paquetes `ZIP-PLUGIN`.
   - El contenido del archivo `.zip` debe seguir la estructura de directorios esperada por el complemento.
   - El valor de `--engine-version` debe estar en el formato `OpenSearch_{MAJOR}.{MINOR}`. Por ejemplo: **`OpenSearch_2.17`**.

   Sustitúyalo por tu propia información: *`placeholder values`*

   ```
   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. (Opcional) Vea el estado de la operación `create-package`, incluidos los resultados de validación y las vulnerabilidades de seguridad, usando el comando describe-packages. Sustitúyalo por su propia información: *placeholder values*

```
aws opensearch describe-packages \
    --region region  \
    --filters '[{"Name": "PackageType","Value": ["ZIP-PLUGIN"]}, {"Name":
 "PackageName","Value": ["package-name"]}]'
```

El comando devuelve información similar a la siguiente:

```
{
    "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 la `create-package` operación, Amazon OpenSearch Service comprueba el `ZIP-PLUGIN` valor para comprobar la compatibilidad de versiones, las extensiones de complementos compatibles y las vulnerabilidades de seguridad. Las vulnerabilidades de

seguridad se escanean mediante el servicio Amazon Inspector. Los resultados de estas verificaciones se muestran en el campo `ErrorDetails` de la respuesta de la API.

3. Usa el comando associate-package para asociar el complemento al dominio de OpenSearch servicio que elijas utilizando el ID de paquete del paquete creado en el paso anterior.

   > **Tip**
   >
   > Si tiene varios complementos, puede usar el comando associate-packages para asociar múltiples paquetes a un dominio en una sola operación.

   Sustitúyalo *`placeholder values`* por tu propia información:

   ```
   aws opensearch associate-package \
       --domain-name domain-name \
       --region region \
       --package-id package-id
   ```

   > **Note**
   >
   > El complemento se instala y desinstala utilizando un proceso de implementación azul/verde.

4. (Opcional) Utilice el list-packages-for-domaincomando para ver el estado de la asociación. El estado de la asociación cambia a medida que el flujo de trabajo avanza de `ASSOCIATING` a `ACTIVE`. El estado de la asociación cambia a ACTIVE cuando se completa la instalación del complemento y el complemento está listo para su uso.

   Sustituya *`placeholder values`* por su propia información.

   ```
   aws opensearch list-packages-for-domain \
       --region region \
       --domain-name domain-name
   ```

Actualización de un complemento personalizado mediante el AWS CLI

Use el comando update-package para realizar cambios en un complemento.

> **Note**
>
> Si lo desea, puede proteger un complemento personalizado al actualizar el complemento especificando una clave AWS Key Management Service (AWS KMS). Para obtener información, consulte the section called “Crea o actualiza un complemento personalizado con una AWS KMS clave de seguridad”.

Para actualizar un complemento personalizado mediante el AWS CLI

- Ejecute el comando siguiente. Sustituya *placeholder values* por su propia información.

```
aws opensearch update-package \
    --region region \
    --package-id package-id \
    --package-source S3BucketName=amzn-s3-demo-bucket,S3Key=s3-key \
    --package-description description
```

Después de actualizar un paquete, puede usar el comando associate-package o associate-packages para aplicar las actualizaciones del paquete a un dominio.

> **Note**
>
> Puede auditar, crear, actualizar, asociar y desasociar operaciones en el complemento mediante AWS CloudTrail. Para obtener más información, consulte Supervisión de las llamadas a la API de Amazon OpenSearch Service con AWS CloudTrail.

Crea o actualiza un complemento personalizado con una AWS KMS clave de seguridad

Puede proteger un complemento personalizado al crearlo o actualizarlo especificando una clave AWS KMS . Para lograr esto, configure `PackageEncryptionOptions` en `true` y especifique el Nombre de recurso de Amazon (ARN) de la clave, como se muestra en los siguientes ejemplos.

Ejemplo: crea un complemento personalizado con una AWS KMS clave de seguridad

```
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"
  }
```

Ejemplo: actualizar un complemento personalizado con AWS KMS claves de seguridad

```
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

Si la AWS KMS clave que especificas está deshabilitada o eliminada, el clúster asociado puede dejar inoperativo.

Para obtener más información sobre AWS KMS la integración con paquetes personalizados,the section called “AWS KMS Integración de paquetes personalizados de Amazon OpenSearch Service”.

Actualizar un dominio de OpenSearch servicio con complementos personalizados a una versión posterior OpenSearch mediante el AWS CLI

Cuando necesite actualizar un dominio de OpenSearch servicio que utiliza complementos personalizados a una versión posterior de OpenSearch, complete los siguientes procesos.

Para actualizar un dominio de OpenSearch servicio con complementos personalizados a una versión posterior OpenSearch mediante el AWS CLI

1. Usa el comando create-package para crear un paquete nuevo para tu complemento especificando la nueva OpenSearch versión.

   Asegúrese de que el nombre del paquete sea el mismo para el complemento en todas las versiones del motor. Si se cambia el nombre del paquete, se produce un error en el proceso de actualización del dominio durante la blue/green implementación.
2. Actualice su dominio a una versión superior siguiendo los pasos que se indican en the section called “Actualización de dominios”.

   Durante este proceso, Amazon OpenSearch Service desasocia la versión anterior del paquete de complementos e instala la nueva versión mediante un blue/green despliegue.

Desinstalar y ver el estado de disociación de un complemento personalizado

Para desinstalar el complemento de cualquier dominio, puede usar el comando dissociate-package. Al ejecutar este comando, también se eliminan todos los paquetes de configuración o licencias relacionados. A continuación, puede utilizar el list-packages-for-domaincomando para ver el estado de la disociación.

> **Tip**
>
> También puede usar el comando dissociate-packages para desinstalar varios complementos de un dominio en una sola operación.

Desinstalar y ver el estado de disociación de un complemento personalizado

1. Deshabilita el complemento en todos los índices. Esto debe hacerse antes de disociar el paquete de complementos.

   Si intentas desinstalar un complemento antes de deshabilitarlo en todos los índices, el proceso de blue/green implementación permanece estancado en ese estado. `Processing`
2. Ejecute el siguiente comando para desinstalar el complemento. Sustituya *`placeholder values`* por su propia información.

```
aws opensearch dissociate-package \
```

```
    --region region \
    --package-id plugin-package-id \
    --domain-name domain name
```

3. (Opcional) Ejecuta el list-packages-for-domaincomando para ver el estado de la disociación.

AWS KMS Integración de paquetes personalizados de Amazon OpenSearch Service

Los paquetes personalizados de Amazon OpenSearch Service proporcionan cifrado de forma predeterminada para proteger tus `ZIP-PLUGIN` paquetes en reposo Claves administradas por AWS.

- Claves propiedad de AWS— Los paquetes personalizados de Amazon OpenSearch Service utilizan estas claves de forma predeterminada para cifrar automáticamente tus `ZIP-PLUGIN` paquetes. No puedes ver, gestionar, usar Claves propiedad de AWS ni auditar su uso. Sin embargo, no necesita realizar ninguna acción ni cambiar ningún programa para proteger las claves que cifran los datos. Para obtener más información, consulte Claves propiedad de AWS en la Guía para desarrolladores de AWS Key Management Service .
- Claves administradas por el cliente: puedes añadir una segunda capa de cifrado sobre la existente si eliges una clave administrada Claves propiedad de AWS por el cliente al crear tu paquete `ZIP-PLUGIN` personalizado.

  Los paquetes personalizados de Amazon OpenSearch Service admiten el uso de una clave simétrica gestionada por el cliente que tú creas, posees y gestionas para añadir una segunda capa de cifrado sobre el cifrado que ya AWS tienes. Debido a que usted tiene control total sobre esta capa de cifrado, puede realizar las siguientes tareas:

  - Establecer y mantener políticas de claves.
  - Establezca y mantenga políticas y subvenciones AWS Identity and Access Management (IAM)
  - Habilitar y deshabilitar políticas de claves.
  - Rotar el material criptográfico.
  - Agregar etiquetas.
  - Crear alias de claves.
  - Programar la eliminación de claves.

Para más información, consulte las claves administradas por el cliente en la Guía para desarrolladores de AWS Key Management Service .

> **Note**
>
> Los paquetes personalizados de Amazon OpenSearch Service permiten automáticamente el cifrado Claves propiedad de AWS en reposo sin coste alguno. Sin embargo, se aplican AWS KMS cargos cuando utilizas una clave gestionada por el cliente. Para obtener más información acerca de los precios, consulte Precios de AWS Key Management Service.

Cómo el OpenSearch servicio de paquetes personalizados de Amazon Service utiliza las subvenciones en AWS KMS

OpenSearch Los paquetes personalizados de Service requieren una autorización para usar tu clave gestionada por el cliente.

Cuando creas un `ZIP-PLUGIN` paquete cifrado con una clave gestionada por el cliente, el OpenSearch servicio de paquetes personalizados de Amazon Service crea una concesión en tu nombre enviando una CreateGrantsolicitud a AWS KMS. Otorga AWS KMS a OpenSearch In Give Service acceso a una AWS KMS clave de tu cuenta. Las concesiones creadas por los paquetes personalizados del OpenSearch Servicio tienen una restricción que permite realizar operaciones únicamente cuando la solicitud incluye un contexto de cifrado con el identificador del paquete personalizado.

Los paquetes personalizados de Amazon OpenSearch Service requieren la autorización para utilizar tu clave gestionada por el cliente en las siguientes operaciones internas:

Operación	Description (Descripción)
`DescribeKey`	Envía `DescribeKey` solicitudes AWS KMS para comprobar que el identificador de clave simétrico gestionado por el cliente introducido al crear el paquete de complementos es válido.
`GenerateDataKeyWithoutPlaintext`	Envía `GenerateDataKeyWithoutPlaintext` solicitudes AWS KMS para generar claves de datos cifradas por la clave gestionada por el cliente.
`GenerateDataKey`	Envía `GenerateDataKey` solicitudes AWS KMS para generar claves de datos para cifrar el paquete al copiarlo internamente.

Operación	Description (Descripción)
`Decrypt`	Envía `Decrypt` solicitudes AWS KMS para descifrar las claves de datos cifradas para que puedan usarse para descifrar los datos.

Puede revocar el acceso a la concesión o eliminar el acceso del servicio a la clave administrada por el cliente en cualquier momento. Si lo hace, el OpenSearch Servicio no podrá acceder a ningún dato cifrado por la clave gestionada por el cliente, lo que afectará a las operaciones que dependen de esos datos. Por ejemplo, si intentas asociar un paquete de complementos al que OpenSearch Service no puede acceder, la operación devuelve un `AccessDeniedException` error.

Creación de una clave administrada por el cliente

Puede crear una clave simétrica gestionada por el cliente mediante el Consola de administración de AWS o el AWS KMS APIs.

Para crear una clave simétrica administrada por el cliente

- Siga los pasos en Crear una clave KMS en la Guía para desarrolladores de AWS Key Management Service .

Política de claves

Las políticas de clave controlan el acceso a la clave administrada por el cliente. Cada clave administrada por el cliente debe tener exactamente una política de clave, que contiene instrucciones que determinan quién puede usar la clave y cómo puede utilizarla. Cuando crea la clave administrada por el cliente, puede especificar una política de clave. Para obtener más información, consulte Políticas de claves en AWS KMS en la Guía para desarrolladores de AWS Key Management Service .

Para usar su clave administrada por el cliente con los recursos de su complemento, debe permitir las siguientes operaciones de API en la política de claves:

- `kms:CreateGrant`: agrega una concesión a una clave administrada por el cliente. Otorga el acceso de control a una AWS KMS clave específica, lo que permite acceder a las operaciones de concesión que requieren los paquetes personalizados del OpenSearch Servicio. Para obtener más información sobre el uso de permisos, consulte la Guía para desarrolladores de AWS KMS.

  Esto permite al OpenSearch Servicio hacer lo siguiente:

  - Llamar a `GenerateDataKeyWithoutPlainText` para generar una clave de datos cifrada y almacenarla para validaciones posteriores.
  - Llamar a `GenerateDataKey` para copiar internamente el paquete del complemento.
  - Llamar a `Decrypt` para acceder internamente al paquete del complemento.
  - Configurar una entidad principal que se retire para permitir que el servicio `RetireGrant`.
- `kms:DescribeKey`— Proporciona los detalles clave gestionados por el cliente para que el OpenSearch Servicio pueda validar la clave.
- `kms:GenerateDataKey`,`kms:GenerateDataKeyWithoutPlaintext`, `kms:Decrypt` — Da acceso a los paquetes personalizados del OpenSearch Servicio para utilizar estas operaciones en la concesión.

Los siguientes son ejemplos de declaraciones de políticas que puede añadir a los paquetes personalizados del OpenSearch Servicio:

```
"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"
      }
    }
  }
]
```

Para obtener más información sobre cómo especificar permisos en una política, consulte Políticas de clave en AWS KMS y en la Guía para desarrolladores de AWS Key Management Service .

Para obtener más información sobre la solución de problemas de acceso a las claves, consulte Solución de problemas de AWS KMS permisos en la Guía para AWS Key Management Service desarrolladores.

Especifica una clave gestionada por el cliente para los paquetes personalizados OpenSearch de Amazon Service

Puede especificar una clave administrada por el cliente como segunda capa de cifrado para sus paquetes `ZIP-PLUGIN`.

Al crear un paquete de complementos, puede especificar la clave de datos introduciendo un identificador de AWS KMS clave, que los paquetes personalizados del OpenSearch Servicio utilizan para cifrar el paquete de complementos.

AWS KMS ID de clave: identificador de clave para una clave gestionada por el AWS KMS cliente. Introduzca el ID de la clave, el ARN de la clave, el nombre de alias o el ARN del alias.

Contexto de cifrado de paquetes personalizados de Amazon OpenSearch Service

Un contexto de cifrado es un conjunto opcional de pares clave-valor que pueden contener información contextual adicional sobre los datos.

AWS KMS utiliza el contexto de cifrado como datos autenticados adicionales para respaldar el cifrado autenticado. Al incluir un contexto de cifrado en una solicitud de cifrado de datos, AWS KMS vincula el contexto de cifrado a los datos cifrados. Para descifrar los datos, debe incluir el mismo contexto de cifrado en la solicitud.

Contexto de cifrado de paquetes personalizados de Amazon OpenSearch Service

Los paquetes personalizados de Amazon OpenSearch Service utilizan el mismo contexto de cifrado en todas las operaciones AWS KMS criptográficas, donde la clave es `packageId` y el valor es el `package-id` del paquete de complementos.

Uso del contexto de cifrado para monitoreo

Cuando utiliza una clave simétrica administrada por el cliente para cifrar su paquete de complemento, puede usar el contexto de cifrado en los registros de auditoría y logs para identificar cómo se está utilizando la clave administrada por el cliente. El contexto de cifrado también aparece en los registros generados por AWS CloudTrail Amazon CloudWatch Logs.

Utilizar el contexto de cifrado para controlar el acceso a la clave administrada por el cliente

Puede utilizar el contexto de cifrado en políticas de claves y políticas de IAM como condiciones para controlar el acceso a su clave simétrica administrada por el cliente. Puede usar también una restricción de contexto de cifrado en una concesión.

OpenSearch Los paquetes personalizados de servicios utilizan una restricción del contexto de cifrado en las concesiones para controlar el acceso a la clave gestionada por el cliente en su cuenta o región. La restricción de concesión requiere que las operaciones que permite la concesión utilicen el contexto de cifrado especificado.

Los siguientes son ejemplos de declaraciones de política de claves para conceder acceso a una clave administrada por el cliente para un contexto de cifrado específico. La condición de esta declaración de política exige que las concesiones tengan una restricción de contexto de cifrado que especifique el contexto de cifrado.

```
{
    "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"
         }
    }
}
```

Servicio de supervisión de las claves de cifrado para paquetes OpenSearch personalizados

Cuando utilizas una clave gestionada por el AWS KMS cliente con los recursos del OpenSearch servicio de paquetes personalizados, puedes utilizar CloudTrail nuestros CloudWatch registros para realizar un seguimiento de las solicitudes a las que se envían los paquetes OpenSearch personalizados AWS KMS.

Más información

Los siguientes recursos proporcionan más información sobre cifrado de datos en reposo.

- Para obtener más información sobre los conceptos AWS KMS básicos, consulte AWS KMS keysla Guía para AWS Key Management Service desarrolladores.

- Para obtener más información sobre las mejores prácticas de seguridad AWS KMS, consulte la guía de orientación AWS prescriptiva para conocer las AWS Key Management Service mejores prácticas.

Instalación de complementos de terceros en Amazon OpenSearch Service

Amazon OpenSearch Service admite complementos de terceros de socios seleccionados. Estos complementos pueden mejorar su OpenSearch configuración con funciones adicionales, como analizadores personalizados, tokenizadores o capacidades de cifrado. Siga las instrucciones específicas de instalación y configuración proporcionadas por los desarrolladores externos para garantizar una integración adecuada con su OpenSearch dominio de servicio.

> **Note**
>
> Debe obtener y mantener licencias válidas directamente de los desarrolladores de terceros. Es posible que algunos proveedores no habiliten todos sus complementos Regiones de AWS, así que consulte con el proveedor del complemento para comprobar su disponibilidad.

Los siguientes complementos de terceros están disponibles para su uso con el OpenSearch Servicio:

- Complemento de cifrado Portal26 (Titanium-Lockbox): Utiliza cifrado certificado NIST FIPS 140-2 para cifrar datos mientras se indexan. Incluye soporte Bring Your Own Key (BYOK), que le permite administrar sus claves de cifrado para mayor seguridad. El complemento lo proporciona Portal26 y requiere la OpenSearch versión 2.15 o superior.
- Concordancia de nombres (RNI): Coincide nombres, organizaciones, direcciones y fechas en más de 24 idiomas, lo que mejora la seguridad y el cumplimiento normativo. El plugin lo proporciona Babel Street y requiere la OpenSearch versión 2.15 o superior.

Temas

- Requisitos previos
- Instalación de complementos de terceros
- Siguientes pasos

Requisitos previos

Antes de instalar un complemento de terceros, realice los siguientes pasos:

- Obtenga los archivos de configuración y licencia del complemento y cárguelos en un bucket de Amazon S3. El bucket debe estar en el mismo dominio que el dominio Región de AWS .
- Un complemento de terceros es un tipo de complemento personalizado. Asegúrese de que el dominio cumpla con los requisitos previos para complementos personalizados.

Instalación de complementos de terceros

Para asociar un complemento de terceros a un dominio de OpenSearch servicio, primero debe cargar tres paquetes independientes: el paquete de licencias, el paquete de configuración y el paquete de complementos.

- Paquete de licencia: incluye la información de licencia o metadatos asociados con el complemento, en formato .json o .xml.
- Paquete de configuración: contiene los archivos de configuración del complemento, los recursos de soporte y la configuración. Estos archivos definen cómo se comporta o se integra el complemento. OpenSearch
- El paquete del complemento contiene el binario del complemento compilado, que es el código ejecutable que OpenSearch se ejecuta. Este es el núcleo de la funcionalidad del complemento.

Después de cargar los paquetes, puede asociar el complemento y la licencia con un dominio compatible.

Consola

Para asociar un complemento de terceros a un dominio, primero importe la licencia y la configuración del complemento como paquetes.

Para instalar un complemento de terceros

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, elija Paquetes.
3. En primer lugar, importe el paquete de licencias. Seleccione Importar paquete.

4. En Tipo de paquete, elija Licencia.
5. En Origen del paquete, ingrese la ruta al archivo JSON o XML de la licencia en Amazon S3.
6. Seleccione Importar. El paquete aparece en la pestaña Licencias de la página Paquetes.
7. Ahora, importe la configuración del complemento. Seleccione nuevamente Importar paquete.
8. En Tipo de paquete, elija Configuración.
9. Para Origen del paquete, ingrese la ruta al archivo ZIP de la configuración del complemento en Amazon S3.
10. Seleccione Importar.
11. Por último, importe el propio complemento. Seleccione Importar paquete.
12. En Tipo de paquete, elija Complemento.
13. En Origen del paquete, ingrese la ruta del archivo ZIP del complemento en Amazon S3.
14. Selecciona la versión de OpenSearch motor compatible con el plugin.
15. Seleccione Importar.

Para asociar un complemento de terceros a un dominio

1. Ahora, asocie la licencia y la configuración del complemento al dominio. En el panel de navegación izquierdo, seleccione Dominios.
2. Elija el nombre del dominio para abrir su configuración de clúster.
3. Vaya a la pestaña Complementos.
4. Elija Asociar paquetes y seleccione los paquetes de complementos, licencias y configuración que acaba de importar.
5. Elija Seleccionar.
6. Elija Siguiente. Revise los paquetes que desea asociar y elija Asociar.

CLI

Primero, use el comando create-package para crear un paquete nuevo que contenga la licencia del complemento. La `S3Key` debe apuntar a un archivo .json o .xml en Amazon S3 que incluya el texto de la licencia o los metadatos.

```
aws opensearch create-package \
```

```
  --package-name plugin-license-package \
  --package-type PACKAGE-LICENSE \
  --package-source S3BucketName=my-bucket,S3Key=licenses/my-plugin-license.json
```

Use el comando create-package nuevamente para crear un paquete que contenga la configuración del complemento. La `S3Key` debe apuntar a un archivo .zip en Amazon S3 que cumpla con la estructura de directorios esperada por el complemento.

```
aws opensearch create-package \
  --package-name plugin-config-package \
  --package-type PACKAGE-CONFIG \
  --package-source S3BucketName=my-bucket,S3Key=path/to/package.zip
```

Use nuevamente el comando create-package para crear un paquete que contenga el complemento en sí. La `S3Key` debe apuntar al archivo .zip del complemento en Amazon S3.

```
aws opensearch create-package \
  --package-name plugin-package \
  --package-type ZIP-PLUGIN \
  --package-source S3BucketName=my-bucket,S3Key=path/to/package.zip
```

Por último, utilice el comando associate-package para vincular el complemento, la licencia y la configuración del socio a un dominio compatible especificando el paquete IDs para cada uno. Especifique el ID del complemento como un requisito previo para los demás paquetes; esto significa que debe asociarse al dominio antes que los otros paquetes.

```
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"]}]'
```

Siguientes pasos

Cuando la asociación se complete, usted podrá habilitar el complemento en índices específicos o configurarlo según sus necesidades. Para aplicar la funcionalidad del complemento de terceros a índices específicos, modifique la configuración del índice durante su creación o actualice índices existentes. Por ejemplo, si su complemento de terceros incluye un analizador personalizado, haga referencia a este en la configuración del índice.

Para aplicar las características del complemento de manera uniforme en varios índices, use plantillas de índice que incluyan las configuraciones del complemento. Consulte siempre la documentación del complemento para saber cómo configurar sus funciones para su OpenSearch configuración.

Consulta tus datos de Amazon OpenSearch Service con SQL

Puedes usar SQL para consultar tu Amazon OpenSearch Service, en lugar de usar la OpenSearch consulta DSL basada en JSON. Consultar con SQL es útil si ya está familiarizado con el lenguaje o si desea integrar su dominio en una aplicación que lo utilice. La compatibilidad con SQL está disponible en los dominios que ejecutan Elasticsearch OpenSearch 6.5 o superior.

> **Note**
>
> Esta documentación describe la compatibilidad de versiones entre OpenSearch Service y varias versiones del complemento SQL, así como los controladores JDBC y ODBC. Consulte la OpenSearchdocumentación de código abierto para obtener información sobre la sintaxis de las consultas básicas y complejas, las funciones, las consultas de metadatos y las funciones agregadas.

Usa la siguiente tabla para encontrar la versión del complemento de SQL compatible con cada versión OpenSearch y con Elasticsearch.

OpenSearch

OpenSearch versión	Versión del complemento de SQL	Características notables
2.19.0	2.19.0.0	
2.18.0	2.18.0.0	
2.17.0	2.17.0.0	
2.15.0	2.15.0.0	
2.13.0	2.13.0.0	

OpenSearch versión	Versión del complemento de SQL	Características notables
2.11.0	2.11.0.0	Adición de compatibilidad con el lenguaje y las consultas de PPL
2.9.0	2.9.0.0	Se añadió el conector Spark y admite funciones de tabla y ProMQL
2.7.0	2.7.0.0	Se añadió API `datasource`
2.5.0	2.5.0.0	
2.3.0	2.3.0.0	Se añadieron funciones de fecha y hora `maketime` y `makedate`
1.3.0	1.3.0.0	Admite tamaño límite de consulta predeterminado y cláusula IN al seleccionar de una lista de valores
1.2.0	1.2.0.0	Se añadió nuevo protocolo para el formato de respuesta de visualización
1.1.0	1.1.0.0	Admite la función de coincidencia como filtro en SQL y PPL
1.0.0	1.0.0.0	Admite la consulta de un flujo de datos

Open Distro para Elasticsearch

Versión de Elasticsearch	Versión del complemento de SQL	Características notables
7,10	1.13.0	NULL FIRST y LAST para funciones de ventana, función CAST(), comandos SHOW y DESCRIBE
7.9	1.11.0	Añada date/time funciones adicionales, palabra clave ORDER BY

Versión de Elasticsearch	Versión del complemento de SQL	Características notables
7.8	1.9.0	
7.7	1.8.0	
7.3	1.3.0	Varios operadores de cadena y número
7.1	1.1.0	

Ejemplo de llamada

Para consultar los datos con SQL, envíe solicitudes HTTP a `_sql` con el formato siguiente:

```
POST domain-endpoint/_plugins/_sql
{
  "query": "SELECT * FROM my-index LIMIT 50"
}
```

> **Note**
>
> Si tu dominio ejecuta Elasticsearch en lugar de hacerlo OpenSearch, el formato será. `_opendistro/_sql`

Notas y diferencias

Las llamadas a `_plugins/_sql` incluyen los nombres de índice en el cuerpo de la solicitud y, por lo tanto, se aplican las mismas consideraciones de política de acceso que en las operaciones bulk, mget y msearch. Como siempre, siga el principio de privilegios mínimos cuando conceda permisos a las operaciones de API.

Para consideraciones de seguridad relacionadas con la utilización de SQL con el control de acceso detallado, consulte the section called “Control de acceso detallado”.

El complemento OpenSearch SQL incluye muchos ajustes que se pueden ajustar. En OpenSearch Service, usa la `_cluster/settings` ruta, no la ruta de configuración del complemento (`_plugins/_query/settings`):

```
PUT _cluster/settings
{
  "transient" : {
    "plugins.sql.enabled" : true
  }
}
```

En los dominios de Elasticsearch heredados, reemplace `plugins` con `opendistro`:

```
PUT _cluster/settings
{
  "transient" : {
    "opendistro.sql.enabled" : true
  }
}
```

SQL Workbench

SQL Workbench es una interfaz de usuario de OpenSearch Dashboards que le permite ejecutar consultas de SQL bajo demanda, traducir SQL a su equivalente en REST y ver y guardar los resultados como texto, JSON, JDBC o CSV. Para más información, consulte Query Workbench.

CLI SQL

La CLI SQL es una aplicación Python independiente que se puede iniciar con el comando `opensearchsql`. Para obtener información sobre los pasos para instalar, configurar y utilizar, consulte SQL CLI.

Controlador JDBC

El controlador de conectividad de bases de datos Java (JDBC) le permite integrar los dominios de OpenSearch servicio con sus aplicaciones de inteligencia empresarial (BI) favoritas. Para descargar el controlador, haga clic aquí. Para obtener más información, consulte el GitHub repositorio.

Controlador ODBC

El controlador de Open Database Connectivity (ODBC) es un controlador de ODBC de solo lectura para Windows y macOS que permite conectar aplicaciones de inteligencia empresarial y visualización de datos como Microsoft Excel al complemento SQL.

Para obtener información sobre la instalación del controlador, consulte el repositorio de SQL en GitHub.

Búsqueda entre clústeres en Amazon Service OpenSearch

La búsqueda entre clústeres en Amazon OpenSearch Service te permite realizar consultas y agregaciones en varios dominios conectados. A menudo tiene más sentido usar varios dominios más pequeños en lugar de un solo dominio grande, especialmente cuando se ejecutan distintos tipos de cargas de trabajo.

Los dominios específicos de la carga de trabajo permiten realizar las siguientes tareas:

- Optimice cada dominio mediante la elección de tipos de instancia para cargas de trabajo específicas.
- Establezca límites de aislamiento de errores entre cargas de trabajo. Esto significa que si una de las cargas de trabajo devuelve un error, el error queda contenido dentro de ese dominio específico y no afecta a las demás cargas de trabajo.
- Escale más fácilmente entre dominios.

La búsqueda entre clústeres es compatible con OpenSearch los paneles, por lo que puede crear visualizaciones y paneles en todos sus dominios. Pagas las tarifas estándar de transferencia AWS de datos por los resultados de búsqueda transferidos entre dominios.

> **Note**
>
> El código abierto OpenSearch también tiene documentación para la búsqueda entre clústeres. La configuración difiere considerablemente entre los clústeres de código abierto y los dominios gestionados OpenSearch de Amazon Service. En particular, en OpenSearch Service, las conexiones entre clústeres se configuran mediante la cURL, Consola de administración de AWS en lugar de a través de ella. Además, el servicio gestionado utiliza AWS Identity and Access Management (IAM) para la autenticación entre clústeres,

además de un control de acceso detallado. Por lo tanto, le recomendamos que utilice esta documentación, en lugar de la documentación de código abierto, para configurar la OpenSearch búsqueda entre clústeres para sus dominios.

Temas

- Limitaciones
- Requisitos previos de búsqueda entre clústeres
- Precio de búsqueda entre clústeres
- Configuración de una conexión
- Eliminación de una conexión
- Configuración de seguridad y explicación de ejemplo
- OpenSearch Cuadros de mando

Limitaciones

La búsqueda en clústeres tiene varias limitaciones importantes:

- No puedes conectar un dominio de Elasticsearch a otro dominio. OpenSearch
- No puedes conectarte a clústeres OpenSearch/Elasticsearch autogestionados.
- Para conectar dominios entre regiones, ambos dominios deben estar en Elasticsearch 7.10 o una versión posterior o. OpenSearch
- Un dominio puede tener un máximo de 20 conexiones salientes. Del mismo modo, un dominio puede tener un máximo de 20 conexiones entrantes. En otras palabras, un dominio puede conectarse a un máximo de 20 dominios.
- El dominio de origen debe estar en la misma versión o superior a la del dominio de destino. Si configura una conexión bidireccional entre dos dominios y desea actualizar uno o ambos, antes debe eliminar una de las conexiones.
- No puede usar diccionarios personalizados o SQL con la búsqueda en clústeres.
- No se puede usar CloudFormation para conectar dominios.
- No se puede utilizar la búsqueda entre clústeres en instancias M3 o bursátiles (T2 y T3).

Requisitos previos de búsqueda entre clústeres

Antes de configurar la búsqueda en clústeres, asegúrese de que los dominios cumplan los siguientes requisitos:

- Dos OpenSearch dominios o dominios de Elasticsearch en la versión 6.7 o posterior
- Control de acceso detallado habilitado
- Node-to-node cifrado activado
- Si alguno de los dominios está en una VPC, los dominios deben conectarse mediante VPC Peering o Transit Gateway, y los grupos de seguridad deben permitir el tráfico entre ellos.

Precio de búsqueda entre clústeres

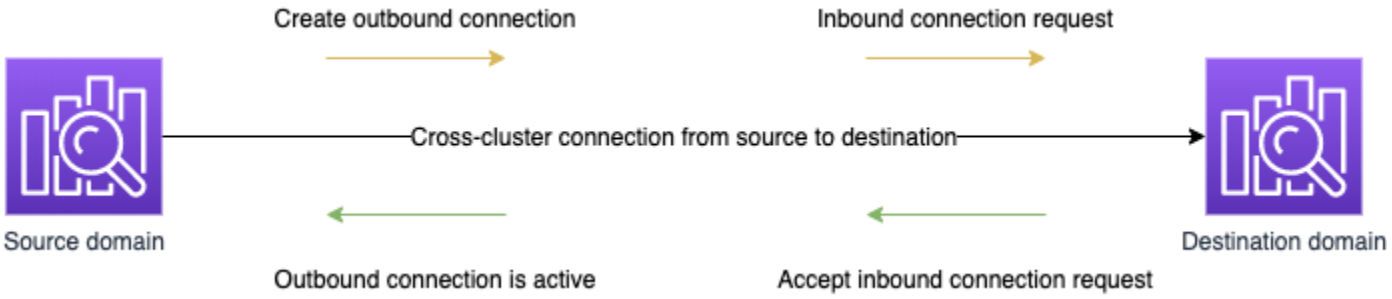
No hay ningún cargo adicional por buscar entre dominios.

Configuración de una conexión

El dominio de “origen” hace referencia al dominio desde el que se origina una petición de búsqueda en clústeres. En otras palabras, el dominio de origen es al que envía la petición de búsqueda inicial.

El dominio “destino” es el dominio que consulta el dominio de origen.

Una conexión entre clústeres es unidireccional desde el dominio de origen hasta el dominio de destino. Esto significa que el dominio de destino no puede consultar el dominio de origen. Sin embargo, puede configurar otra conexión en la dirección opuesta.



El dominio de origen crea una conexión “saliente” al dominio de destino. El dominio de destino recibe una solicitud de conexión “entrante” del dominio de origen.

Para configurar una conexión

1. En el panel del dominio, seleccione un dominio y vaya a la pestaña Conexiones.

2. En la sección Conexiones de salida, seleccione Solicitar.
3. En Alias de conexión, ingrese un nombre para la conexión.
4. Elige entre conectarte a un dominio de tu Cuenta de AWS región o a otra cuenta o región.

   - Para conectarte a un clúster de tu Cuenta de AWS región, selecciona el dominio en el menú desplegable y selecciona Solicitar.
   - Para conectarse a un clúster de otra región Cuenta de AWS o región, seleccione el ARN del dominio remoto y elija Solicitar. Para conectar dominios entre regiones, ambos dominios deben ejecutar Elasticsearch versión 7.10 o posterior o. OpenSearch
5. Para omitir los clústeres no disponibles para las consultas de clústeres, seleccione Omitir los no disponibles. Esta configuración garantiza que las consultas entre clústeres devuelvan resultados parciales a pesar de que se produzcan errores en uno o más clústeres remotos.
6. La búsqueda entre clústeres valida primero la solicitud de conexión para asegurarse de que se cumplen los requisitos previos. Si los dominios son incompatibles, la solicitud de conexión entra en el estado `Validation failed`.
7. Una vez correctamente validada la solicitud de conexión, se envía al dominio de destino, donde tiene que aprobarse. Mientras no se obtenga la aprobación, la conexión permanecerá en el estado de `Pending acceptance`. Cuando se acepta la solicitud de conexión en el dominio de destino, el estado cambia a `Active` y el dominio de destino se vuelve disponible para consultas.

   - La página de dominio muestra el estado general del dominio y los detalles del estado de la instancia del dominio de destino. Los propietarios de dominios tienen la flexibilidad de crear, visualizar, eliminar y monitorear conexiones desde o hacia sus dominios.

Una vez establecida la conexión, se cifra todo el tráfico que circule entre los nodos de los dominios conectados. Si conecta un dominio de VPC a un dominio que no es VPC y el dominio que no es VPC es un punto de conexión público que puede recibir tráfico de Internet, el tráfico entre clústeres entre los dominios sigue cifrado y seguro.

Eliminación de una conexión

La eliminación de una conexión detiene cualquier operación entre clústeres en sus índices.

1. En el panel del dominio, vaya a la pestaña Conexiones.
2. Seleccione las conexiones de dominio que desea eliminar, luego elija Eliminar y, a continuación, confirme la eliminación.

Puede realizar estos pasos en el dominio fuente o de destino para eliminar la conexión. Una vez quitada la conexión, seguirá visible con el estado `Deleted` durante un periodo de 15 días.

No se puede eliminar un dominio con conexiones activas entre clústeres. Para eliminar un dominio, primero quite todas las conexiones entrantes y salientes de ese dominio. Esto es para asegurarse de tener en cuenta a los usuarios del dominio entre clústeres antes de eliminar el dominio.

Configuración de seguridad y explicación de ejemplo

1. Envíe una petición de búsqueda en clústeres al dominio de origen.
2. El dominio de origen evalúa esa solicitud en comparación con la política de acceso al dominio. Dado que la búsqueda en clústeres requiere un control de acceso detallado, recomendamos una política de acceso abierto en el dominio de origen.

   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**
>
> Si incluye índices remotos en la ruta, debe codificar la URL del URI en el ARN del dominio. Por ejemplo, use `arn:aws:es:us-east-1:123456789012:domain/`

`my-domain/local_index,dst%3Aremote_index` en lugar de `arn:aws:es:us-east-1:123456789012:domain/my-domain/local_index,dst:remote_index`.

Si elige utilizar una política de acceso restrictivo además de un control de acceso detallado, la política debe permitir el acceso a `es:ESHttpGet` como mínimo.

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. El control de acceso detallado en el dominio de origen evalúa la solicitud:

   - ¿La solicitud está firmada con credenciales básicas IAM o HTTP válidas?
   - Si es así, ¿tiene el usuario permiso para realizar la búsqueda y acceder a los datos?

   Si la solicitud solo busca datos en el dominio de destino (por ejemplo, `dest-alias:dest-index/_search`), solo necesitará permisos en el dominio de destino.

   Si la solicitud busca datos en ambos dominios (por ejemplo, `source-index,dest-alias:dest-index/_search`), necesita permisos en ambos dominios.

   En el control de acceso detallado, los usuarios deben tener el permiso `indices:admin/shards/search_shards` además de los permisos `read` o `search` estándar para los índices pertinentes.

4. El dominio de origen pasa la solicitud al dominio de destino. El dominio de destino evalúa esta solicitud frente a su política de acceso al dominio. Debe incluir el permiso `es:ESCrossClusterGet` en el dominio de destino:

   JSON

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

   Asegúrese de que el permiso `es:ESCrossClusterGet` se aplica para `/dst-domain` y no `/dst-domain/*`.

   Sin embargo, esta política mínima solo permite búsquedas en clústeres. Para realizar otras operaciones, como indexar documentos y realizar búsquedas estándar, necesita permisos adicionales. Recomendamos la siguiente política en el dominio de destino:

   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**
>
> Todas las solicitudes de búsqueda entre clústeres entre dominios se cifran en tránsito de forma predeterminada como parte del cifrado. node-to-node

5. El dominio de destino realiza la búsqueda y devuelve los resultados al dominio de origen.
6. El dominio de origen combina sus propios resultados (si los hay) con los resultados del dominio de destino y los devuelve.
7. Recomendamos Postman para probar las solicitudes:

   - En el dominio de destino, indexe un documento:

   ```
   POST https://dst-domain.us-east-1.es.amazonaws.com/books/_doc/1

   {
     "Dracula": "Bram Stoker"
   }
   ```

   - Para consultar este índice desde el dominio de origen, incluya el alias de conexión del dominio de destino dentro de la consulta.

   ```
   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"
      }
    }
  ]
}
```

Puede encontrar el alias de conexión en la pestaña Conexiones del panel de control del dominio.

- Si configura una conexión entre `domain-a -> domain-b` con alias de conexión `cluster_b` y `domain-a -> domain-c` con alias de conexión `cluster_c`, busque `domain-a`, `domain-b` y `domain-c` de la siguiente manera:

```
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"
    }
  }
}
```

Respuesta

```
{
  "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
        }
      }
    ]
```

```
  }
}
```

Si no eligió omitir los clústeres no disponibles en la configuración de la conexión, todos los clústeres de destino que busca tienen que estar disponibles para que la petición de búsqueda se ejecute correctamente. De lo contrario, se produce un error en toda la solicitud; incluso si uno de los dominios no está disponible, la búsqueda no devuelve ningún resultado.

OpenSearch Cuadros de mando

Puede visualizar datos de varios dominios conectados de la misma manera que desde un solo dominio, excepto que debe acceder a los índices remotos mediante `connection-alias:index`. Por lo tanto, su patrón de índice debe coincidir `connection-alias:index`.

Aprender a clasificar para Amazon OpenSearch Service

OpenSearch utiliza un marco de clasificación probabilístico llamado BM-25 para calcular las puntuaciones de relevancia. Si una palabra clave distintiva aparece con más frecuencia en un documento, BM-25 asigna una puntuación de relevancia más alta a ese documento. Este marco, sin embargo, no tiene en cuenta el comportamiento del usuario, como los datos de clics, lo que puede mejorar aún más la relevancia.

Learning to Rank es un complemento de código abierto que permite utilizar machine learning y los datos de comportamiento para ajustar la relevancia de los documentos. Utiliza modelos de las bibliotecas Ranklib XGBoost y Ranklib para volver a puntuar los resultados de la búsqueda. El plugin Elasticsearch LTR fue desarrollado inicialmente por OpenSource Connections, con importantes contribuciones de la Fundación Wikimedia, Snagajob Engineering, Bonsai y Yelp Engineering. La OpenSearch versión del complemento se deriva del complemento LTR de Elasticsearch.

Para aprender a clasificar se necesita Elasticsearch OpenSearch 7.7 o una versión posterior. Para utilizar el complemento Aprender a clasificar, debe tener permisos de administrador completos. Para obtener más información, consulte the section called “Modificar el usuario maestro”.

> **Note**
>
> Esta documentación proporciona información general del complemento Aprender a clasificar y le ayuda a empezar a usarlo. La documentación completa, incluidos los pasos detallados

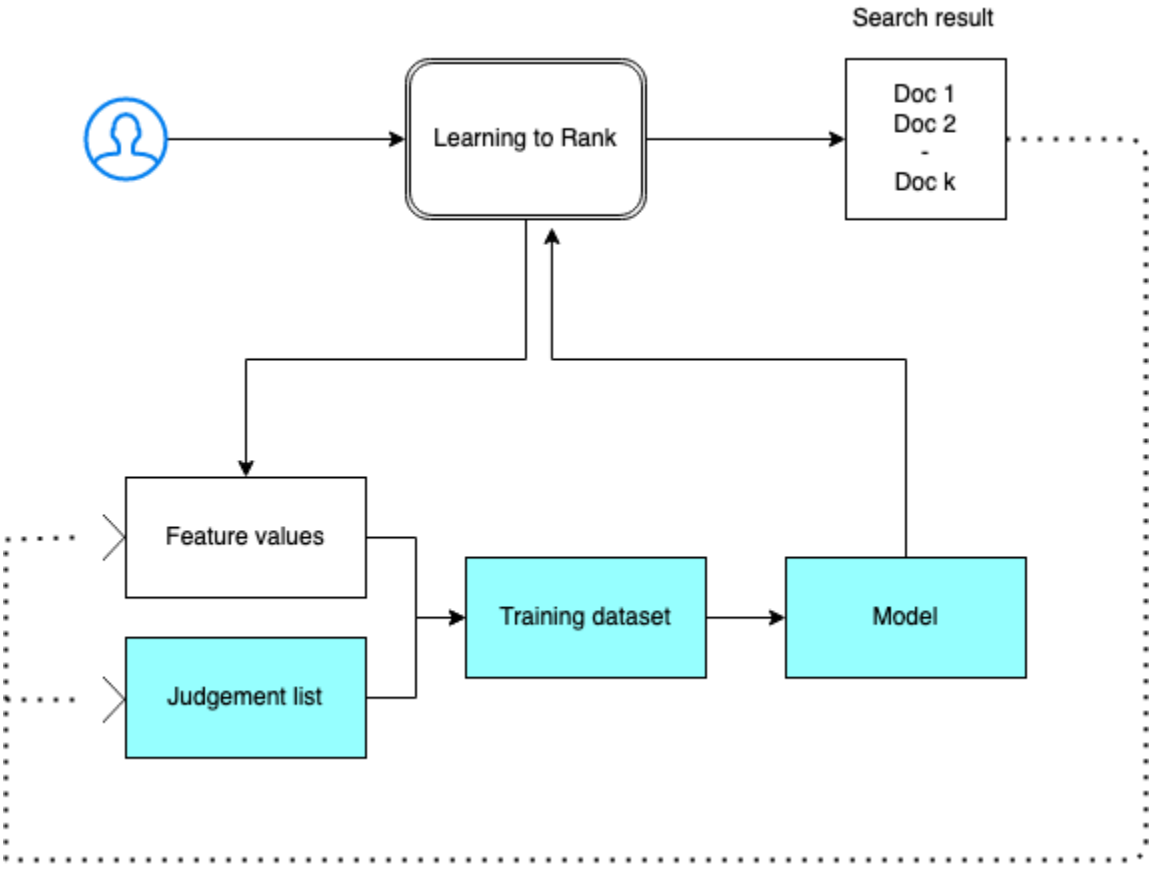
y las descripciones de la API, está disponible en la documentación de Learning to Rank (Aprender a clasificar).

Temas

- Introducción a Aprender a clasificar
- API de Aprender a clasificar

Introducción a Aprender a clasificar

Debes proporcionar una lista de opiniones, preparar un conjunto de datos de formación y entrenar el modelo fuera de Amazon OpenSearch Service. Las partes en azul se encuentran fuera del OpenSearch Servicio:



Paso 1: inicializar el complemento

Para inicializar el complemento Learning to Rank, envía la siguiente solicitud a tu dominio OpenSearch de servicio:

```
PUT _ltr
```

```
{
  "acknowledged" : true,
  "shards_acknowledged" : true,
  "index" : ".ltrstore"
}
```

Este comando crea un índice `.ltrstore` oculto que almacena información de metadatos, como conjuntos de características y modelos.

Paso 2: crear una lista de criterios

> **Note**
>
> Debe realizar este paso fuera del OpenSearch Servicio.

Una lista de criterios es un conjunto de ejemplos de los cuales aprende un modelo de machine learning. La lista de criterios debe incluir palabras clave importantes para usted y un conjunto de documentos calificados para cada palabra clave.

En este ejemplo, tenemos una lista de criterios para un conjunto de datos de películas. Un grado de 4 indica una coincidencia perfecta. Un grado de 0 indica la peor coincidencia.

Grado	Palabra clave	ID del documento	Nombre de la película
4	rambo	7555	Rambo
3	rambo	1370	Rambo III
3	rambo	1369	Rambo: First Blood Part II
3	rambo	1368	First Blood

Prepare la lista de criterios con el siguiente formato:

```
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"
```

Para ver un ejemplo más completo de una lista de criterios, consulte Criterios sobre películas.

Puede crear esta lista de criterios de forma manual con la ayuda de anotadores humanos o inferirla de manera programática a partir de datos de análisis.

Paso 3: crear un conjunto de características

Una característica es un campo que corresponde a la relevancia de un documento, por ejemplo, `title`, `overview`, `popularity score` (número de vistas) y así sucesivamente.

Cree un conjunto de características con una plantilla de Mustache para cada característica. Para obtener más información acerca de las características, consulte Trabajo con características.

En este ejemplo, crearemos un conjunto de características `movie_features` con los campos `title` y `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}}"
            }
          }
        }
      ]
    }
}
```

Si consulta el índice `.ltrstore` original, recupera el conjunto de características:

```
GET _ltr/_featureset
```

Paso 4: registrar los valores de la característica

Los valores de la característica son las puntuaciones de relevancia calculadas por BM-25 para cada característica.

Combine el conjunto de características y la lista de sentencias para registrar los valores de la característica. Para obtener más información acerca de las características de registro, consulte Puntuaciones de características.

En este ejemplo, la consulta `bool` recupera los documentos calificados con el filtro y, a continuación, elige el conjunto de características con la consulta `sltr`. La consulta `ltr_log` combina los documentos y las características para registrar los valores de característica correspondientes:

```
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"
      }
    }
  }
}
```

Un ejemplo de respuesta se vería así:

```
{
  "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"
      ]
```

```
      }
    ]
  }
}
```

En el ejemplo anterior, la primera característica no posee un valor de característica, ya que la palabra clave “rambo” no aparece en el campo de título del documento con un ID igual a 1368. Se trata de un valor de característica que falta en los datos de entrenamiento.

Paso 5: crear un conjunto de datos de entrenamiento

> **Note**
>
> Debe realizar este paso fuera del OpenSearch Servicio.

El siguiente paso es combinar la lista de sentencias y los valores de característica para crear un conjunto de datos de entrenamiento. Si la lista de criterios original se ve así:

```
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
```

Conviértala en el conjunto de datos de entrenamiento final, que se ve así:

```
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
```

Puede realizar este paso de forma manual o escribir un programa para automatizarlo.

Paso 6: elegir un algoritmo y crear el modelo

> **Note**
>
> Debe realizar este paso fuera del OpenSearch Servicio.

Una vez creado el conjunto de datos de entrenamiento, el siguiente paso es utilizar nuestras XGBoost bibliotecas de Ranklib para crear un modelo. XGBoost y las bibliotecas de Ranklib te permiten crear modelos populares como LambdaMart, Random Forests, etc.

Para ver los pasos a seguir XGBoost y usar Ranklib para construir el modelo, consulte la documentación y, respectivamente. XGBoostRankLib Para usar Amazon SageMaker para crear el XGBoost modelo, consulta XGBoostAlgorithm.

Paso 7: implementar el modelo

Después de haber creado el modelo, impleméntelo en el complemento Learning to Rank (Aprender a clasificar). Para obtener más información acerca de la implementación de un modelo, consulte Carga de un modelo entrenado.

En este ejemplo, crearemos un modelo de `my_ranklib_model` con la biblioteca 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>
"""
    }
  }
}
```

Para ver el modelo, envíe la siguiente solicitud:

```
GET _ltr/_model/my_ranklib_model
```

Paso 8: buscar con Aprender a clasificar

Después de implementar el modelo, estará listo para realizar la búsqueda.

Realice la consulta `sltr` con las características que utiliza y el nombre del modelo que desea ejecutar:

```
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 Aprender a clasificar, usted ve “Rambo” como el primer resultado porque le asignamos la calificación más alta en la lista de criterios:

```
{
  "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"
        }
      }
    ]
  }
}
```

Si buscas sin usar el plugin Learning to Rank, OpenSearch obtendrás resultados diferentes:

```
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"
        }
      }
    ]
  }
}
```

En función de cuán bien considere que el modelo funciona, ajuste la lista de criterios y las características. A continuación, repita los pasos 2 a 8 para mejorar los resultados de la clasificación a lo largo del tiempo.

API de Aprender a clasificar

Utilice las operaciones de Aprender a clasificar para trabajar de manera programática con conjuntos de características y modelos.

Crear tienda

Crea un índice `.ltrstore` oculto que almacena información sobre metadatos, como conjuntos de características y modelos.

```
PUT _ltr
```

Eliminar tienda

Elimina el índice `.ltrstore` oculto y restablece el complemento.

```
DELETE _ltr
```

Crear conjunto de características

Crea un conjunto de características.

```
POST _ltr/_featureset/<name_of_features>
```

Eliminar un conjunto de características

Elimina un conjunto de características.

```
DELETE _ltr/_featureset/<name_of_feature_set>
```

Obtener un conjunto de características

Recupera un conjunto de características.

```
GET _ltr/_featureset/<name_of_feature_set>
```

Crear un modelo

Crea un modelo.

```
POST _ltr/_featureset/<name_of_feature_set>/_createmodel
```

Eliminar modelo

Elimina un modelo.

```
DELETE _ltr/_model/<name_of_model>
```

Obtener modelo

Recupera un modelo.

```
GET _ltr/_model/<name_of_model>
```

Obtener estadísticas

Proporciona información acerca de cómo se comporta el complemento.

```
GET _ltr/_stats
```

También puede usar filtros para recuperar una sola estadística:

```
GET _ltr/_stats/<stat>
```

Además, puede limitar la información a un único nodo en el clúster:

```
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
    }
  }
}
```

Las estadísticas se proporcionan en dos niveles, nodo y clúster, como se especifica en las siguientes tablas:

Estadísticas a nivel de nodo

Nombre del campo	Description (Descripción)
request_total_count	Recuento total de solicitudes de clasificación.

Nombre del campo	Description (Descripción)
request_error_count	Recuento total de solicitudes fallidas.
cache	Estadísticas de todas las cachés (caracter ísticas, conjuntos de características, modelos). Se produce un acierto de caché cuando un usuario consulta el complemento y el modelo ya está cargado en la memoria.
cache.eviction_count	Número de expulsiones de caché.
cache.hit_count	Número de aciertos de caché.
cache.miss_count	Número de errores de caché. Se produce un error de caché cuando un usuario consulta el complemento y el modelo aún no se cargó en la memoria.
cache.entry_count	Número de entradas en la caché.
cache.memory_usage_in_bytes	Memoria total utilizada en bytes.
cache.cache_capacity_reached	Indica si se alcanza el límite de caché.

Estadísticas de nivel de clúster

Nombre del campo	Description (Descripción)
tiendas	Indica dónde se almacenan los conjuntos de características y los metadatos del modelo. (El valor predeterminado es “.ltrstore”. De lo contrario, tiene el prefijo “.ltrstore_”, con un nombre proporcionado por el usuario).
stores.status	Estado del índice.
stores.feature_sets	Número de conjuntos de características.

Nombre del campo	Description (Descripción)
stores.features_count	Número de características.
stores.model_count	Número de modelos.
status	Estado del complemento basado en el estado de los índices del almacén de características (rojo, amarillo o verde) y el estado del interrupt or de circuito (abierto o cerrado).
cache.cache_capacity_reached	Indica si se alcanza el límite de caché.

Obtener estadísticas de caché

Devuelve estadísticas sobre el uso de la memoria y la caché.

```
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": {
                ...
            }
        }
    }
}
```

Borrar caché

Borra la caché del complemento. Utilice esta opción para actualizar el modelo.

```
POST _ltr/_clearcache
```

Búsqueda asíncrona en Amazon Service OpenSearch

Con la búsqueda asíncrona de Amazon OpenSearch Service, puedes enviar una consulta de búsqueda que se ejecute en segundo plano, supervisar el progreso de la solicitud y recuperar los resultados más adelante. Puede recuperar resultados parciales, a medida que estén disponibles, antes de que se complete la búsqueda. Una vez finalizada la búsqueda, guarde los resultados para una posterior recuperación y análisis.

La búsqueda asíncrona requiere la OpenSearch versión 1.0 o una versión posterior, o Elasticsearch 7.10 o una versión posterior.

Esta documentación proporciona una breve información general de la búsqueda asíncrona. También se analizan las limitaciones del uso de la búsqueda asíncrona con un dominio de Amazon OpenSearch Service gestionado en lugar de con un clúster de código abierto. OpenSearch Para obtener la documentación completa sobre la búsqueda asincrónica, incluida la configuración disponible, los permisos y una referencia completa de la API, consulta la búsqueda asincrónica en la documentación. OpenSearch

Ejemplo de llamada de búsqueda

Para realizar una búsqueda asíncrona, envíe solicitudes HTTP a `_plugins/_asynchronous_search` con el siguiente formato:

```
POST opensearch-domain/_plugins/_asynchronous_search
```

> **Note**
>
> Si utilizas Elasticsearch 7.10 en lugar de una OpenSearch versión, sustitúyela por en todas las solicitudes de búsqueda asincrónica. `_plugins` `_opendistro`

Puede especificar las siguientes opciones de búsqueda asíncrona:

Opciones	Description (Descripción)	Predeterminado	Obligatorio
`wait_for_completion_timeout`	Especifica la cantidad de tiempo durante la cual planea esperar los resultados. Puede ver los resultados que obtiene dentro de este tiempo, igual que en una búsqueda normal. Puede sondear los resultados restantes en función de un ID. El valor máximo es 300 segundos.	1 segundos	No
`keep_on_completion`	Especifica si desea guardar los resultados en el clúster una vez finalizada la búsqueda. Puede examinar los resultados almacenados más adelante.	`false`	No
`keep_alive`	Especifica la cantidad de tiempo que se guarda el resultado en el clúster. Por ejemplo, `2d` significa que los resultados se almacenan en el clúster durante 48 horas. Los resultados de búsqueda guardados se eliminan después de este periodo o si se cancela la búsqueda. Tenga en cuenta que esto incluye el tiempo de ejecución de la consulta. Si la consulta sobrepasa este tiempo, el proceso cancela esta consulta automáticamente.	12 horas	No

Solicitud de ejemplo

```
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**
>
> Todos los parámetros de solicitud que se aplican a una consulta `_search` estándar son compatibles. Si utilizas Elasticsearch 7.10 en lugar de una versión, sustitúyela por. OpenSearch `_plugins` `_opendistro`

Permisos de búsqueda asíncrona

La búsqueda asíncrona admite el control de acceso detallado. Para obtener más información sobre la combinación y la coincidencia de permisos para adaptarlos al caso de uso, consulte Asynchronous search security.

Para los dominios con control de acceso detallado habilitado, necesita los siguientes permisos mínimos para un rol:

```
# 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'
```

Para dominios con control de acceso detallado deshabilitado, utilice el acceso de IAM y la clave secreta para firmar todas las solicitudes. Puede acceder a los resultados con el ID de búsqueda asíncrona.

Configuración de búsqueda asincrónica

OpenSearch te permite cambiar todos los ajustes de búsqueda asíncrona disponibles mediante la API. `_cluster/settings` En OpenSearch Service, solo puedes cambiar los siguientes ajustes:

- `plugins.asynchronous_search.node_concurrent_running_searches`
- `plugins.asynchronous_search.persist_search_failures`

Búsqueda en clústeres

Puede realizar una búsqueda asíncrona entre clústeres con las siguientes limitaciones menores:

- Puede ejecutar una búsqueda asíncrona solo en el dominio fuente.
- No puede minimizar los viajes de ida y vuelta de la red como parte de una consulta de búsqueda entre clústeres.

Si configura una conexión entre `domain-a -> domain-b`, con alias de conexión `cluster_b`, y `domain-a -> domain-c`, con alias de conexión `cluster_c`, realice una búsqueda asíncrona de `domain-a`, `domain-b` y `domain-c`, de la siguiente manera:

```
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": []
    }
  }
```

```
}
```

Respuesta

```
{
  "id" :
 "Fm9pYzJyVG91U19xb0hIQUJnMHJfRFEAAAAAAAknghQ1OWVBczNZQjVEa2dMYTBXaTdEagAAAAAAAAAB",
  "state" : "RUNNING",
  "start_time_in_millis" : 1609329314796,
  "expiration_time_in_millis" : 1609761314796
}
```

Para obtener más información, consulte the section called “Búsqueda en clústeres”.

UltraWarm

Las búsquedas asincrónicas con UltraWarm índices siguen funcionando. Para obtener más información, consulte the section called “UltraWarm almacenamiento”.

> **Note**
>
> Puede supervisar las estadísticas de búsqueda asíncrona en. CloudWatch Para obtener una lista completa de las métricas, consulte the section called “Métricas de búsqueda asíncrona”.

Búsqueda puntual en Amazon OpenSearch Service

El punto en el tiempo (PIT) es un tipo de búsqueda que permite ejecutar diferentes consultas en un conjunto de datos fijo en el tiempo. Normalmente, cuando ejecuta la misma consulta en el mismo índice en distintos puntos en el tiempo, recibe resultados diferentes porque los documentos se indexan, actualizan y eliminan constantemente. Con PIT, puede realizar consultas en función de un estado constante de su conjunto de datos.

El uso principal de la búsqueda es combinarla con la funcionalidad `search_after`. Este es el método de paginación preferido OpenSearch, especialmente para la paginación profunda, porque funciona en un conjunto de datos que está congelado en el tiempo, no está vinculado a una consulta y admite una paginación coherente hacia adelante y hacia atrás. Puede usar PIT con un dominio que ejecute OpenSearch la versión 2.5.

> **Note**
>
> En este tema se proporciona información general sobre PIT y algunos aspectos a tener en cuenta a la hora de utilizarlo en un dominio gestionado de Amazon OpenSearch Service en lugar de en un OpenSearch clúster autogestionado. Para obtener la documentación completa de PIT, incluida una referencia exhaustiva sobre la API, consulte Point in Time en la OpenSearch documentación de código abierto.

Consideraciones

Al configurar las búsquedas PIT, tenga en cuenta lo siguiente:

- Si vas a realizar una actualización desde un dominio que ejecuta la OpenSearch versión 2.3 y necesitas un control de acceso detallado para las acciones de PIT, tendrás que añadir esas acciones y funciones manualmente.
- El PIT no tiene capacidad de recuperación. El reinicio de los nodos, la terminación de los nodos, las implementaciones en azul o verde y los reinicios de los OpenSearch procesos provocan la pérdida de todos los datos PIT.
- Si un fragmento se reubica durante el blue/green despliegue, solo los segmentos de datos activos se transfieren al nuevo nodo. Los segmentos de particiones retenidos por PIT (tanto de forma exclusiva como el que se comparte con los datos activos) permanecen en el nodo anterior.
- Actualmente, las búsquedas PIT no funcionan con la búsqueda asíncrona.

Creación de un PIT

Para ejecutar una consulta de PIT, envíe solicitudes HTTP a `_search/point_in_time` con el siguiente formato:

```
POST opensearch-domain/my-index/_search/point_in_time?keep_alive=time
```

Puede especificar las siguientes opciones de PIT:

Opciones	Description (Descripción)	Predeterminado	Obligatorio
`keep_alive`	La cantidad de tiempo que debe durar la conservación del PIT. Cada vez que se accede a un PIT con una solicitud de búsqueda, la vida útil del PIT se prolonga en un tiempo igual al parámetro `keep_alive` . Este parámetro de consulta es obligatorio cuando se crea un PIT, pero es opcional en una solicitud de búsqueda.		Sí
`preference`	Una cadena que especifica el nodo o la partición utilizados para realizar la búsqueda.	Random	No
`routing`	Una cadena que especifica que las solicitudes de búsqueda se dirijan a una partición específica.	El `_id` del documento	No
`expand_wildcards`	Una cadena que especifica el tipo de índice que puede coincidir con el patrón comodín. Admite valores separados por comas. Los valores válidos son los siguientes: • `all`: Coincide con cualquier índice o flujo de datos, incluidos los ocultos. • `open`: Coincide con índices abiertos, no ocultos o flujos de datos no ocultos. • `closed`: Coincide con índices cerrados, no ocultos o flujos de datos no ocultos. • `hidden`: Coincide con índices o flujos de datos ocultos. Debe combinarse con índices abiertos, cerrados o tanto abiertos como cerrados. • `none`: No se aceptan patrones comodín.	`open`	No

Opciones	Description (Descripción)	Predeterminado	Obligatorio
`allow_partial_pit_creation`	Un valor booleano que especifica si se debe crear un PIT con errores parciales.	`true`	No

Respuesta de ejemplo

```
{
    "pit_id":
 "o463QQEPbXktaW5kZXgtMDAwMDAxFnNOWU43ckt3U3IyaFVpbGE1UWEtMncAFjFyeXBsRGJmVFM2RTB6eVg1aVVqQncAA
    "_shards": {
        "total": 1,
        "successful": 1,
        "skipped": 0,
        "failed": 0
    },
    "creation_time": 1658146050064
}
```

Al crear un PIT, recibirá un ID de PIT en la respuesta. Este es el ID que se utiliza para realizar búsquedas con el PIT.

Permisos de puntos en el tiempo

Los PIT admiten el control de acceso detallado. Si vas a actualizar a un dominio de la OpenSearch versión 2.5 y necesitas un control de acceso detallado, tendrás que crear roles manualmente con los siguientes permisos:

```
# 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"
```

Para los dominios con OpenSearch la versión 2.5 o superior, puedes usar el rol integrado`point_in_time_full_access`. Para obtener más información, consulte el modelo de seguridad en la OpenSearch documentación.

Configuración de PIT

OpenSearch permite cambiar todos los ajustes PIT disponibles mediante la `_cluster/settings` API. En OpenSearch Service, actualmente no puedes modificar la configuración.

Búsqueda en clústeres

Puede crear PITs, buscar con PIT IDs PITs, enumerar y eliminar PITs en todos los clústeres con las siguientes limitaciones menores:

- Puede enumerarlos todos y eliminarlos todos PITs solo en el dominio de origen.
- No puede minimizar los viajes de ida y vuelta de la red como parte de una consulta de búsqueda entre clústeres.

Para obtener más información, consulte the section called “Búsqueda en clústeres”.

UltraWarm

Las búsquedas PIT con UltraWarm índices siguen funcionando. Para obtener más información, consulte the section called “UltraWarm almacenamiento”.

> **Note**
>
> Puede supervisar las estadísticas de búsqueda PIT en CloudWatch. Para obtener una lista completa de las métricas, consulte the section called “Métricas de un momento dado”.

Búsqueda de agentes en Amazon Service OpenSearch

A partir de OpenSearch la versión 3.3, la búsqueda de agentes habilita un proceso basado en la IA que utiliza un agente autónomo para ejecutar búsquedas complejas en su nombre.

La búsqueda de agentes presenta un sistema de agentes inteligente que comprende las intenciones del usuario, organiza las herramientas adecuadas y genera consultas optimizadas. Proporciona resúmenes transparentes de sus decisiones a través de una interfaz de lenguaje natural.

Con OpenSearch Service, puede configurar conectores de IA para servicios externos Servicios de AWS y otros. Con la consola, también puede crear un modelo de aprendizaje automático con una CloudFormation plantilla que se puede utilizar para crear su agente. Para obtener más información, consulte Configuración de la búsqueda de agentes con Bedrock Claude.

Para obtener la documentación y la step-by-step implementación completas, consulte Agentic Search en la documentación. OpenSearch

Búsqueda semántica en Amazon Service OpenSearch

A partir de la versión 2.19, puede utilizar el enriquecimiento semántico automático para implementar la búsqueda semántica con un mínimo esfuerzo de configuración. Esta función genera automáticamente una codificación dispersa para los campos de texto, lo que elimina la necesidad de configurar el proceso de ingesta manual. Para obtener más información, consulte el enriquecimiento semántico automático en la documentación. OpenSearch

A partir de OpenSearch la versión 2.9, puede utilizar la búsqueda semántica para comprender las consultas de búsqueda y mejorar su relevancia. Puede usar la búsqueda semántica de dos maneras: con la búsqueda neuronal y con la búsqueda de k-Nearest Neighbor (k-NN).

Con OpenSearch Service, puede configurar conectores de IA Servicios de AWS y servicios externos. Con la consola, también puede crear un modelo de aprendizaje automático con una CloudFormation plantilla. Para obtener más información, consulte the section called “CloudFormation integraciones de plantillas”.

Para obtener una documentación completa sobre la búsqueda semántica, incluida una step-by-step guía para utilizar la búsqueda semántica, consulte la búsqueda semántica en la documentación de código abierto. OpenSearch

Búsqueda simultánea de segmentos en Amazon Service OpenSearch

A partir de OpenSearch la versión 2.17, la búsqueda de segmentos simultáneos utiliza una nueva configuración para controlar el comportamiento de las búsquedas simultáneas.

- Los dominios nuevos creados con la versión 2.17 tienen la búsqueda simultánea de segmentos configurada en modo automático de manera predeterminada en los nodos de tamaño 2xl o superior.
- Los dominios existentes que se actualizan a 2.17 tienen la búsqueda simultánea de segmentos configurada en modo automático según el tipo de instancia para todos los nodos de tamaño 2xl o superior y si la utilización general de CPU del clúster fue inferior al 45 % durante la última semana.
- Para obtener más información, consulte Búsqueda de segmentos concurrentes versión 2.17.

A partir de la OpenSearch versión 2.13, puede utilizar la búsqueda de segmentos simultánea para ayudarle a buscar segmentos en paralelo durante la fase de consulta. Para obtener una documentación completa sobre la búsqueda simultánea de segmentos, consulte la búsqueda simultánea de segmentos en la documentación de código abierto. OpenSearch Para obtener información sobre CloudWatch las métricas de Amazon relacionadas con la búsqueda simultánea de segmentos, consulta Métricas y UltraWarm métricas de instancias.

Existen algunas limitaciones adicionales que se aplican cuando utilizas la búsqueda de segmentos actual con Amazon OpenSearch Service:

- De forma predeterminada, OpenSearch Service utiliza un recuento de 2 segmentos con el mecanismo de recuento máximo de segmentos.

Generación de consultas en lenguaje natural en Amazon OpenSearch Service

La función de generación de consultas en lenguaje natural de Amazon OpenSearch Service te permite consultar los datos del registro operativo y de seguridad mediante un lenguaje natural. OpenSearch es una opción ideal para explorar los datos de registro, ya que es un motor de búsqueda y análisis de registros altamente escalable y eficiente, y ahora puede utilizar el lenguaje natural para explorar estos registros. Esta función le permite identificar problemas sin tener que recurrir al lenguaje de procesamiento OpenSearch canalizado (PPL) ni tener que buscar las definiciones de los datos al crear las consultas. Puede utilizar la función de generación de consultas en lenguaje natural en los dominios de OpenSearch servicio con la versión 2.13 y posteriores. Debe tener el control de acceso detallado habilitado.

Esta función se creó con el kit de herramientas OpenSearch Assistant. Si desea crear características similares que se conecten a sus modelos lingüísticos de gran tamaño, puede utilizar el kit de herramientas para configurar sus propios agentes y herramientas.

Requisitos previos

Antes de poder utilizar la característica de generación de consultas en lenguaje natural, su dominio debe tener lo siguiente:

- Versión 2.13 o posterior.
- Software del servicio R20240520-P4 o superior.
- Control de acceso detallado habilitado. Para obtener más información, consulte the section called “Habilitar el control de acceso detallado”.

Introducción

La generación de consultas en lenguaje natural está habilitada de forma predeterminada en todos los dominios creados con la versión 2.13 o posterior que tengan el control de acceso detallado habilitado.

Para otros dominios, habilítelo seleccionando Habilitar generación de consulta en lenguaje natural y características de Amazon Q Developer.

Después de habilitarla, vaya a la página de registros en los OpenSearch paneles de control. Seleccione Explorador de eventos y haga una pregunta con el asistente para consultas.

Configuración de permisos

Si habilita la generación de consultas en lenguaje natural en un dominio de OpenSearch servicio preexistente, es posible que la función query_assistant_access no esté definida en el dominio. Los usuarios que no sean administradores deben estar asignados a este rol para poder administrar índices templados en los dominios mediante un control de acceso detallado. Para crear el rol query_assistant_access de forma manual, siga estos pasos:

1. En los OpenSearch paneles, vaya a Seguridad y elija Roles.
2. Elija Crear rol y configure los siguientes permisos de clúster:
   - `cluster:admin/opensearch/ml/config/get`
   - `cluster:admin/opensearch/ml/execute`
   - `cluster:admin/opensearch/ml/predict`
   - `cluster:admin/opensearch/ppl`
3. Asigne al rol el nombre query_assistant_access.
4. Elija Crear rol. El rol query_assistant_access ya está disponible.

> **Note**
>
> También debe tener los permisos de índice `indices:admin/mappings/get` y `read` para los índices con los que desee utilizar las preguntas en lenguaje natural.

Automatización de la configuración

Flow Framework es un OpenSearch complemento que proporciona una forma de automatizar OpenSearch las configuraciones para casos de uso, como la generación de consultas y el chat conversacional. Como el complemento rastrea los recursos que permiten la característica de generación de consultas en lenguaje natural, el índice de Flow Framework almacena una plantilla para cada dominio que utiliza el asistente para consultas.

Flow Framework te permite seleccionar entre un conjunto de plantillas predefinidas o crear tus propias automatizaciones para los conectores, las herramientas, los agentes y otros componentes del aprendizaje automático que sirven de base OpenSearch para los modelos generativos.

Búsqueda vectorial

La búsqueda vectorial en Amazon OpenSearch Service te permite buscar contenido semánticamente similar mediante incrustaciones de aprendizaje automático en lugar de la tradicional coincidencia de palabras clave. La búsqueda vectorial convierte sus datos (texto, imágenes, audio, etc.) en vectores numéricos de alta dimensión (incrustaciones) que capturan el significado semántico del contenido. Al realizar una búsqueda, OpenSearch compara la representación vectorial de la consulta con los vectores almacenados para encontrar los elementos más similares.

La búsqueda vectorial incluye los siguientes componentes clave.

Campos vectoriales

OpenSearch admite el tipo de `knn_vector` campo para almacenar vectores densos con dimensiones configurables (hasta 16 000).

Métodos de búsqueda

- k-NN (k-vecinos más cercanos): encuentra los k vectores más similares
- K-nN aproximado: utiliza algoritmos como HNSW (Hierarchical Navigable Small World) para realizar búsquedas más rápidas en conjuntos de datos de gran tamaño

Métricas de distancia

Admite varios cálculos de similitud, entre los que se incluyen:

- Distancia euclidiana
- Similitud coseno
- Producto Dot

Casos de uso comunes

La búsqueda vectorial admite los siguientes casos de uso habituales.

- Búsqueda semántica: busca documentos con un significado similar, no solo con palabras clave coincidentes
- Sistemas de recomendación: sugiera productos, contenido o usuarios similares
- Búsqueda de imágenes: busca imágenes visualmente similares
- Detección de anomalías: identifique valores atípicos en los patrones de datos

- RAG (Retrieval Augmented Generation): mejore las respuestas de LLM con un contexto relevante

Integración con machine learning

OpenSearch se integra con los siguientes servicios y modelos de aprendizaje automático:

- Amazon Bedrock: para generar incrustaciones mediante modelos de base
- Amazon SageMaker AI: para la implementación de modelos de aprendizaje automático personalizados
- Modelos Hugging Face: modelos de incrustación previamente entrenados
- Modelos personalizados: sus propios modelos de incrustación entrenados

Con la búsqueda vectorial, puede crear aplicaciones sofisticadas basadas en inteligencia artificial que entiendan el contexto y el significado, y van mucho más allá de las capacidades tradicionales de coincidencia de texto.

Importación de vectores de Amazon S3 a OpenSearch Serverless

Amazon S3 Vectors ofrece el primer almacén de objetos en la nube con compatibilidad nativa para almacenar y consultar vectores. S3 Vectors proporciona un almacenamiento vectorial rentable, elástico y duradero que se puede consultar en función del significado y la similitud semánticos. Ofrece tiempos de respuesta a las consultas inferiores a un segundo y reduce los costos de carga, almacenamiento y consulta de vectores hasta un 90 %.

Amazon S3 Vectors presenta buckets vectoriales S3, que puede utilizar para almacenar y consultar datos vectoriales, así como acceder a ellos, sin aprovisionar ninguna infraestructura. En un bucket vectorial, puede organizar los datos vectoriales en índices vectoriales. Su bucket vectorial puede tener varios índices vectoriales y cada índice vectorial puede contener millones de vectores. Para obtener más información, consulte Cómo trabajar con Amazon S3 Vectors y buckets de vectores en la Guía del usuario de Amazon S3.

Cada vector se compone de los siguientes elementos:

- Clave única
- Datos vectoriales
- Metadatos opcionales en formato JSON

Los índices vectoriales admiten las funciones de distancia euclidiana y cosenoidal para las operaciones de búsqueda de similitudes.

> **Note**
>
> La principal ventaja de los buckets de vectores es su capacidad para almacenar conjuntos de datos masivos a un costo extremadamente bajo y, al mismo tiempo, proporcionar acceso directo a la API para las operaciones con vectores.

Para obtener más información sobre los buckets de Amazon S3 Vectors, incluido cómo crear uno, consulte Uso de vectores y buckets de Amazon S3 Vectors en la Guía del usuario de Amazon S3. Para obtener más información sobre la integración con OpenSearch Service además de lo que se describe en este tema, consulte Uso de S3 Vectors with OpenSearch Service

Puede utilizar S3 Vectors con Amazon OpenSearch Service para reducir el coste del almacenamiento vectorial cuando las consultas son menos frecuentes y, después, trasladar rápidamente esos conjuntos de datos a otros conjuntos de datos a OpenSearch medida que aumente la demanda o para mejorar las capacidades de búsqueda.

OpenSearch El servicio se integra con Amazon S3 Vectors para ofrecer un rendimiento y una funcionalidad mejorados que van más allá de lo que ofrecen los buckets vectoriales de Amazon S3 por sí solos. Considere esta integración cuando necesite:

- Consultas de mayor rendimiento
- Latencia de búsqueda de menos de un segundo
- Capacidades de análisis avanzadas, como las agregaciones
- Búsqueda híbrida que combina datos de texto y vectores

Esta integración es particularmente útil cuando varias aplicaciones consumen los mismos datos de vectores con diferentes requisitos de rendimiento. Puede hacer que algunas aplicaciones interactúen directamente con los buckets vectoriales de Amazon S3 para casos de uso sensibles a los costes, mientras que otras aprovechan la OpenSearch integración para operaciones de rendimiento crítico.

Arquitecturas de integración

La integración utiliza Amazon OpenSearch Ingestion (OSI) como canalización de datos entre los índices vectoriales de Amazon S3 y las colecciones vectoriales de Amazon OpenSearch

Serverless. OpenSearch Ingestion exporta automáticamente los datos vectoriales del índice vectorial especificado y los ingiere en colecciones vectoriales OpenSearch sin servidor para realizar operaciones de búsqueda de alto rendimiento.

> **Note**
>
> Tras la exportación, los datos siguen presentes en el índice vectorial S3. Tiene dos copias de los datos.

Cada índice vectorial se asigna a un índice correspondiente de la OpenSearch colección Serverless. En la integración:

- Conserva las dimensiones vectoriales
- Conserva los metadatos
- Optimiza la estructura de datos para las capacidades OpenSearch de búsqueda vectorial

Tras la configuración, OpenSearch Ingestion comienza el proceso de exportación de datos consumiendo los vectores del índice vectorial especificado mediante la ListVectors API Amazon S3. El servicio procesa los vectores en paralelo para optimizar la velocidad de ingesta y, al mismo tiempo, respeta los límites de escalado de OpenSearch Ingestion y Amazon OpenSearch Serverless.

Durante la ingesta, el servicio hace lo siguiente:

- Transforma los datos vectoriales para que coincidan con el formato esperado para Service OpenSearch
- Conserva la información esencial, incluidos los valores de vectores, los metadatos y las métricas de distancia
- Maneja los escenarios de fallo mediante mecanismos de reintento inteligentes
- Coloca los registros problemáticos en un bucket de Amazon S3 que se utiliza como cola de mensajes fallidos para su posterior análisis

La integración gestiona conjuntos de datos masivos de manera eficiente y el rendimiento depende de las dimensiones de vectores, el tamaño del conjunto de datos y los límites de escalado configurados. OSI puede escalar hasta 16 trabajadores por canalización, mientras que OpenSearch Serverless ajusta automáticamente la capacidad en función de las demandas de ingesta. De forma

predeterminada, OpenSearch aumenta la unidad `maxSearch` OpenSearch computacional (OCU) del lado OpenSearch sin servidor a 100.

> **Note**
>
> La integración prioriza la rentabilidad mediante:
>
> - Cierre automático de la canalización una vez finalizada la exportación
> - OpenSearch Escalado de colecciones sin servidor
> - Pay-per-use modelo de recursos

Permisos de IAM necesarios

La integración requiere una configuración cuidadosa de los permisos de IAM para permitir una comunicación segura entre los servicios. OpenSearch Ingestion necesita permisos para leer los índices vectoriales de Amazon S3, escribir en colecciones vectoriales de OpenSearch Service y gestionar las políticas de seguridad asociadas.

Al habilitar la integración mediante el procedimiento que se describe más adelante en este tema, puede elegir una de las siguientes opciones para la administración de permisos:

- Permita que el sistema cree automáticamente un rol de servicio con los permisos necesarios
- Proporcione un rol existente que cumpla con los requisitos

El rol creado automáticamente incluye políticas para:

- Acceso al índice vectorial de Amazon S3 APIs
- Gestión de las operaciones OpenSearch de recopilación de servicios
- Administración de las operaciones de cola de mensajes fallidos en caso de intentos de ingesta fallidos

Si decide especificar un rol existente, compruebe que dicho rol tenga los siguientes permisos de IAM:

(Obligatorio): permisos de canalización de datos entre OpenSearch Ingestion y 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" ]
               }
            }
        }
    ]
}
```

(Obligatorio): permisos de ingesta de datos entre OpenSearch Ingestion y la cola de cartas muertas de Amazon S3

JSON

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

```
    "Statement": [
        {
            "Sid": "s3Access",
            "Effect": "Allow",
            "Action": [
              "s3:PutObject"
            ],
            "Resource": [ "arn:aws:s3:::bucket/*" ]
        }
    ]
}
```

(Obligatorio): permisos de ingesta de datos entre OpenSearch Ingestion y 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"
            ]
        }
    ]
}
```

(Obligatorio si el AWS KMS cifrado está activado): permisos de descifrado para la comunicación entre OpenSearch Ingestion y 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"
                        ]
                 }
             }
        }
    ]
}
```

Configuración de la integración de Amazon S3 Vectors con OpenSearch

Utilice el siguiente procedimiento para configurar la integración de Amazon S3 Vectors con OpenSearch Serverless.

> **Note**
>
> Si ha iniciado el proceso de configuración de la integración desde la consola de Amazon S3 seleccionando la OpenSearch opción Exportar a en la página de cubos vectoriales, algunos de los pasos del siguiente procedimiento no son aplicables, como se indica en el procedimiento.

Para configurar la integración de Amazon S3 Vectors con OpenSearch Serverless

1. Abre la página Importar índice vectorial S3 a motor OpenSearch vectorial en la consola OpenSearch de Amazon Service. La página se muestra automáticamente si ha hecho clic en Exportar a OpenSearch en la consola de Amazon S3. Si está empezando en la OpenSearch consola, seleccione Integración en el menú de navegación de la izquierda y, a continuación, seleccione Importar índice vectorial de S3.
2. En la sección Source, si comenzó en la consola de Amazon S3, compruebe que el nombre del índice vectorial y su nombre de recurso de Amazon (ARN) ya estén especificados. Si comenzó en la OpenSearch consola, introduzca el índice ARN en el campo ARN del índice vectorial S3.
3. En la sección Agregar cuentas elija una opción. Si elige un rol existente, verifique que tenga todos los permisos necesarios para la integración, como se describe en Permisos de IAM necesarios.
4. (Opcional) Amplíe Configuración adicional. Para Habilitar la redundancia (réplicas activas), recomendamos dejar esta opción seleccionada para los entornos de producción. Al crear la primera colección, OpenSearch Serverless crea dos instancias OCUs: una para la indexación y otra para la búsqueda. Para garantizar la alta disponibilidad, también lanza un conjunto de nodos en espera en otra zona de disponibilidad. Para fines de desarrollo y pruebas, puede deshabilitar la configuración Habilitar redundancia para una colección, lo que elimina las dos réplicas en espera y solo crea instancias de dos. OCUs De forma predeterminada, las réplicas activas redundantes están habilitadas, lo que significa que se OCUs crean instancias de un total de cuatro para la primera recopilación de una cuenta.

   En Añadir AWS KMS clave gestionada por el cliente para el vector Amazon OpenSearch Serverless, elija esta opción para cifrar los datos de la colección vectorial mediante una clave gestionada por el cliente. De forma predeterminada, utiliza un. OpenSearch Clave administrada de AWS
5. Si ha iniciado este proceso haciendo clic en la OpenSearch opción Exportar a de la consola de Amazon S3, en la sección Detalles de exportación se muestran los pasos OpenSearch que se seguirán a continuación. Cuando esté listo para continuar, seleccione Exportar.

   Si inició este proceso en la consola de OpenSearch servicio, en la sección Detalles de importación se enumeran los pasos OpenSearch que se seguirán a continuación. Cuando esté listo para continuar, elija Importar.

   OpenSearch abre la página de historial para mostrar todos los índices vectoriales exports/imports de Amazon S3 en índices de OpenSearch Serverless.

Tras una ingesta correcta, OSI detiene automáticamente la canalización para evitar costes innecesarios y, al mismo tiempo, conservar los datos exportados. OpenSearch Puede supervisar el progreso de la integración mediante CloudWatch métricas y acceder a registros detallados para solucionar problemas.

La OpenSearch colección permanece activa y disponible para consultas una vez completada la ingesta inicial. Puede realizar:

- Búsqueda de similitud
- Agregaciones
- Análisis operativo

Capacidades de búsqueda avanzada con un motor vectorial Amazon S3

Amazon OpenSearch Service ofrece la posibilidad de utilizar Amazon S3 como motor vectorial para índices vectoriales. Esta característica le permite descargar datos vectoriales a Amazon S3 y, al mismo tiempo, mantener las capacidades de búsqueda vectorial en menos de un segundo a un bajo costo.

Con esta función, OpenSearch almacena las incrustaciones vectoriales en un índice vectorial de Amazon S3 y, al mismo tiempo, mantiene otros campos del documento en el almacenamiento del OpenSearch clúster. Este método ofrece las siguientes ventajas:

- Durabilidad: Los datos escritos en S3 Vectors se almacenan en S3, que se ha diseñado para ofrecer una durabilidad de datos del 99,999999999 % (11 nueves).
- Escalabilidad: transfiera grandes conjuntos de datos vectoriales a S3 sin consumir almacenamiento en clústeres.
- Rentabilidad: optimice los costos de almacenamiento para cargas de trabajo con gran cantidad de vectores.

OpenSearch tiene los siguientes requisitos para usar índices vectoriales de S3:

- OpenSearch versión 2.19 o posterior
- OpenSearch Instancias optimizadas
- Última versión del parche para su OpenSearch lanzamiento

Activación de S3 Vectors

Al crear un dominio nuevo o actualizar un dominio existente, puede elegir la opción Habilitar S3 Vectors como motor en la sección Características avanzadas. Esta configuración permite OpenSearch crear un depósito vectorial de S3 al utilizar S3 Vectors como motor. Al habilitar esta opción, OpenSearch configura S3 Vectors para su dominio de la siguiente manera:

1. Crea dos nuevas concesiones en la AWS KMS clave configurada con tu dominio:
   - Una concesión para las tareas de indexación en segundo plano de S3 Vectors con privilegios de descifrado
   - Una subvención para OpenSearch crear cubos de vectores S3 con permisos `GenerateDataKey`
2. Configurar la clave KMS utilizada por su OpenSearch dominio como CMK para el cifrado del resto de los datos del índice vectorial.

Creación de índices con el motor vectorial de S3

Después de configurar un dominio, puede crear uno o más índices k-NN con campos, con el uso de `s3vector` como motor vectorial de backend en las asignaciones de índices. Puede configurar diferentes campos vectoriales con diferentes tipos de motores en función de su caso de uso.

> **Important**
>
> Solo puede utilizar el motor `s3vector` para asignar una definición de campo durante la creación del índice. No puede agregar ni actualizar la asignación con el motor `s3vector` después de crear el índice.

Estos son algunos ejemplos que crean índices de motores vectoriales de S3.

Ejemplo: creación de un índice k-NN con el motor vectorial de 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"
        }
    }
  }
}
```

Ejemplo: creación de un índice k-NN con motores vectoriales de S3 y FAISS

Este ejemplo destaca el hecho de que se pueden utilizar varios motores vectoriales dentro del mismo índice.

```
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
                }
            }
        }
    }
  }
}
```

Ejemplo no compatible: agregar el motor vectorial de S3 tras la creación del índice

El siguiente enfoque no es compatible y fallará.

```
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"
        }
```

```
    }
}
```

Limitaciones de funcionalidad

Tenga en cuenta las siguientes limitaciones antes de utilizar el motor `s3vector` en un índice:

Características y comportamientos no compatibles con el motor s3vector

Característica	Comportamiento
Índice de Split/Shrink/Clone	Estos APIs errores se producen cuando se utilizan con un índice configurado con el `s3vector` motor en el `knn_vector` campo.
Snapshots	Los índices que utilizan el motor `s3vector` no admiten instantáneas. Para dominios administrados: • Las instantáneas automatizadas solo incluyen índices que no utilizan el motor `s3vector`. • Las solicitudes manuales de instantáneas para los índices `s3vector` fallan. **Note** Si bien las instantáneas no son compatibles con la point-in-time recuperación, el `s3vector` motor, junto con las instancias OpenSearch optimizadas, ofrecen una durabilidad de 11 nueves.
UltraWarm nivel	Los índices configurados con `s3vector` el motor no pueden migrar al UltraWarm nivel.
Replicación entre clústeres	Los índices configurados con el motor `s3vector` no admiten la replicación entre clústeres.
Protección contra eliminaci ones accidentales	Como las instantáneas no son compatibles con los índices que utilizan el motor `s3vector`, la protección contra la eliminaci

Característica	Comportamiento
	ón accidental no está disponible. Aún puedes restaurar otros índices en el dominio.
Búsqueda radial	Las consultas con búsqueda radial no se admiten en los campos que utilizan el motor `s3vector`.

Indexación de documentos

Tras crear un índice con el motor vectorial S3, puede ingerir documentos mediante la `_bulk` API estándar. OpenSearch transfiere automáticamente los datos vectoriales de los `knn_vector` campos que utilizan el `s3vector` motor al índice vectorial S3 en tiempo real. Los datos que pertenezcan a otros campos o a `knn_vector` campos que utilicen diferentes motores se conservarán OpenSearch en su propia capa de almacenamiento.

Para todas las solicitudes masivas que se confirmen, OpenSearch garantiza que todos los datos (vectoriales y no vectoriales) sean duraderos. Si una solicitud recibe un acuse de recibo negativo, no se garantiza la durabilidad de los documentos incluidos en esa solicitud masiva. Deberías volver a intentar realizar dichas solicitudes, preferiblemente después de eliminar la solicitud anterior fallida utilizando el identificador del documento para evitar la duplicación de documentos en estos raros casos.

Ejemplo de indexación masiva

```
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 }
```

Buscar documentos

Puede buscar en su índice mediante la API `_search` estándar para ejecutar consultas de texto, k-NN o híbridas. En el caso de consultas en `knn_vector` campos configurados con el `s3vector` motor, transfiere OpenSearch automáticamente la consulta al índice de vectores S3 correspondiente.

> **Note**
>
> Con el `s3vector` motor, las consultas de búsqueda k-NN admiten un valor máximo k de 100. Esto significa que en los resultados de la búsqueda se puede encontrar un máximo de 100 vecinos más cercanos.

Ejemplo de consulta de búsqueda

```
GET my-first-s3vector-index/_search
{
  "size": 2,
  "query": {
    "knn": {
      "my_vector_1": {
        "vector": [2.5, 3.5],
        "k": 2
      }
    }
  }
}
```

Puede ejecutar una búsqueda vectorial filtrada en el índice OpenSearch kNN utilizando el motor s3vector. OpenSearch aplica el filtro como filtro posterior y utiliza un mecanismo de sobremuestreo basado en ciertas heurísticas para equilibrar la recuperación con la latencia.

Ejemplo de consulta de búsqueda 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
            }
          }
        }
      }
    }
  }
}
```

Parámetros de asignación compatibles

Con el motor `s3vector`, el campo `knn_vector` admite los siguientes parámetros en las asignaciones.

Parámetros de campo vectorial

Parámetro	Obligatorio	Description (Descripción)	Valores admitidos
`type`	Sí	El tipo de campo presente en el documento.	`knn_vector`
`dimension`	Sí	La dimensión de cada vector que se ingerirá en el índice.	>0, <=4096
`space_type`	No	El espacio vectorial utilizado para calcular la distancia entre los vectores.	`l2`, `cosinesimil`
`method.engine`	Sí	El motor k-NN aproximado que se utilizará para la indexación y la búsqueda.	`s3vector`
`method.name`	No	El método del vecino más cercano	""

Parámetro	Obligatorio	Description (Descripción)	Valores admitidos
`store`	N/A	No es posible activar o desactivar este parámetro de mapeo, ya que los datos de knn_vector no se almacenan en él. OpenSearch	No es compatible
`doc_values`	N/A	No es posible activar o desactivar este parámetro de mapeo, ya que los datos de knn_vector no se almacenan en él. OpenSearch	No es compatible

> **Important**
>
> El motor `knn_vector` no admite los tipos de campos anidados de `s3vector`

Facturación y medición

Para obtener información sobre la medición y la facturación de esta función, consulta los precios de Amazon OpenSearch Service.

Deshabilitación del motor s3vector

Antes de deshabilitar el motor `s3vector`, elimine todos los índices que lo utilizan actualmente. Si no lo hace, cualquier intento de deshabilitar el motor fallará.

Tenga en cuenta también que al habilitar o deshabilitar el motor `s3vector` se desencadena una implementación azul/verde en su dominio.

Para deshabilitar el motor `s3vector`, edite la configuración de su dominio y configure `S3VectorsEngine.Enabled: false`.

Búsqueda de K-vecino más cercano (K-nn) en Amazon Service OpenSearch

K-nn for Amazon OpenSearch Service, abreviatura de su algoritmo asociado de k-vecinos más cercanos, te permite buscar puntos en un espacio vectorial y encontrar los «vecinos más cercanos» de esos puntos por distancia euclidiana o similitud de coseno. Los casos de uso incluyen recomendaciones (por ejemplo, una característica de “otras canciones que podrían gustarte” en una aplicación de música), reconocimiento de imágenes y detección de fraude.

> **Note**
>
> Esta documentación proporciona una breve descripción general del complemento k-NN, así como las limitaciones a la hora de utilizarlo con un servicio gestionado. OpenSearch Para obtener una documentación completa del complemento k-NN, que incluye ejemplos simples y complejos, referencias de parámetros y la referencia completa de la API, consulta la documentación de código abierto. OpenSearch La documentación de código abierto también incluye el ajuste del rendimiento y la configuración k-NN-specific del clúster.

Introducción a k-NN

Para utilizar k-NN, debe crear un índice con la configuración `index.knn` y agregar uno o más campos del tipo de datos `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
```

```
      }
    }
  }
}
```

El tipo de datos `knn_vector` admite una sola lista de hasta 10 000 flotadores, con el número de flotadores definido por el parámetro `dimension` requerido. Después de crear el índice, agregue algunos datos.

```
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 }
```

A continuación, puede buscar los datos mediante el tipo de consulta `knn`.

```
GET my-index/_search
{
  "size": 2,
  "query": {
    "knn": {
      "my_vector2": {
        "vector": [2, 3, 5, 6],
        "k": 2
      }
    }
```

```
  }
}
```

En este caso, k es el número de vecinos que desea que devuelva la consulta, pero también debe incluir la opción `size`. De lo contrario, obtendrá k resultados para cada partición (y cada segmento) en lugar de k resultados para toda la consulta. k-NN admite un valor de k máximo de 10 000.

Si mezcla la consulta knn con otras cláusulas, es posible que reciba menos resultados k. En este ejemplo, la cláusula `post_filter` reduce el número de resultados de 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
      }
    }
  }
}
```

Si necesita gestionar un gran volumen de consultas y, al mismo tiempo, mantener un rendimiento óptimo, puede utilizar la API `_msearch` para crear una búsqueda masiva con JSON y enviar una única solicitud para realizar varias búsquedas:

```
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 }} } }
```

En el siguiente video se muestra cómo configurar búsquedas vectoriales masivas para consultas K-NN.

Diferencias, ajuste y limitaciones de k-NN

OpenSearch permite modificar todos los ajustes de k-NN mediante la `_cluster/settings` API. En el OpenSearch Servicio, puedes cambiar todos los ajustes excepto `knn.memory.circuit_breaker.enabled` y`knn.circuit_breaker.triggered`. Las estadísticas de K-nn se incluyen como métricas de Amazon CloudWatch .

En concreto, compara la `KNNGraphMemoryUsage` métrica de cada nodo de datos con la `knn.memory.circuit_breaker.limit` estadística y la RAM disponible para el tipo de instancia. OpenSearch El servicio utiliza la mitad de la RAM de una instancia para el montón de Java (hasta un tamaño de pila de 32 GiB). De forma predeterminada, k-NN utiliza hasta el 50 % de la mitad restante, por lo que un tipo de instancia con 32 GiB de RAM puede acomodar 8 GiB de gráficos (32 * 0,5 * 0,5). El rendimiento puede verse afectado si el uso de la memoria gráfica supera este valor.

Puedes migrar un índice k-NN creado en la versión 2.x o posterior a un dominio con la versión 2.17 UltraWarmo posterior o almacenarlo en frío.

La API de borrado de caché y las API de calentamiento para los índices k-NN están bloqueadas para los índices cálidos. Cuando se inicia la primera consulta para el índice, descarga los archivos de gráficos de Amazon S3 y carga el gráfico en la memoria. Del mismo modo, cuando el TTL de los gráficos caduca, los archivos se eliminan automáticamente de la memoria.

Ingestión de vectores

La ingesta vectorial le ayuda a incorporar e indexar OpenSearch dominios y OpenSearch colecciones sin servidor con rapidez. El servicio examina su dominio o colección y crea una canalización de ingestión en su nombre para cargar sus datos. OpenSearch Vector ingestion gestiona por usted la ingesta e indexación de su dominio o colección.

Puede acelerar y optimizar el proceso de indexación habilitando las funciones. Aceleración de GPU para indexación vectorial Optimización automática Con la ingestión de vectores, no necesita administrar la infraestructura subyacente, parchear el software ni escalar los clústeres para respaldar la indexación e ingestión de sus bases de datos vectoriales. Esto le permite crear rápidamente su base de datos vectorial para satisfacer sus necesidades.

Funcionamiento

La ingestión de vectores examina su dominio o colección y su índice. Puede configurar manualmente los campos de índice vectorial o OpenSearch permitir el uso de la configuración automática.

La ingestión de vectores utiliza OpenSearch la ingestión (OSI) como canalización de datos entre Amazon S3 y. OpenSearch El servicio procesa los vectores en paralelo para optimizar la velocidad de ingesta y, al mismo tiempo, respetar los límites de escalado de OSI y OpenSearch.

OpenSearch Precios de ingestión vectorial

En un momento específico, solo pagas por la cantidad de vectores ingeridos OCUs que se asignen a una canalización, independientemente de si hay datos circulando por la canalización. OpenSearch La ingestión vectorial se adapta inmediatamente a tus cargas de trabajo al aumentar o reducir la capacidad de la canalización en función del uso.

Para obtener información completa sobre los precios, consulta los precios OpenSearch de Amazon Service.

Requisitos previos

Antes de utilizar la ingestión de vectores, asegúrate de disponer de los siguientes recursos:

- Depósito de Amazon S3 que contiene sus documentos OpenSearch JSON en formato Parquet o JSONL
- OpenSearch recurso: un dominio o una colección
- OpenSearch versión `2.19` o posterior (necesaria para la optimización automática de la integración)

Crear una base de datos vectorial

Utilice el flujo de trabajo de creación de trabajos de ingestión de vectores para configurar el ajuste automatizado de índices vectoriales y acelerar la creación de índices a gran escala.

> **Note**
>
> El contenido procedimental de esta sección está sujeto a cambios a medida que se finalice la interfaz de usuario. Es posible que el flujo de trabajo se actualice en futuras versiones para reflejar la experiencia de consola más reciente.

Para crear un trabajo de ingestión de vectores

1. En la sección de detalles del trabajo de ingestión de vectores, en Nombre, introduzca un nombre para el trabajo de ingestión.
2. En la sección Fuente de datos, configure lo siguiente:
   a. Para el URI de Amazon S3, introduzca la ubicación del bucket de Amazon S3 que contiene los documentos JSON de OpenSearch servicio.
   b. Elija Browse Amazon S3 para seleccionar entre los buckets disponibles o seleccione View para obtener una vista previa del contenido del bucket.
   c. En Tipo de contenido, selecciona una de las siguientes opciones:
      - Vectores: los documentos ya contienen vectores y no es necesario generar más vectores incrustados.
      - Texto, imagen o audio: los documentos contienen contenido, como texto, imágenes o bytes de audio, que deben codificarse en incrustaciones vectoriales.
3. En la sección Permisos de la fuente de datos, configure los permisos de acceso:
   a. En Rol de IAM, elija una de las opciones siguientes:
      - Create a new role (Crear un nuevo rol)
      - Utilice un rol existente
   b. En el nombre del rol de IAM, introduzca un nombre para el rol.
4. En la sección Destino, configure el punto final del OpenSearch servicio:
   a. Para Endpoint, selecciona Elige una opción para seleccionar entre tus dominios o colecciones compatibles en la región actual.
   b. Seleccione Siguiente para continuar con el punto final seleccionado.
5. Seleccione Siguiente para continuar con el paso siguiente o Cancelar para salir sin guardar nada.

Funciones relacionadas

La ingestión de vectores funciona con las siguientes funciones de Amazon OpenSearch Service para optimizar el rendimiento de las bases de datos vectoriales:

Aceleración de GPU para indexación vectorial

La aceleración de la GPU reduce el tiempo necesario para crear, actualizar y eliminar índices vectoriales. Si se utiliza con la ingesta de vectores, se puede acelerar considerablemente el proceso de ingesta e indexación de bases de datos vectoriales a gran escala.

Optimización automática

La optimización automática descubre automáticamente las compensaciones óptimas entre la latencia de búsqueda, la calidad y los requisitos de memoria. La ingestión de vectores puede aplicar recomendaciones de optimización automática durante el proceso de ingestión para garantizar que los índices vectoriales estén configurados de forma óptima.

Para obtener los mejores resultados, considere habilitar tanto la aceleración de la GPU como la optimización automática cuando utilice la ingestión de vectores para crear bases de datos vectoriales a gran escala.

Exporte el índice vectorial de Amazon S3 al motor vectorial de OpenSearch servicios

Una point-in-time exportación del índice vectorial de Amazon S3 seleccionado a OpenSearch Service. El motor vectorial OpenSearch Service proporciona un almacén vectorial simple y escalable con funcionalidad de búsqueda avanzada.

Para exportar el índice vectorial de Amazon S3 al motor vectorial de OpenSearch servicios

1. En la sección Source, compruebe los detalles del índice vectorial de Amazon S3:
   - Índice vectorial de Amazon S3: el nombre del índice de origen
   - Índice vectorial de Amazon S3 (ARN): el nombre del recurso de Amazon de su índice
2. En la sección Acceso al servicio, configure la autorización OpenSearch del servicio:
   a. En Elija un método para autorizar el OpenSearch servicio, seleccione una de las siguientes opciones:

- Crear y usar un nuevo rol de servicio
- Usar un rol de servicio existente

b. En Nombre del rol de servicio, introduzca un nombre para el rol de servicio.

> **Note**
>
> El nombre del rol de servicio debe tener entre 1 y 64 caracteres. Los caracteres válidos son a-z, A-Z, 0-9 y puntos (.).

c. Seleccione Ver detalles de los permisos para revisar los permisos necesarios.

3. Amplíe Configuración adicional (opcional) para configurar opciones avanzadas si es necesario.
4. En la sección Detalles de exportación, configure las siguientes opciones:

- Automatice OpenSearch la creación de colecciones de vectores de OpenSearch servicios: las colecciones de servicios se utilizan para almacenar datos vectoriales. La capacidad de cómputo sin servidor se mide en unidades de cómputo de OpenSearch servicio (OCUs); de forma predeterminada, la capacidad máxima de la OCU es de 50.
- Automatice la creación de roles de IAM para el acceso al servicio: OpenSearch Service utiliza este rol para leer el índice vectorial de Amazon S3 y escribir en la colección de OpenSearch servicios.
- Automatice OpenSearch la creación de canalizaciones de ingestión de OpenSearch servicios: las canalizaciones de ingestión de servicios se utilizan para ingerir datos. Se ha creado un bucket de Amazon S3 como práctica recomendada para capturar y almacenar los eventos fallidos en un bucket de Amazon S3 Dead Letter Queue (DLQ), lo que permite un fácil acceso para la resolución de problemas y el análisis.

5. Seleccione Exportar para iniciar el proceso de exportación o Cancelar para salir sin exportar.

Importación del espacio de nombres vectorial de Amazon S3 al motor vectorial OpenSearch de servicios

El análisis de sus datos vectoriales con el OpenSearch Servicio requiere una recopilación única del OpenSearch Servicio y una configuración de permisos de IAM.

Para importar el espacio de nombres vectorial de Amazon S3 al motor vectorial OpenSearch de servicio

1. En la sección Source, configure el índice vectorial de Amazon S3:

   - Para el ARN del índice vectorial de Amazon S3, introduzca el ARN de su índice vectorial de Amazon S3.

     > **Note**
     >
     > Debe tener el formato arn:aws:iam: :account-id: /*:index vector-bucket-name

2. En la sección Acceso al servicio, configure la autorización del OpenSearch servicio:

   a. En Elija un método para autorizar el OpenSearch servicio, seleccione una de las siguientes opciones:

      - Crear y usar un nuevo rol de servicio
      - Usar un rol de servicio existente

   b. En Nombre del rol de servicio, introduzca un nombre para el rol de servicio.

      > **Note**
      >
      > El nombre del rol de servicio debe tener entre 1 y 64 caracteres. Los caracteres válidos son a-z, A-Z, 0-9 y puntos (.).

   c. Seleccione Ver detalles de los permisos para revisar los permisos necesarios.

3. Amplíe Configuración adicional (opcional) para configurar opciones avanzadas si es necesario.
4. En la sección Pasos de importación, configura las siguientes opciones de automatización:

   - Automatice OpenSearch la creación de colecciones de vectores de OpenSearch servicios: las colecciones de servicios se utilizan para almacenar datos vectoriales. La capacidad de cómputo sin servidor se mide en unidades de cómputo de OpenSearch servicio (OCUs); de forma predeterminada, la capacidad máxima de la OCU es de 50.
   - Automatice la creación de roles de IAM para el acceso al servicio: OpenSearch Service utiliza este rol para leer el índice vectorial de Amazon S3 y escribir en la colección de OpenSearch servicios.

- Automatice OpenSearch la creación de canalizaciones de ingestión de OpenSearch servicios: las canalizaciones de ingestión de servicios se utilizan para ingerir datos. Se ha creado un bucket de Amazon S3 como práctica recomendada para capturar y almacenar los eventos fallidos en un bucket de Amazon S3 Dead Letter Queue (DLQ), lo que permite un fácil acceso para la resolución de problemas y el análisis.

5. Seleccione Importar para iniciar el proceso de importación o Cancelar para salir sin importar.

Ver los trabajos de ingestión de vectores y el historial de importación

Los trabajos de ingestión de vectores crean una canalización para vectorizar conjuntos de datos, automatizar el ajuste de los índices vectoriales y acelerar la creación de índices a gran escala.

Para ver los trabajos de ingestión de vectores

1. En la sección de trabajos de ingestión de vectores, consulte la información resumida:
   - Trabajos: número total de trabajos de ingestión
   - Seleccione Crear base de datos vectorial para crear un nuevo trabajo de ingestión
2. En la sección de importaciones de vectores de Amazon S3, consulte el resumen de la importación:
   - Importaciones totales: número de importaciones finalizadas
   - Elija Importar vectores de Amazon S3 para iniciar una nueva importación
3. En la tabla de trabajos de ingestión de vectores, supervise los trabajos activos con la siguiente información:
   - Nombre: el nombre del trabajo
   - Estado: estado actual del trabajo (por ejemplo, activo)
   - Fuente de datos: ubicación de la fuente (p. ej., s3://location)
   - Destino: destino objetivo
   - Última actualización: fecha y hora de la actualización más reciente
4. Utilice el cuadro de búsqueda para encontrar un trabajo de ingestión de vectores para localizar trabajos específicos.
5. Para gestionar los trabajos, elija una de las siguientes acciones:
   - Elija Eliminar para eliminar los trabajos seleccionados

- Elija Crear base de datos vectorial para crear trabajos adicionales

6. En la sección del historial de importación de vectores de Amazon S3, realice un seguimiento de los eventos de importación:

   a. Utilice el filtro de intervalo de fechas para especificar un período de tiempo para el historial de importaciones.

   b. Usa el menú desplegable Estado para filtrar por estado de importación (por ejemplo, Cualquier estado).

   c. Utilice el cuadro de búsqueda para buscar importaciones por índice vectorial de Amazon S3 na... para localizar importaciones específicas.

   d. Vea los detalles de la importación, que incluyen:

      - La importación se inició el (UTC+ 5:30): cuando se inició la importación
      - Estado de la importación: estado actual (en curso, completa, fallida, parcialmente completada)
      - ARN del índice vectorial de Amazon S3: identificador del índice de origen
      - OpenSearch Colección de vectores de servicio: colección de destino

7. Elija Importar vector de Amazon S3 para iniciar un nuevo proceso de importación.

Optimización automática

La optimización automática es un servicio que automatiza las optimizaciones de los índices vectoriales, lo que permite a los usuarios equilibrar la calidad, la velocidad y el coste de la búsqueda sin necesidad de semanas de ajustes manuales por parte de un experto. Evalúa las configuraciones de los índices en función de los requisitos de latencia y recuperación definidos por el usuario y genera recomendaciones de optimización, por lo que se requiere una experiencia mínima. Por lo general, las recomendaciones se entregan en un plazo de 30 a 60 minutos.

La configuración tradicional de índices vectoriales requiere una gran experiencia y experimentación para lograr un rendimiento óptimo. Parámetros como `ef_construction` (que controla la calidad de creación del índice), `m` (que determina el número de conexiones gráficas), `ef_search` (que controla la búsqueda en HNSW) y los métodos de cuantificación (cuantificación binaria (32x, 16x, 8x) y cuantificación escalar (4x)) tienen un impacto significativo tanto en la precisión de la búsqueda como en la utilización de los recursos. La optimización automática utiliza algoritmos de optimización

de hiperparámetros para descubrir las configuraciones de índice que son especialmente óptimas para su conjunto de datos, dentro de los requisitos de latencia y recuperación definidos.

Ventajas

La optimización automática para OpenSearch ofrece las siguientes ventajas:

- Ajuste automático de parámetros: elimina la experimentación manual con los parámetros del algoritmo (HNSW), la cuantificación, la repuntuación y el motor, lo que ahorra tiempo y reduce la curva de aprendizaje para la optimización de la búsqueda vectorial.
- Optimice la velocidad de búsqueda: de forma predeterminada, OpenSearch está configurado para funcionar en memoria. La optimización automática descubre ventajas y desventajas que mejoran la calidad de la búsqueda y ahorran costes, a la vez que mantienen una velocidad de búsqueda aceptable.
- Optimización de costos: reduce los costos al encontrar opciones para reducir los requisitos de memoria de indexación y, al mismo tiempo, minimizar las compensaciones entre calidad y velocidad de búsqueda.
- Optimice la calidad de las búsquedas: puede ofrecer una mayor capacidad de recuperación que la configuración predeterminada, o bien descubra ventajas que supongan un importante ahorro de costes con una pérdida mínima de recuperaciones.

La optimización automática funciona junto con otras OpenSearch funciones, por ejemplo, Aceleración de GPU para indexación vectorial para proporcionar una optimización integral del rendimiento de las cargas de trabajo de búsqueda vectorial.

Funcionamiento

La optimización automática funciona mediante una arquitectura basada en tareas que analiza los datos vectoriales y proporciona recomendaciones de optimización. Puntos clave:

- Los usuarios comparten sus conjuntos de datos en formato Parquet o JSONL en un bucket de Amazon S3.
- Configuran los trabajos de optimización automática sin servidor mediante la configuración de sus umbrales aceptables de recuperación y latencia. Unos umbrales más relajados permiten al servicio descubrir optimizaciones de costes más significativas.
- Optimice automáticamente los trabajos ejecutados en una infraestructura totalmente gestionada por Amazon OpenSearch Service. Los trabajos no consumen recursos de tu dominio ni de tus

colecciones. Los trabajadores trabajan en paralelo para evaluar las configuraciones de los índices y utilizan el muestreo en conjuntos de datos de gran tamaño para obtener resultados, por lo general, en un plazo de 30 a 60 minutos.

- Cada trabajo se factura con una tarifa fija predecible. Para obtener información sobre precios, consulta Amazon OpenSearch Service Pricing.

Requisitos previos

- Formato y permisos del conjunto de datos: debe tener su conjunto de datos disponible como uno o más archivos Parquet o JSONL en una carpeta de bucket de Amazon S3. Por ejemplo:
  - Parquet: y `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`

  Proporcione el URI de la carpeta adjunta (por ejemplo,). `s3://dataset-bucket-us-east-1/dataset_folder/` La carpeta debe contener archivos de un solo formato; no mezcle archivos Parquet y JSONL en la misma carpeta. Este conjunto de datos se utilizará para generar las recomendaciones. Asegúrese de que su función federada tenga los siguientes permisos de Amazon S3 en ese recurso:`"s3:Get*", "s3:List*", "s3:Describe*"`.
- Especifique los metadatos correctos del conjunto de datos: el conjunto de datos proporcionado debe contener filas de valores flotantes. El nombre de cada columna y la dimensionalidad de cada vector deben coincidir con las opciones proporcionadas en la consola. Por ejemplo, si el conjunto de datos contiene vectores `train_data` cuyo nombre corresponde a cada 768 dimensión, estos valores deben coincidir con la consola de optimización automática.
- (Si usa la función de ingestión de vectores): si planea utilizar la función de ingestión (siguiendo las recomendaciones de optimización automática para crear datos de índice e ingestión automáticamente), debe configurar su OpenSearch clúster para conceder el permiso de optimización automática a fin de incorporar su conjunto de datos al clúster. OpenSearch En el caso de OpenSearch los dominios con una política de acceso a dominios, conceda acceso al rol recién creado mediante esa política. En el OpenSearch caso de los dominios con un control de acceso detallado, añada la función de canalización como función de back-end. En el caso de las colecciones OpenSearch sin servidor, añada la función de canalización a la política de acceso a los datos.
- Permisos de IAM: necesita los siguientes permisos de IAM para utilizar la optimización automática:
  - `opensearch:SubmitAutoOptimizeJob`

- `opensearch:GetAutoOptimizeJob`
- `opensearch:DeleteAutoOptimizeJob`
- `opensearch:CancelAutoOptimizeJob`
- `opensearch:ListAutoOptimizeJobs`

> **Note**
>
> Se trata de políticas basadas en la identidad. La optimización automática no admite políticas basadas en recursos.

- Caducidad de las credenciales: configure la sesión de usuario federado para que las credenciales tengan una caducidad mínima de al menos 1 hora. Para conjuntos de datos muy grandes o de grandes dimensiones, considere la posibilidad de aumentar la duración de caducidad hasta 3 horas.

Casos de uso para la optimización automática

La optimización automática es particularmente valiosa en los siguientes escenarios:

Optimización de la configuración inicial

Al implementar por primera vez aplicaciones de búsqueda vectorial, determinar los parámetros óptimos de HNSW a menudo requiere pruebas exhaustivas y experiencia en el campo. La optimización automática elimina este trial-and-error proceso al analizar las características de los datos y la carga de trabajo para recomendar configuraciones listas para la producción.

Este caso de uso es ideal para los equipos que se inician en la búsqueda vectorial o para aquellos que migran desde otras plataformas de bases de datos vectoriales y necesitan establecer rápidamente las configuraciones de referencia.

Optimización del escalado

A medida que el conjunto de datos vectoriales pase de miles a millones de vectores, es posible que los parámetros que funcionaron bien inicialmente no sean óptimos. La optimización automática recomienda ajustes para mantener el rendimiento a escala.

Reducción de costos

Los índices vectoriales pueden consumir importantes recursos de cómputo y almacenamiento, especialmente en el caso de incrustaciones de alta dimensión. La optimización automática identifica las oportunidades de reducir los costos mediante la búsqueda de configuraciones de parámetros más eficientes que mantienen los niveles de rendimiento requeridos y, al mismo tiempo, utilizan menos recursos.

Por ejemplo, la optimización automática podría descubrir que su `m` valor actual (conectividad de gráficos) es superior al necesario para sus requisitos de precisión, lo que le permitirá reducir el tiempo de indexación y el almacenamiento sin afectar a la calidad de la búsqueda.

Soluciones de problemas de rendimiento

Cuando el rendimiento de las consultas es lento o se produce una latencia alta en las operaciones de búsqueda vectorial, la optimización automática puede analizar el conjunto de datos e identificar una configuración más óptima. El servicio proporciona recomendaciones específicas para abordar los cuellos de botella en el rendimiento, como ajustar la conectividad de los gráficos o los parámetros de búsqueda.

Limitaciones

- Disponibilidad regional: la optimización automática solo está disponible en las siguientes regiones: 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
- Tipos de colecciones: la optimización automática solo se admite en las colecciones y OpenSearch dominios de búsqueda vectorial (2.19, 3.1 y 3.3).
- Soporte de motor

Soporte de motores por tipo de despliegue

Motor	Sin servidor	OpenSearch Gestionado
Lucene	No	Sí
Faiss	Sí	Sí
Nmslib	No	No

- Compatibilidad con algoritmos: la optimización automática solo admite índices vectoriales basados en HNSW.
- Trabajos simultáneos: puede ejecutar hasta 10 trabajos de optimización simultáneos por cuenta y región. No se aceptarán nuevos trabajos si se alcanza el límite.
- Duración del trabajo: los trabajos de optimización pueden tardar entre 15 minutos y varias horas, según el tamaño del conjunto de datos, las dimensiones y las métricas de rendimiento requeridas.
- Recomendaciones: la optimización automática sugiere solo un máximo de 3 recomendaciones.
- Conjunto de datos
  - Formatos compatibles: Parquet, JSONL
  - Almacén de datos: Amazon S3

Facturación y costos

La optimización automática utiliza un modelo de precios por trabajo en el que se paga por cada trabajo de optimización exitoso, independientemente del tamaño del conjunto de datos y de las configuraciones de optimización. No se le cobrará por los trabajos fallidos o cancelados. Además, la optimización automática se ejecuta en una infraestructura diferente a la de los OpenSearch clústeres gestionados o sin servidor, por lo que no afecta a la utilización de los recursos de los clústeres preexistentes.

Modelo de precios

Los costes de optimización automática se facturan por separado de los costes de almacenamiento e informática estándar de dominios OpenSearch gestionados o OpenSearch sin servidor.

Para obtener información sobre precios, consulta Amazon OpenSearch Service Pricing.

Formatos de datos admitidos

La optimización automática admite los siguientes formatos de datos para los conjuntos de datos vectoriales almacenados en Amazon S3:

Formato Parquet

Parquet es un formato de almacenamiento en columnas optimizado para cargas de trabajo analíticas. Cada archivo de Parquet debe contener una columna de matrices flotantes que representen sus datos vectoriales.

Ejemplo de estructura de archivos de Parquet (vista como una tabla):

```
| 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 (líneas JSON) es un formato de texto en el que cada línea es un objeto JSON válido. Cada línea debe contener un campo con una matriz flotante que represente los datos vectoriales.

Ejemplo de archivo JSONL:

```
{"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]}
```

Conversión entre formatos

Si los datos están en un formato diferente, puede usar los siguientes scripts de Python para convertirlos.

Convierta JSON o JSONL a 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)
```

Convierte Parquet a 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)
```

Uso de la optimización automática en la consola

Puede utilizar la consola de Amazon OpenSearch Service para crear trabajos de ingestión de vectores, supervisar su progreso, ver las recomendaciones de optimización y crear índices basados en esas recomendaciones.

Requisitos previos

Para poder utilizar la optimización automática en la consola, debe disponer de lo siguiente:

- Una AWS cuenta activa con acceso a la OpenSearch consola.
- Una colección existente OpenSearch sin servidor de tipo búsqueda vectorial o OpenSearch dominio gestionado.
- Permisos de IAM para las siguientes acciones:
  - `opensearch:SubmitAutoOptimizeJob`
  - `opensearch:GetAutoOptimizeJob`
  - `opensearch:DeleteAutoOptimizeJob`
  - `opensearch:CancelAutoOptimizeJob`
  - `opensearch:ListAutoOptimizeJobs`

  > **Note**
  >
  > Se trata de políticas basadas en la identidad. AWS no admite políticas basadas en recursos para optimizar automáticamente los recursos.

- Configure su sesión de usuario federado para que tenga una caducidad mínima de las credenciales de al menos 1 hora. Para conjuntos de datos muy grandes o de dimensiones elevadas, considere la posibilidad de aumentar la duración de caducidad hasta 3 horas.

Crear un trabajo de ingestión de vectores

Un trabajo de ingestión de vectores analiza los datos vectoriales y proporciona recomendaciones de optimización para la configuración del índice.

Para crear un trabajo de ingestión de vectores

1. Inicia sesión en la consola OpenSearch de Amazon Service en Consola de administración de AWS.
2. En el panel de navegación, selecciona Optimización automática.
3. Seleccione Crear trabajo de ingestión de vectores.
4. En Detalles del trabajo, introduzca un nombre para su trabajo de ingestión de vectores. Este nombre le ayuda a identificar el trabajo en la consola.
5. En Fuente de datos, configure lo siguiente:

   a. Para el URI de Amazon S3, introduzca el URI de Amazon S3 de la carpeta que contiene los archivos de datos (Parquet o JSONL). El URI debe apuntar a la carpeta adjunta, no a

archivos individuales. Por ejemplo, si el archivo está en `s3://my-bucket/my-folder/file1.parquet` o`s3://my-bucket/my-folder/data.jsonl`, introduzcas`3://my-bucket/my-folder/`.

> **Note**
>
> La carpeta debe contener archivos de un solo formato. No mezcle archivos Parquet y JSONL en la misma carpeta.

b. En Región, selecciona la AWS región en la que se encuentra tu bucket de Amazon S3. La región debe coincidir con la ubicación del depósito.

6. En OpenSearch Dominio, selecciona un dominio o colección existente o selecciona Crear nuevo para crear uno.

> **Note**
>
> Puede especificar un dominio OpenSearch gestionado o una colección OpenSearch sin servidor y sin servidor.

7. En Permisos de fuente de datos, especifique el rol de IAM que tiene permisos para acceder a su bucket, OpenSearch dominio o colección de Amazon S3. El rol debe tener los permisos necesarios en función de la configuración de su dominio o colección:

   - En el OpenSearch caso de los dominios con una política de acceso al dominio, conceda acceso al rol a través de esa política.
   - En el OpenSearch caso de los dominios con un control de acceso detallado, añada la función como función de back-end.
   - En el caso de las colecciones OpenSearch sin servidor, añada la función a la política de acceso a los datos.

8. Elija Siguiente.

9. En Configurar índice, especifique lo siguiente:

   a. En Nombre de campo, introduzca el nombre del campo del conjunto de datos que contiene los datos vectoriales.

   b. En Tipo de espacio, seleccione la métrica de distancia utilizada para calcular la distancia entre los vectores:

- l2: distancia euclidiana
- cosinesimil - Similitud de coseno
- producto interno - Producto interno

c. En Dimensión, introduzca el número de valores de punto flotante en cada vector.

10. En Requisitos de rendimiento, configure lo siguiente:

a. Para Recall, especifique la calidad de búsqueda deseada como un valor decimal entre 0 y 1. Los valores de recuperación más altos devuelven resultados más relevantes. Por ejemplo:

- Un valor de 0,95 indica que, de media, se devuelven 19 de los 20 vectores de documento verdaderos más cercanos a un vector de consulta
- 0.9 indica 9 de cada 10
- 0.8 indica 8 de cada 10

b. Para conocer los requisitos de latencia de búsqueda, selecciona tu tolerancia de latencia. Los requisitos modestos permiten ahorrar más costes gracias a los métodos de compresión que reducen los requisitos de memoria.

11. Elija Siguiente.
12. Revise la configuración y seleccione Crear.

El trabajo comienza a procesarse. Puede supervisar su progreso en la tabla de trabajos de ingestión de vectores.

Supervisar los trabajos de optimización

Puede supervisar el estado de sus trabajos de ingestión de vectores desde la página de destino de optimización automática.

Para supervisar los trabajos de optimización

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación, selecciona Optimización automática.
3. La tabla de trabajos de ingestión de vectores muestra todos los trabajos con su estado actual. Actualice la página para ver la información de estado actualizada.

> **Note**
>
> No hay ningún mecanismo automático de actualización o notificación. Debe actualizar manualmente la consola para ver cuándo se completa un trabajo.

Comprender los estados del estado del trabajo

Los trabajos de optimización automática pueden tener los siguientes valores de estado:

Pending (Pendiente)

El trabajo está en cola y esperando para empezar.

Ejecutar

El trabajo de optimización automática consiste en analizar activamente sus datos y generar recomendaciones.

Completado

El trabajo de optimización automática ha finalizado correctamente. Todos los análisis, las evaluaciones y las recomendaciones están completos y están disponibles para su visualización.

Con error

Se ha detectado un error en el trabajo. Consulte los detalles del error en la página de detalles del trabajo para determinar la causa.

Activo

Se ha creado un índice en el clúster adjunto y se han incorporado los datos.

La duración del trabajo depende principalmente del tamaño del conjunto de datos y de la carga de servicio actual. Los trabajos típicos se completan entre 15 minutos y varias horas.

Consultar detalles del trabajo

Puede ver información detallada sobre un trabajo de optimización específico, incluida su configuración y estado.

Para ver los detalles del trabajo

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación, selecciona Optimización automática.
3. En la tabla Trabajos de ingestión de vectores, elija el nombre del trabajo.
4. La página de detalles del trabajo muestra la siguiente información:

   - Nombre y estado del trabajo
   - Configuración de la fuente de datos (región y URI de Amazon S3)
   - OpenSearch dominio o colección
   - Configuración del índice (nombre de campo, tipo de espacio, dimensión)
   - Requisitos de rendimiento (recuperación y latencia)
   - Mensajes de error (si el trabajo ha fallado)

Visualización y comprensión de los resultados

Una vez que un trabajo se complete correctamente, podrá ver las recomendaciones de optimización.

Para ver los resultados de la optimización

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación, selecciona Optimización automática.
3. En la tabla Trabajos de ingestión de vectores, elija un trabajo con el estado Completado.
4. La página de resultados muestra las siguientes secciones:

   - Resumen de los resultados: muestra la calidad de búsqueda estimada en comparación con sus requisitos y el espacio de memoria del índice en comparación con la configuración más recomendada.
   - Recomendaciones: enumera hasta tres recomendaciones de optimización, ordenadas con la recomendación principal como la que mejor se adapte a su configuración. Cada recomendación incluye:
     - Parámetros de configuración del índice
     - Parámetros de configuración de búsqueda

- Métricas de rendimiento esperadas
- Estimaciones del espacio de memoria

> **Note**
>
> Si bien las recomendaciones se ordenan según la mejor opción, puedes seleccionar la que mejor se adapte a tu caso de uso específico. Optimice automáticamente los intentos para encontrar las coincidencias más parecidas a los criterios de recuperación que haya elegido.

Crear un índice a partir de las recomendaciones

Tras revisar las recomendaciones de optimización, puede crear un índice manualmente con la configuración recomendada o crear automáticamente un índice con la recomendación seleccionada.

Para crear un índice automáticamente

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación, selecciona Optimización automática.
3. En la tabla Trabajos de ingestión de vectores, elija un trabajo con el estado Completado.
4. Revise las recomendaciones y seleccione la que desee usar.
5. Elige Crear índice.
6. El sistema crea automáticamente un índice en el clúster según la recomendación seleccionada e ingiere los datos vectoriales del conjunto de datos.

Para crear un índice manualmente

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación, selecciona Optimización automática.
3. En la tabla Trabajos de ingestión de vectores, elija un trabajo con el estado Completado.
4. Revise las recomendaciones y anote la configuración del índice y los parámetros de configuración de búsqueda de la recomendación que elija.

5. Utilice la OpenSearch API o la consola para crear manualmente un índice con los parámetros recomendados.

Características relacionadas

La optimización automática funciona junto con otras funciones de Amazon OpenSearch Service para ayudarte a crear y optimizar aplicaciones de búsqueda vectorial:

- Aceleración de GPU para indexación vectorial- Acelere la creación de índices vectoriales mediante la aceleración de la GPU para reducir el tiempo y los costes de indexación.
- Ingestión de vectores- Incorpore e indexe rápidamente datos vectoriales de Amazon S3 en su dominio o colección.

Aceleración de GPU para indexación vectorial

La aceleración de la GPU le ayuda a crear bases de datos vectoriales a gran escala de forma más rápida y eficiente. Puede habilitar esta función en OpenSearch dominios y OpenSearch colecciones sin servidor nuevos o existentes. Esta función utiliza la aceleración de la GPU para reducir el tiempo necesario para indexar los datos en índices vectoriales.

Con la aceleración por GPU, puede aumentar la velocidad de indexación vectorial hasta 10 veces a una cuarta parte del costo de indexación.

Requisitos previos

La aceleración por GPU se admite en los OpenSearch dominios que ejecutan una versión o posterior y en las colecciones Serverless. OpenSearch `3.1` OpenSearch Para obtener más información, consulteActualización de los dominios OpenSearch de Amazon Service, y. UpdateDomainConfigUpdateCollection APIs

Funcionamiento

Los índices vectoriales requieren importantes recursos de cómputo para crear estructuras de datos, como los gráficos jerárquicos de mundos pequeños navegables (HNSW). Cuando habilita la aceleración por GPU en su dominio o colección, detecta OpenSearch automáticamente las oportunidades para acelerar la creación de índices y transfiere las compilaciones de índices a las instancias de GPU. OpenSearch El servicio administra las instancias de GPU en tu nombre y las

asigna a tu dominio o colección cuando es necesario. Esto significa que no administras la utilización ni pagas por el tiempo de inactividad.

Solo paga por el procesamiento útil mediante unidades de cómputo (OCU): aceleración vectorial. Cada OCU de aceleración vectorial es una combinación de aproximadamente 8 GiB de memoria de CPU, 2 CPUs V y 6 GiB de memoria de GPU. Para obtener más información, consulte the section called “Precios de aceleración de GPU”.

Para habilitar la aceleración por GPU en su dominio o colección, consulte. Habilitar la aceleración por GPU

Precios de aceleración de GPU

AWS le cobra cuando OpenSearch detecta oportunidades para acelerar las cargas de trabajo de creación de índices de su dominio o colección. Cada OCU de aceleración vectorial es una combinación de aproximadamente 8 GiB de memoria de CPU, 2 CPUs V y 6 GiB de memoria de GPU.

AWS factura a OCU con una granularidad de segundo nivel. En tu estado de cuenta, verás una entrada para calcular en OCU-hours.

Por ejemplo, si utilizas la aceleración de la GPU durante una hora para crear un índice con 2 vCPU y 1 GiB de memoria de GPU, se te facturará 1 OCU. Si utilizas 9 GiB de memoria de CPU mientras utilizas la aceleración de la GPU, se te facturarán 2 OCU.

OpenSearch Serverless añade más OCU OCUs en incrementos de 1 OCU en función de la potencia de cálculo y el almacenamiento necesarios para sus colecciones. Puede configurar un número máximo de OCUs para su cuenta a fin de controlar los costes.

> **Note**
>
> El número de OCUs aprovisionados en cualquier momento puede variar y no es exacto. Con el tiempo, el algoritmo que OpenSearch utiliza OpenSearch Serverless seguirá mejorando para minimizar mejor el uso del sistema.

Para obtener información completa sobre los precios, consulta los precios OpenSearch de Amazon Service.

Operaciones de escritura y aceleración de la GPU

La aceleración de la GPU se activa cuando la tasa de ingestión OpenSearch de vectores (MB/seg) está dentro de un rango. En OpenSearch los dominios, tiene la flexibilidad de configurar este rango mediante y. `index.knn.remote_index_build.size.min` `index.knn.remote_index_build.size.max` Por ejemplo, con el rango inferior predeterminado de 50 MB, escribir 15 000 vectores de precisión total con una dimensión de 768 entre los intervalos de actualización activará la aceleración de la GPU de forma predeterminada.

Los datos se escriben con las siguientes operaciones de API:

- Flush
- Bulk
- Reindexar
- Índice
- Actualización
- Delete
- Fusión forzada

La aceleración de la GPU se activa con la combinación de segmentos automática y manual.

Configuraciones de índice compatibles

El motor Faiss admite la aceleración de la GPU.

Las siguientes configuraciones no admiten la aceleración por GPU:

- Falsa la cuantificación del producto
- Índice de archivos invertido (IVF)
- Biblioteca de espacios no métricos
- Motor Lucene

Compatible Regiones de AWS

La aceleración por GPU está disponible en los siguientes casos: Regiones de AWS

- Este de EE. UU. (Norte de Virginia)
- Oeste de EE. UU. (Oregón)
- Asia-Pacífico (Sídney)
- Asia Pacífico (Tokio)
- Europa (Irlanda)

Habilitar la aceleración por GPU

Puedes habilitar la aceleración por GPU al crear o actualizar un OpenSearch dominio o una colección OpenSearch sin servidor con el, o el SDK. Consola de administración de AWS AWS CLI AWS

Una vez que habilites la aceleración por GPU en tu dominio o colección, esta función estará habilitada de forma predeterminada en todos los índices. Si necesita deshabilitar esta función a nivel de índice, consulte. Creación de índices vectoriales acelerados por GPU

Consola

Los siguientes procedimientos permiten la aceleración por GPU para OpenSearch dominios y colecciones OpenSearch sin servidor mediante la consola de administración sin servidor. OpenSearch

Create new domain

Para crear un OpenSearch dominio con la aceleración por GPU habilitada, consulte. Creación de dominios OpenSearch de servicio

Edit existing domain

1. Abra la consola de administración del OpenSearch servicio.
2. En el panel de navegación, elija Dominios.
3. Elige tu nombre de dominio para abrir la página de detalles del dominio.
4. Selecciona Acciones y, a continuación, Editar dominio.
5. En la sección Funciones avanzadas, selecciona Activar la aceleración de la GPU. Una vez habilitada esta función, las operaciones de indexación vectorial se aceleran.
6. Seleccione Save changes (Guardar cambios).

Create new collection

Para crear una colección OpenSearch sin servidor con la aceleración por GPU habilitada, consulte. Tutorial: Introducción a Amazon OpenSearch Serverless Durante la creación de la colección, asegúrese de seleccionar el tipo de colección de búsqueda vectorial y habilitar la aceleración por GPU en la configuración de búsqueda vectorial.

Edit existing collection

1. Abra la consola de administración del OpenSearch servicio.
2. En el panel de navegación, elija Colecciones.
3. Elija el nombre de la colección para abrir la página de detalles de la colección.
4. En la sección de opciones de despliegue, edita la aceleración vectorial de la GPU.
5. Activa o desactiva la aceleración de la GPU.
6. Seleccione Save changes (Guardar cambios).

AWS CLI

Create new domain

En el siguiente AWS CLI ejemplo, se crea un OpenSearch dominio con la aceleración por GPU habilitada en EE. UU. Este (Norte de Virginia). *text*Sustitúyalo por el de su propia configuración.

```
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

El siguiente AWS CLI ejemplo habilita la aceleración por GPU para un dominio existente. OpenSearch *text*Sustitúyalo por el de su propia configuración.

```
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

En el siguiente AWS CLI ejemplo, se crea una colección OpenSearch sin servidor con la aceleración por GPU habilitada en EE. UU. Este (Virginia del Norte). Sustitúyala por *text* la de su propia configuración.

```
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

El siguiente AWS CLI ejemplo habilita la aceleración por GPU para una colección Serverless existente OpenSearch . Sustitúyalo por *text* el de su propia configuración.

```
aws opensearchserverless update-collection \
    --id 07tjusf2h91cunochc \
    --vector-options '{"ServerlessVectorAcceleration": "ENABLED"}' \
    --region us-east-1
```

Creación de índices vectoriales acelerados por GPU

Después de habilitar la aceleración por GPU en su dominio o colección, cree índices vectoriales que puedan aprovechar el procesamiento de la GPU.

> **Note**
>
> Cuando creas un dominio con la aceleración por GPU habilitada, la configuración es la predeterminada. `index.knn.remote_index_build.enabled` `true` No es necesario establecer esta configuración de forma explícita al crear índices. En el caso de las colecciones, debes especificar explícitamente un valor para esta configuración.

Creating index with GPU-acceleration

El siguiente ejemplo crea un índice vectorial optimizado para el procesamiento de la GPU. Este índice almacena vectores de 768 dimensiones (habituales en las incrustaciones de texto).

```
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"
      }
    }
  }
```

```
}
```

Elementos clave de configuración:

- `"index.knn": true`- Habilita la funcionalidad k-near neighbor
- `"index.knn.remote_index_build.enabled": true`- Habilita el procesamiento por GPU para este índice. Cuando el dominio tiene habilitada la aceleración por GPU, el valor predeterminado es no especificar esta configuración. `true` En el caso de las colecciones, debes especificar explícitamente un valor para esta configuración.
- `"dimension": 768`- Especifica el tamaño del vector (se ajusta en función del modelo de incrustación)

Creating index without GPU-acceleration

El siguiente ejemplo crea un índice vectorial en el que el procesamiento de la GPU está desactivado. Este índice almacena vectores de 768 dimensiones (habituales en las incrustaciones de texto).

```
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"
      }
    }
  }
}
```

Indexación de datos vectoriales y fusión forzada

Una vez que haya creado un índice vectorial acelerado por GPU en su dominio o colección, puede añadir datos vectoriales y optimizar el índice mediante operaciones estándar. OpenSearch La aceleración por GPU mejora automáticamente tanto el rendimiento de indexación como las operaciones de fusión forzada, lo que agiliza la creación y el mantenimiento de aplicaciones de búsqueda vectorial a gran escala sin necesidad de cambiar los flujos de trabajo existentes.

Indexación de datos vectoriales

Indexe los datos vectoriales como lo haría normalmente. La aceleración de la GPU se aplica automáticamente a las operaciones de indexación y fusión forzada. En el siguiente ejemplo, se muestra cómo añadir documentos vectoriales al índice mediante la API masiva. Cada documento contiene un campo vectorial con valores numéricos y contenido de texto asociado:

```
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"}
```

Operaciones de fusión forzada

La aceleración de la GPU también se aplica a las operaciones de fusión forzada, lo que puede reducir considerablemente el tiempo necesario para optimizar los índices vectoriales. Ten en cuenta que las colecciones no admiten las operaciones de fusión forzada. El siguiente ejemplo muestra cómo optimizar el índice vectorial mediante la consolidación de todos los segmentos en un solo segmento:

```
POST my-vector-index/_forcemerge?max_num_segments=1
```

Prácticas recomendadas

Siga estas prácticas recomendadas para maximizar los beneficios de la aceleración de la GPU para sus cargas de trabajo de búsqueda vectorial:

- Aumente el número de clientes indexados: para aprovechar al máximo la GPUs creación del índice, aumente el número de clientes indexados en los que ingiere datos. OpenSearch Esto permite una mejor paralelización y utilización de los recursos de la GPU.

- Ajuste el umbral aproximado: cambie la `index.knn.advanced.approximate_threshold` configuración para garantizar que no se generen índices de segmentos más pequeños, lo que mejora la velocidad general de ingesta. Un valor de 10 000 es un buen punto de partida. En el caso de las colecciones, debes especificar explícitamente un valor para esta configuración.
- Optimice el tamaño de los fragmentos: intente crear fragmentos que tengan al menos 1 millón de documentos. Es posible que los fragmentos con menos de este número de documentos no se beneficien en general de la aceleración de la GPU.

Uso de OpenSearch cuadros de mando con Amazon Service OpenSearch

OpenSearch Dashboards es una herramienta de visualización de código abierto diseñada para funcionar con ella. OpenSearch Amazon OpenSearch Service proporciona una instalación de paneles de control con cada dominio OpenSearch de servicio. OpenSearch Dashboards se ejecuta en los nodos de datos más importantes del dominio.

OpenSearch Los paneles son una herramienta de visualización para explorar y analizar datos dentro de un único OpenSearch dominio. Por el contrario, la interfaz de OpenSearch usuario centralizada (también denominada OpenSearch aplicación) es una interfaz de usuario basada en la nube que se conecta a varios OpenSearch dominios, colecciones OpenSearch sin servidor y fuentes de AWS datos. Incluye espacios de trabajo para casos de uso específicos, como los análisis de observabilidad y seguridad y proporciona una experiencia unificada en todos los conjuntos de datos. Si bien Dashboards está vinculado a dominios individuales, la interfaz de usuario centralizada permite la integración y el análisis de datos entre dominios. Para obtener más información, consulte OpenSearch INTERFAZ DE USUARIO.

Encontrarás un enlace a los OpenSearch paneles en el panel de control de tu dominio, en la consola de OpenSearch servicio. Para los dominios en ejecución OpenSearch, la URL es*domain-endpoint*`/_dashboards/`. Para los dominios que ejecutan una versión heredada de Elasticsearch, la URL es *domain-endpoint*`/_plugin/kibana`.

Las consultas que utilizan esta instalación de OpenSearch Dashboards predeterminada tienen un tiempo de espera de 300 segundos.

> **Note**
>
> En esta documentación se analizan los OpenSearch paneles de control en el contexto de Amazon OpenSearch Service, incluidas las diferentes formas de conectarse a ellos. Para obtener una documentación completa, que incluye una guía de introducción, instrucciones para crear un panel, administrar los paneles y el lenguaje de consulta de paneles (DQL), consulte los OpenSearch paneles en la documentación de código abierto. OpenSearch

Control del acceso a Dashboards

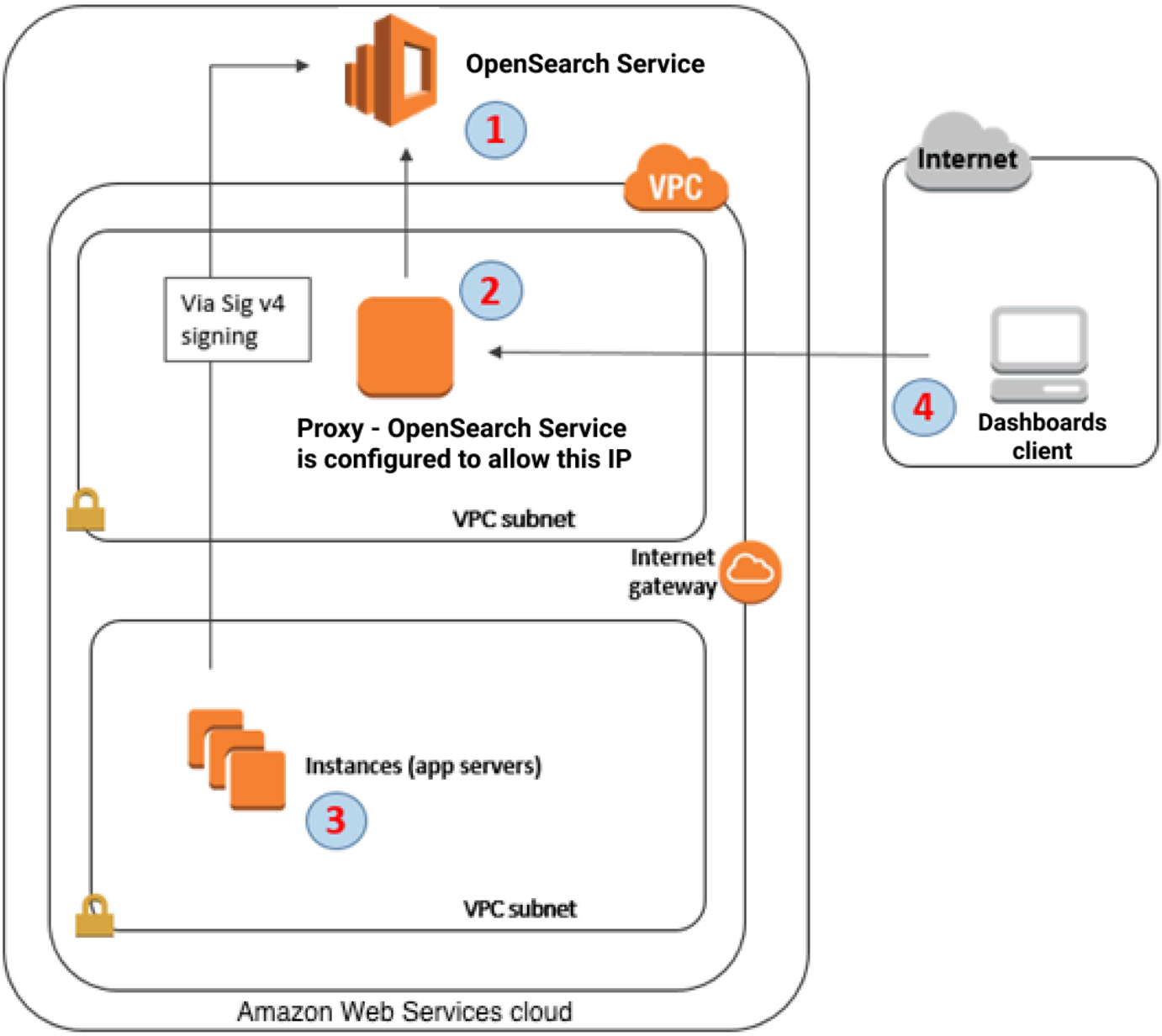
Dashboards no admite de forma nativa los usuarios y roles de IAM, pero OpenSearch Service ofrece varias soluciones para controlar el acceso a los paneles:

- Habilite SAML authentication for Dashboards (Autenticación SAML para Dashboards).
- Utilice el control de acceso detallado con autenticación básica de HTTP.
- Configure la autenticación de Cognito para Dashboards.
- Para los dominios de acceso público, configure una política de acceso basada en IP que utilice o no un servidor proxy.
- Para los dominios de acceso a la VPC, configure una política de acceso abierto que utilice o no un servidor proxy, y grupos de seguridad para controlar el acceso. Para obtener más información, consulte the section called “Acerca de las políticas de acceso en los dominios de VPC”.

Uso de un proxy para acceder al servicio desde los paneles OpenSearch

> **Note**
>
> Este proceso solo es de aplicación si el dominio utiliza acceso público y no se desea emplear la autenticación de Cognito. Consulte the section called “Control del acceso a Dashboards”.

Como Dashboards es una JavaScript aplicación, las solicitudes se originan en la dirección IP del usuario. El control de acceso basado en IP podría resultar poco práctico debido al gran número de direcciones IP a las que sería necesario brindar acceso para que cada usuario tenga acceso a Dashboards. Una solución alternativa consiste en colocar un servidor proxy entre los paneles y el servicio. OpenSearch A continuación, puede agregar una política de acceso basada en IP que permita solicitudes desde solo una única dirección IP, la del proxy. En el siguiente diagrama se muestra esta configuración.



1. Este es su dominio de OpenSearch servicio. IAM proporciona acceso autorizado a este dominio. Otra política de acceso basada en IP proporciona acceso al servidor proxy.
2. Este es el servidor proxy que se ejecuta en una instancia de Amazon EC2.
3. Otras aplicaciones pueden usar el proceso de firma de la versión 4 de Signature para enviar solicitudes autenticadas al OpenSearch Servicio.
4. Los clientes de Dashboards se conectan a su dominio OpenSearch de servicio a través del proxy.

Para habilitar esta clase de configuración, necesita una política basada en recursos que especifica funciones y direcciones IP. A continuación, mostramos un ejemplo de política:

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/*"
      }
   ]
}
```

Recomendamos configurar la instancia EC2 mediante la ejecución del el servidor proxy con una dirección IP elástica. De esta forma, puede sustituir la instancia cuando sea necesario y continuar adjuntando la misma dirección IP pública a la misma. Para obtener más información, consulte Direcciones IP elásticas en la Guía del usuario de Amazon EC2.

Si utiliza un servidor proxy y la autenticación de Cognito, es posible que tenga que agregar parámetros de configuración para Dashboards y Amazon Cognito con el fin de evitar errores de tipo `redirect_mismatch`. Consulte el siguiente ejemplo de `nginx.conf`:

```
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**
>
> (Opcional) Si decide aprovisionar un nodo coordinador dedicado, este empezará a alojar automáticamente los OpenSearch paneles de control. Como resultado, aumenta la disponibilidad de los recursos del nodo de datos, como la CPU y la memoria. Esta mayor disponibilidad de los recursos de los nodos de datos puede ayudar a mejorar la resiliencia general de su dominio.

Configuración de OpenSearch Dashboards para utilizar un servidor de mapas de WMS

La instalación predeterminada de Dashboards for OpenSearch Service incluye un servicio de mapas, excepto para los dominios de las regiones de India y China. El servicio de mapas admite hasta 10 niveles de acercamiento/alejamiento.

Con independencia de la región, puede configurar Dashboards para que utilice un servidor de Web Map Service (WMS) diferente para las visualizaciones de mapas de coordenadas. Las visualizaciones de mapas de región solo admiten el servicio de mapas predeterminado.

Para configurar Dashboards para que utilice un servidor de mapas de WMS:

1. Abra Dashboards.
2. Seleccione Stack Management (Gestión de pilas).
3. Seleccione Advanced Settings (Configuración avanzada).
4. Visualización de ubicaciones: mapa de mosaicos:. WMSdefaults
5. Cambie `enabled` a `true` y `url` a la URL de un servidor de mapas WMS válido.

```
{
  "enabled": true,
  "url": "wms-server-url",
  "options": {
    "format": "image/png",
    "transparent": true
```

```
  }
}
```

6. Seleccione Save changes (Guardar cambios).

Para aplicar el nuevo valor predeterminado a las visualizaciones, es posible que deba volver a cargar Dashboards. Si guardó visualizaciones, elija Options (Opciones) después de abrir la visualización. Verifique que el Servidor de mapas WMS esté habilitado y que la URL de WMS contenga el servidor de mapas preferido y, a continuación, elija Apply changes (Aplicar cambios).

> **Note**
>
> Los servicios de mapa suelen estar sujetos a restricciones o cuotas de licencias. El usuario es responsable de todas estas consideraciones en cualquier servidor de mapas que especifique. Los servicios de mapa de U.S. Geological Survey podrían resultarle útiles para hacer pruebas.

Conexión de un servidor de paneles de control local al servicio OpenSearch

Si ya ha invertido una cantidad considerable de tiempo en configurar su propia instancia de Dashboards, puede utilizarla en lugar de (o además de) la instancia de Dashboards predeterminada que OpenSearch proporciona Service. El siguiente procedimiento sirve para dominios que utilicen el control de acceso detallado con una política de acceso abierto.

Para conectar un servidor de Dashboards local a Service OpenSearch

1. En su dominio OpenSearch de servicio, cree un usuario con los permisos adecuados:
   a. En Dashboards, vaya a Security (Seguridad), Internal users (Usuarios internos) y elija Create internal user (Crear usuario interno).
   b. Proporcione un nombre de usuario y una contraseña y elija Create (Crear).
   c. Vaya a Roles (Roles) y elija un rol.
   d. Seleccione Mapped users (Usuarios asignados) y Seleccione Manage mapping (Gestión de mapas).
   e. En Users (Usuarios), agregue el nombre de usuario y elija Map (Asignar).

2. Descargue e instale la versión adecuada del complemento de OpenSearch seguridad en su instalación autogestionada de Dashboards OSS.
3. En su servidor de Dashboards local, abra el `config/opensearch_dashboards.yml` archivo y añada su terminal de OpenSearch servicio con el nombre de usuario y la contraseña que creó anteriormente:

```
opensearch.hosts: ['https://domain-endpoint']
opensearch.username: 'username'
opensearch.password: 'password'
```

Puede utilizar el siguiente ejemplo de archivo `opensearch_dashboards.yml`:

```
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'

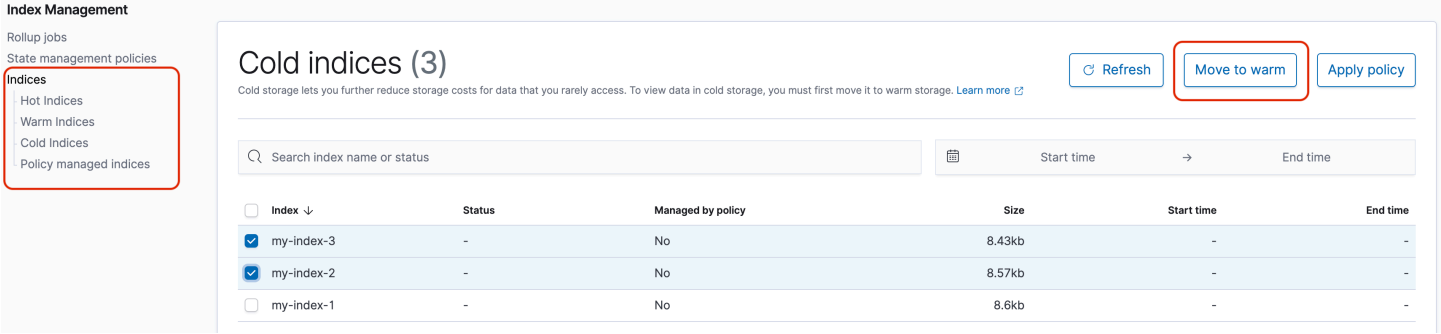
opensearch.username: 'username'
opensearch.password: 'password'
opensearch.requestHeadersWhitelist: [authorization, securitytenant,
 security_tenant]
```

Para ver sus índices OpenSearch de servicio, inicie su servidor de Dashboards local, vaya a Dev Tools y ejecute el siguiente comando:

```
GET _cat/indices
```

Administración de índices en OpenSearch Dashboards

La instalación de Dashboards en su dominio de OpenSearch servicio proporciona una interfaz de usuario útil para administrar los índices en los diferentes niveles de almacenamiento de su dominio. Seleccione Administración de índices en el menú principal de los paneles para ver todos los índices almacenados en caliente y en frío UltraWarm, así como los índices gestionados por las políticas de Index State Management (ISM). Utilice la administración de índices para mover índices entre el almacenamiento en frío y el almacenamiento en caliente y para monitorear las migraciones entre los tres niveles.



Tenga en cuenta que no verá las opciones de índice de calor, calor y frío a menos que tenga activado el almacenamiento en frío UltraWarm o esté activado.

Características adicionales

La instalación predeterminada de Dashboards en cada dominio OpenSearch de servicio tiene algunas funciones adicionales:

- Interfaces de usuario para los distintos complementos OpenSearch
- Inquilinos
- Informes

  Utilice Reporting (Informes) para generar informes CSV bajo demanda desde la página Discover e informes PDF o PNG de paneles o visualizaciones. Los informes CSV tienen un límite de 10 000 filas.
- Diagramas de Gantt
- Cuadernos

Uso de la OpenSearch interfaz de usuario en Amazon OpenSearch Service

OpenSearch La interfaz de usuario (UI) es una experiencia de análisis operativo modernizada para Amazon OpenSearch Service que proporciona una vista unificada para que puedas interactuar con los datos de múltiples fuentes. A diferencia de OpenSearch Dashboards, que solo funciona con el dominio o colección que lo aloja, la OpenSearch interfaz de usuario está alojada en. Nube de AWS Esto permite que la OpenSearch interfaz de usuario alcance una alta disponibilidad y siga funcionando durante las actualizaciones del clúster, además de conectarse de forma nativa con varias fuentes de datos. Para obtener información sobre los OpenSearch paneles de control, consulte. OpenSearch Cuadros de mando

Las siguientes son las principales características de la OpenSearch interfaz de usuario:

- Soporte para múltiples fuentes de datos: la OpenSearch interfaz de usuario puede conectarse con múltiples fuentes de datos para crear una vista completa. Esto incluye OpenSearch dominios y colecciones sin servidor, así como fuentes de AWS datos integradas como Amazon CloudWatch, Amazon Security Lake y Amazon Simple Storage Service (Amazon S3).
- Sin tiempo de inactividad durante las actualizaciones: la OpenSearch interfaz de usuario está alojada en. Nube de AWS Esto significa que OpenSearch permanece operativo y puede recuperar datos de los clústeres durante los procesos de actualización.
- Espacios de trabajo: espacios seleccionados para la colaboración en equipo para diversos flujos de trabajo, como observabilidad, análisis de seguridad y búsqueda. Puede definir la configuración de privacidad y administrar los permisos de los colaboradores en su espacio de trabajo.
- Inicio de sesión único: la OpenSearch interfaz de usuario funciona con una federación de SAML a través de AWS Identity and Access Management IAM para integrarse con sus proveedores de identidad AWS IAM Identity Center y crear una experiencia de inicio de sesión único para sus usuarios finales.
- Análisis basado en Genai: la OpenSearch interfaz de usuario admite la generación de consultas en lenguaje natural para ayudar a generar las consultas adecuadas para su análisis. OpenSearch La interfaz de usuario también trabaja con Amazon Q Developer para proporcionar el chat de Amazon Q y ayudar a generar visualizaciones, resúmenes de alertas, información y detectores de anomalías recomendados.

- Compatibilidad con varios lenguajes de consulta: la OpenSearch interfaz de usuario es compatible con el lenguaje de procesamiento canalizado (PPL), SQL, Lucene y el lenguaje de consultas de paneles (DQL).
- Soporte multiregional y multicuenta: la OpenSearch interfaz de usuario puede utilizar la función de búsqueda entre clústeres para conectarse con OpenSearch dominios de diferentes cuentas y regiones para realizar análisis y visualizaciones agregados.

Para empezar y crear tu primera OpenSearch interfaz de usuario, sigue las instrucciones que se indican en. the section called “Introducción”

Para obtener información sobre las últimas funciones publicadas para la OpenSearch interfaz de usuario, consultethe section called “Historial de versiones”.

Temas

- Historial de versiones de la interfaz de usuario de Amazon OpenSearch Service
- Introducción a la interfaz OpenSearch de usuario de Amazon OpenSearch Service
- La IA de la agencia en Amazon Service OpenSearch
- Cifrar los metadatos de las aplicaciones de OpenSearch interfaz de usuario con claves administradas por el cliente
- Habilitar la federación de SAML con AWS Identity and Access Management
- Administración de las asociaciones de orígenes de datos y los permisos de acceso a la nube privada virtual
- Uso de los espacios de trabajo OpenSearch de Amazon Service
- Acceso a datos entre regiones y cuentas
- Administrar el acceso a la OpenSearch interfaz de usuario desde un punto final de VPC
- Migración de objetos guardados de los OpenSearch paneles a la interfaz de usuario OpenSearch
- OpenSearch Puntos finales y cuotas de la interfaz de usuario

Historial de versiones de la interfaz de usuario de Amazon OpenSearch Service

En la siguiente tabla se enumeran todas las versiones del soporte de Amazon OpenSearch Service para la OpenSearch interfaz de usuario y las funciones y mejoras incluidas en cada versión.

Cambio	Fecha de publicación	Description (Descripción)
Nueva característica	10/03/2020	OpenSearch La interfaz de usuario ahora admite el acceso a los datos de varios dominios entre cuentas. OpenSearch Puede configurar las aplicaciones de una cuenta para acceder a los dominios de diferentes cuentas mediante IAM o AWS IAM Identity Center autenticación, tanto para dominios públicos como de VPC. Para obtener más información, consulte the section called “Acceso a los datos de varias cuentas a los OpenSearch dominios”.
Nueva característica	29 de diciembre de 2025	OpenSearch La interfaz de usuario ahora admite claves administradas por el cliente (CMK) para cifrar los metadatos de las aplicaciones. Esta función le ayuda a cumplir los requisitos normativos y de conformidad al proporcionarle un control total sobre las claves de cifrado. Este lanzamiento también aumenta el límite de tamaño de los metadatos de los objetos guardados en la OpenSearch interfaz de usuario, lo que permite crear y almacenar consultas complejas, visualizaciones extensas y espacios de trabajo a gran escala. Para obtener más información, consulte the section called “Cifrado y clave gestionada por el cliente”.
Nueva característica	16 de abril de 2025	OpenSearch La interfaz de usuario ahora funciona con la búsqueda entre clústeres. Esto te permite usar la OpenSearch interfaz de usuario en uno para acceder Región de AWS a los clústeres de una región diferente. Esto se hace configurándolo como un clúster remoto conectado a un clúster de la misma región. Para obtener más información, consulte the section called “Búsqueda en clústeres”.
Nueva característica	31-03-2021	Amazon Q Developer ya está disponible de forma general en Amazon OpenSearch Service. Para obtener información, consulte Asistente de IA.

Cambio	Fecha de publicación	Description (Descripción)
Nueva característica	05-02-05	La federación del lenguaje de marcado de aserciones de seguridad 2.0 (SAML) mediante la federación AWS Identity and Access Management (IAM) ahora funciona con la interfaz de usuario. OpenSearch Esto permite crear una experiencia de inicio de sesión único (SSO) iniciada por un proveedor de identidad para los usuarios finales. Para obtener más información, consulte the section called “Habilitar la federación SAML con IAM”.
Nueva integración	01 de diciembre de 2022	La integración de Zero-ETL con Amazon CloudWatch simplifica el análisis y la visualización de los datos de registro, lo que reduce los gastos técnicos y operativos. Para obtener más información, consulte New Amazon CloudWatch y Amazon OpenSearch Service lanzan una experiencia de análisis integrada en el blog de AWS noticias.
Nueva integración	01/12/2020	La integración sin ETL con Amazon Security Lake permite a las organizaciones buscar, analizar y obtener información procesable a partir de sus datos de seguridad de manera eficiente. Para obtener más información, consulte Introducción a la integración de Amazon OpenSearch Service y Amazon Security Lake para simplificar el análisis de seguridad en el blog de AWS noticias.
Versión inicial	07-11-2020	La versión pública inicial de UI. OpenSearch Para obtener más información, consulta Amazon OpenSearch Service lanza la OpenSearch interfaz de usuario de próxima generación en el blog sobre AWS macrodatos.

Introducción a la interfaz OpenSearch de usuario de Amazon OpenSearch Service

En Amazon OpenSearch Service, una aplicación es una instancia de la interfaz de OpenSearch usuario (OpenSearch UI). Cada aplicación se puede asociar a varios orígenes de datos y un solo origen se puede asociar a varias aplicaciones. Puede crear varias aplicaciones para distintos administradores mediante diferentes opciones de autenticación compatibles.

Utilice la información de este tema como guía en el proceso de creación de una aplicación de OpenSearch interfaz de usuario mediante el Consola de administración de AWS o el AWS CLI.

Temas

- Permisos necesarios para crear aplicaciones OpenSearch de Amazon Service
- Crear una aplicación de interfaz de usuario OpenSearch
- Administrar los administradores de aplicaciones

Permisos necesarios para crear aplicaciones OpenSearch de Amazon Service

Antes de crear una aplicación, compruebe que se le han concedido los permisos necesarios para la tarea. Si es necesario, póngase en contacto con un administrador de cuentas para obtener ayuda.

Permisos generales

Para trabajar con las aplicaciones de OpenSearch Service, necesitas los permisos que se muestran en la siguiente política. Los permisos sirven para los siguientes propósitos:

- Los cinco permisos `es:*Application` son necesarios para crear y administrar una aplicación.
- Los tres permisos `es:*Tags` son necesarios para agregar, enumerar y eliminar etiquetas de la aplicación.
- Los permisos `aoss:BatchGetCollection`, `es:DescribeDomain` y `es:GetDirectQueryDataSource` son necesarios para asociar los orígenes de datos.
- Los permisos `aoss:APIAccessAll`, `es:ESHttp*` y 4 `opensearch:*DirectQuery*` son necesarios para acceder a los orígenes de datos.
- `iam:CreateServiceLinkedRole`Da permiso a Amazon OpenSearch Service para crear un rol vinculado a un servicio (SLR) en tu cuenta. Esta función se utiliza y permite que la aplicación

de OpenSearch interfaz de usuario publique las CloudWatch métricas de Amazon en tu cuenta. Para obtener más información, consulte the section called “Permisos” en el tema Uso de roles vinculados al servicio para crear dominios de VPC y orígenes de datos de consulta directa.

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"
        }
    ]
}
```

Permisos para crear una aplicación que utilice la autenticación del Centro de identidades de IAM (opcional)

De forma predeterminada, las aplicaciones del panel de control se autentican mediante AWS Identity and Access Management (IAM) para gestionar los permisos de los usuarios de los AWS recursos. Sin embargo, puede optar por ofrecer una experiencia de inicio de sesión único mediante el Centro de identidad de IAM, que le permite utilizar sus proveedores de identidad actuales para iniciar sesión en las aplicaciones de interfaz de usuario. OpenSearch En este caso, seleccionará la opción Autenticación con el Centro de Identidad de IAM en el procedimiento que aparece más adelante en este tema y, a continuación, concederá a los usuarios del Centro de Identidad de IAM los permisos necesarios para acceder a la aplicación de interfaz de usuario.) OpenSearch

Para crear una aplicación que utilice la autenticación del Centro de identidades de IAM, necesitará los siguientes permisos. Sustituya *`placeholder values`* por su propia información. Si es necesario, póngase en contacto con un administrador de cuentas para obtener ayuda.

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"
        }
    ]
}
```

Crear una aplicación de interfaz de usuario OpenSearch

Cree una aplicación que especifique el nombre de la aplicación, el método de autenticación y los administradores mediante uno de los siguientes procedimientos.

Temas

- Crear una aplicación de OpenSearch interfaz de usuario que utilice la autenticación de IAM en la consola
- Crear una aplicación de OpenSearch interfaz de usuario que utilice la AWS IAM Identity Center autenticación en la consola
- Crear una aplicación de OpenSearch interfaz de usuario que utilice la AWS IAM Identity Center autenticación mediante el AWS CLI

Crear una aplicación de OpenSearch interfaz de usuario que utilice la autenticación de IAM en la consola

Para crear una aplicación de OpenSearch interfaz de usuario que utilice la autenticación de IAM en la consola

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, selecciona OpenSearch UI (paneles).
3. Elija Creación de aplicación.
4. En Nombre de la aplicación, especifique un nombre para su aplicación.
5. No active la casilla de verificación Autenticación con el Centro de identidades de IAM. Para obtener información sobre cómo crear una aplicación con autenticación mediante autenticación AWS IAM Identity Center, consulte the section called “Crear una aplicación de OpenSearch interfaz de usuario que utilice la AWS IAM Identity Center autenticación en la consola” más adelante en este tema.
6. (Opcional) Se lo agrega automáticamente como administrador de la aplicación que está creando. En el área de administración de administradores de OpenSearch aplicaciones, puede conceder permisos de administrador a otros usuarios.

> **Note**
>
> El rol de administrador de la aplicación de OpenSearch interfaz de usuario otorga permisos para editar y eliminar una aplicación de OpenSearch interfaz de usuario. Los administradores de aplicaciones también pueden crear, editar y eliminar espacios de trabajo en una aplicación de OpenSearch interfaz de usuario.

Para conceder permisos de administrador a otros usuarios, elija una de las siguientes opciones:

- Otorgue el permiso de administrador a usuarios específicos: en el campo de administradores de la OpenSearch aplicación, en la lista emergente Propiedades, seleccione usuarios de IAM o

  AWS IAM Identity Center usuarios y, a continuación, elija los usuarios individuales a los que va a conceder los permisos de administrador.
- Otorgue permisos de administrador a todos los usuarios: todos los usuarios de su organización o cuenta tienen permisos de administrador.

7. (Opcional) Configure los ajustes de cifrado. De forma predeterminada, los metadatos de la OpenSearch interfaz de usuario se cifran con las claves AWS propias. Para utilizar su propia clave gestionada por el cliente (CMK) para el cifrado, consultethe section called “Cifrado y clave gestionada por el cliente”.
8. (Opcional) En el área Etiquetas, aplique uno o más name/value pares de claves de etiquetas a la aplicación.

   Las etiquetas son metadatos opcionales que usted asigna a un recurso. Las etiquetas permiten clasificar los recursos de diversas maneras, por ejemplo, según la finalidad, el propietario o el entorno.
9. Seleccione Crear.

Crear una aplicación de OpenSearch interfaz de usuario que utilice la AWS IAM Identity Center autenticación en la consola

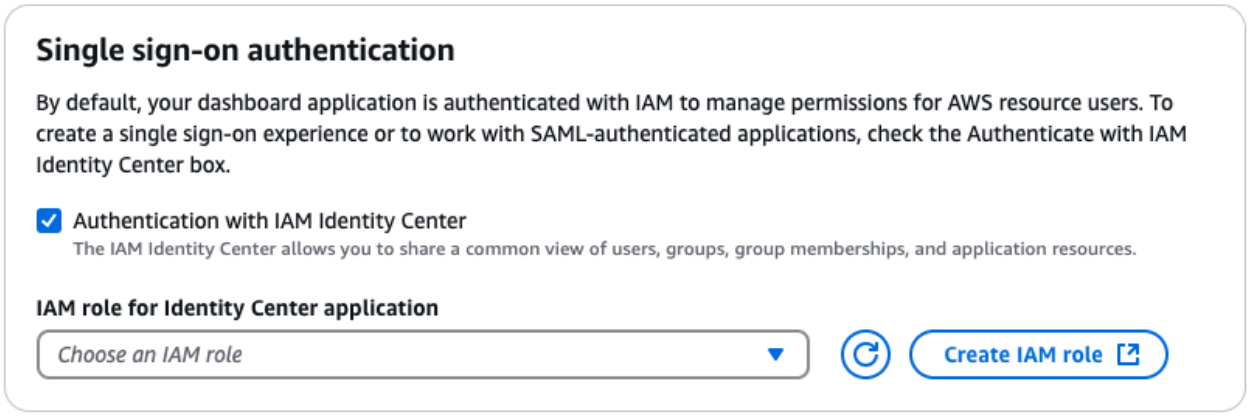
Para crear una aplicación de OpenSearch interfaz de usuario que utilice la AWS IAM Identity Center autenticación, debe tener los permisos de IAM descritos anteriormente en este tema enthe section

called “Permisos para crear una aplicación que utilice la autenticación del Centro de identidades de IAM (opcional)”.

Para crear una aplicación de OpenSearch interfaz de usuario que utilice la AWS IAM Identity Center autenticación en la consola

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, selecciona OpenSearch UI (paneles).
3. Elija Creación de aplicación.
4. En Nombre de la aplicación, especifique un nombre para su aplicación.
5. (Opcional) Para habilitar el inicio de sesión único en su organización o cuenta, haga lo siguiente:

   a. Seleccione la casilla Autenticación con el Centro de identidades de IAM, como se muestra en la siguiente imagen:



   b. Realice una de las siguientes acciones:

   - En la lista de aplicaciones de IAM para Identity Center, elija una función de IAM existente que proporcione los permisos necesarios para que IAM Identity Center acceda a la OpenSearch interfaz de usuario y a las fuentes de datos asociadas. Consulte las políticas en el siguiente punto para ver los permisos que debe tener el rol.
   - Creación de un rol de IAM con los permisos necesarios. Utilice los siguientes procedimientos de la Guía del usuario de IAM con las opciones especificadas para crear un nuevo rol y con la política de permisos y la política de confianza necesarias.
     - Procedimiento: Creación de políticas de IAM (consola)

Siguiendo los pasos de este procedimiento, pegue la siguiente política en el campo JSON del editor de políticas:

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": "*"
        }
    ]
}
```

- Procedimiento: Cree un rol mediante políticas de confianza personalizadas

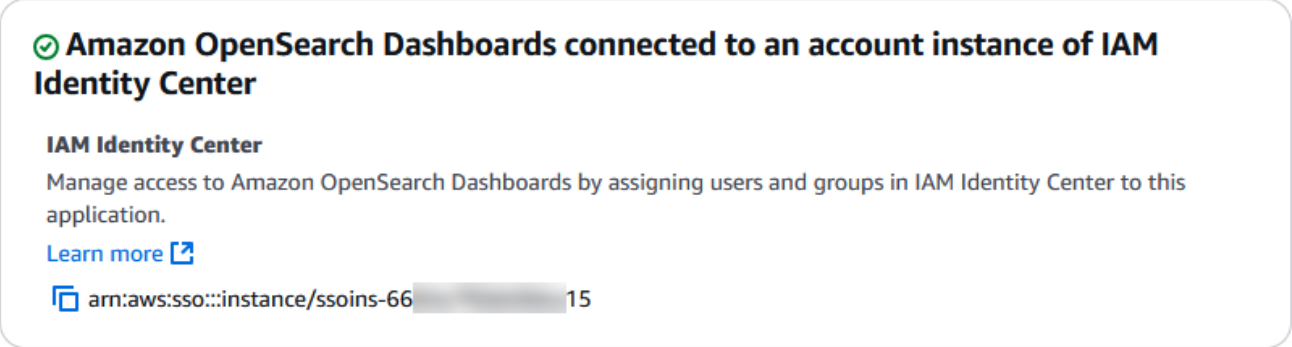
Siguiendo los pasos de este procedimiento, reemplace el marcador de posición JSON del cuadro Política de confianza personalizada con lo siguiente:

> **Tip**
>
> Si va a agregar la política de confianza a un rol existente, agregue la política en la pestaña Relación de confianza del rol.

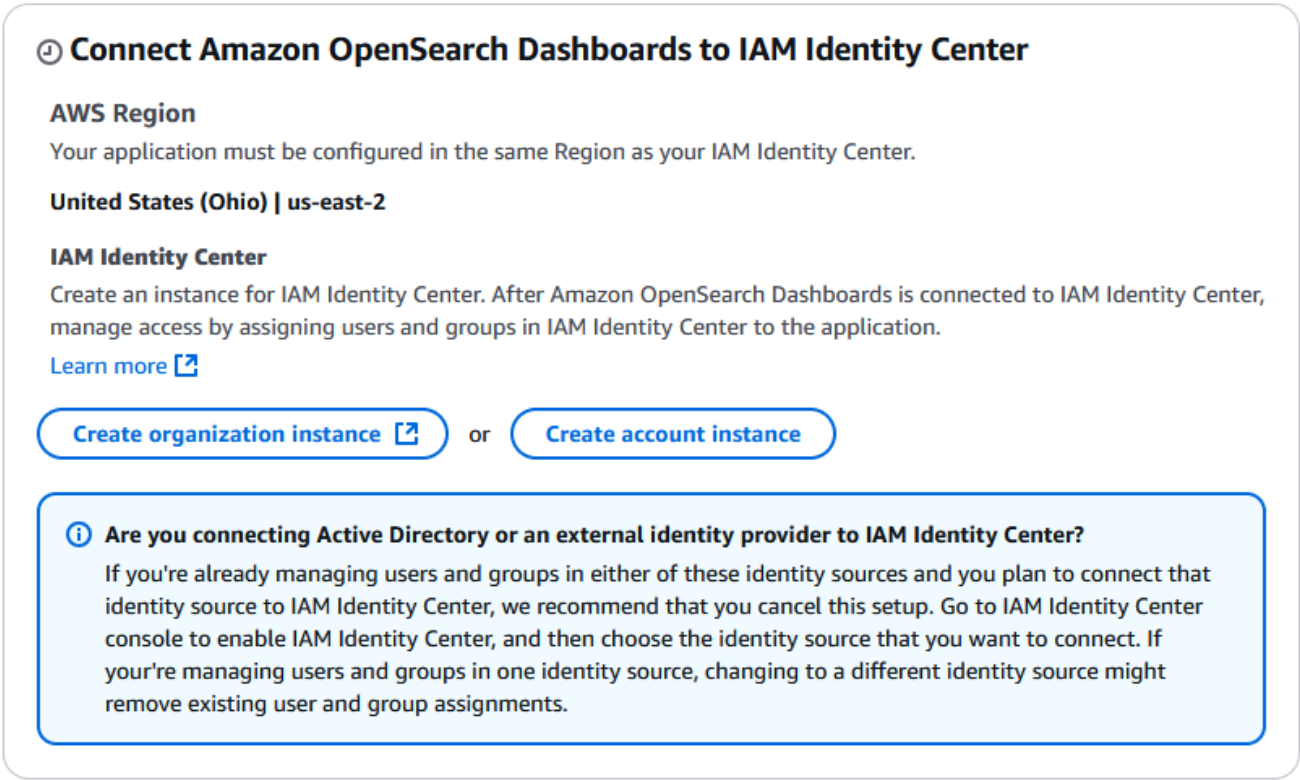
JSON

```
{
    "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. Si ya se ha creado una instancia de IAM Identity Center en su organización o cuenta, la consola informa de que Amazon OpenSearch Dashboards ya está conectado a una instancia de organización de IAM Identity Center, como se muestra en la siguiente imagen.



Si el Centro de Identidad de IAM aún no está disponible en su organización o cuenta, usted o un administrador con los permisos necesarios pueden crear una instancia de organización o de cuenta. El área Connect Amazon OpenSearch Dashboards to IAM Identity Center ofrece opciones para ambos, como se muestra en la siguiente imagen:



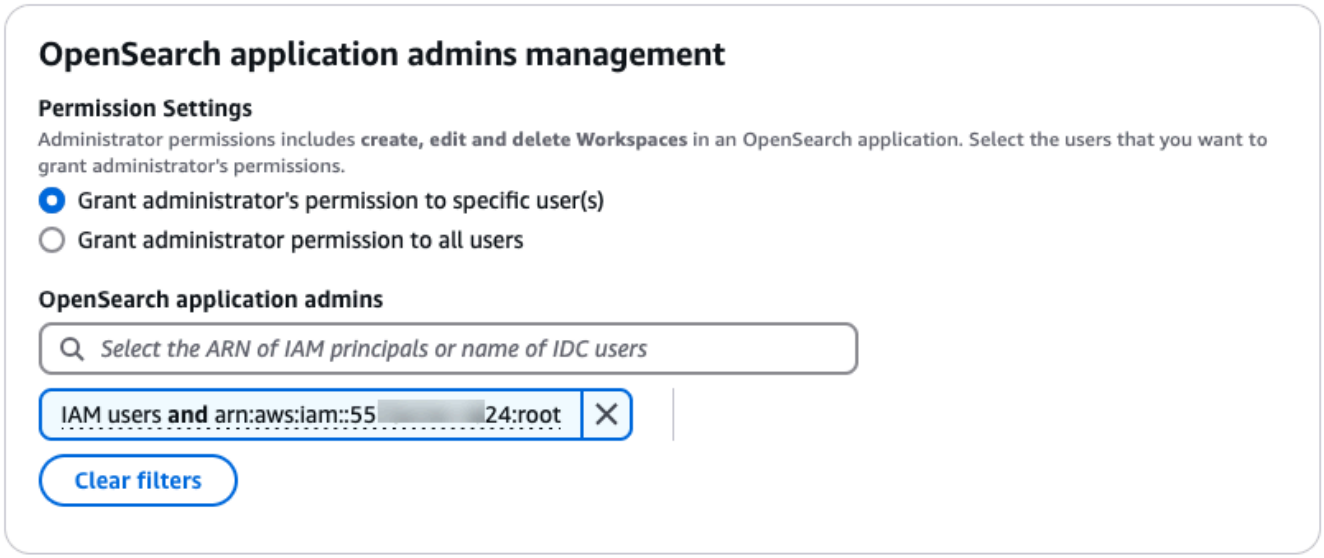
En este caso, puede crear una instancia de cuenta en el Centro de Identidad de IAM para realizar pruebas o solicitar que un administrador cree una instancia de organización en el

Centro de Identidad de IAM. Para obtener más información, consulte los siguientes temas en la Guía del usuario de AWS IAM Identity Center :

> **Note**
>
> Si desea crear aplicaciones de OpenSearch interfaz de usuario en una instancia organizativa diferente a la Región de AWS de su IAM Identity Center, consulte Uso del IAM Identity Center en múltiples instancias. Regiones de AWS

- Instancias de organización del IAM Identity Center
- Instancias de cuenta de IAM Identity Center
- Habilitado AWS IAM Identity Center

6. (Opcional) Se lo agrega automáticamente como administrador de la aplicación que está creando. En el área de administración de administradores de OpenSearch aplicaciones, puede conceder permisos de administrador a otros usuarios, como se muestra en la siguiente imagen:



> **Note**
>
> El rol de administrador de aplicaciones de OpenSearch interfaz de usuario otorga permisos para editar y eliminar una aplicación de OpenSearch interfaz de usuario. Los administradores de aplicaciones también pueden crear, editar y eliminar espacios de trabajo en una aplicación de OpenSearch interfaz de usuario.

Para conceder permisos de administrador a otros usuarios, elija una de las siguientes opciones:

- Otorgue el permiso de administrador a usuarios específicos: en el campo de administradores de la OpenSearch aplicación, en la lista emergente Propiedades, seleccione usuarios de IAM o

  AWS IAM Identity Center usuarios y, a continuación, elija los usuarios individuales a los que va a conceder los permisos de administrador.
- Otorgue permisos de administrador a todos los usuarios: todos los usuarios de su organización o cuenta tienen permisos de administrador.

7. (Opcional) En el área Etiquetas, aplique uno o más name/value pares de claves de etiquetas a la aplicación.

   Las etiquetas son metadatos opcionales que usted asigna a un recurso. Las etiquetas permiten clasificar los recursos de diversas maneras, por ejemplo, según la finalidad, el propietario o el entorno.
8. Seleccione Crear.

Crear una aplicación de OpenSearch interfaz de usuario que utilice la AWS IAM Identity Center autenticación mediante el AWS CLI

Para crear una aplicación de OpenSearch interfaz de usuario que utilice la AWS IAM Identity Center autenticación mediante el AWS CLI, utilice el comando create-application con las siguientes opciones:

- `--name`: nombre de la aplicación.
- `--iam-identity-center-options`— (Opcional) La instancia del IAM Identity Center y el rol de IAM que OpenSearch se utilizará para la autenticación y el control de acceso.

Sustituya *`placeholder values`* por su propia información.

```
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\"
          }
    "
```

Administrar los administradores de aplicaciones

Un administrador de aplicaciones de OpenSearch interfaz de usuario es un rol definido con permiso para editar y eliminar una aplicación de OpenSearch interfaz de usuario.

De forma predeterminada, como creador de una aplicación de OpenSearch interfaz de usuario, usted es el primer administrador de la aplicación de OpenSearch interfaz de usuario.

Administrar los administradores de la OpenSearch interfaz de usuario mediante la consola

Puede añadir administradores adicionales a una aplicación de OpenSearch interfaz de usuario en la Consola de administración de AWS, ya sea durante el flujo de trabajo de creación de la aplicación o en la página de edición una vez creada la aplicación.

La función de administrador de aplicaciones de OpenSearch interfaz de usuario otorga permisos para editar y eliminar una aplicación de OpenSearch interfaz de usuario. Los administradores de aplicaciones también pueden crear, editar y eliminar espacios de trabajo en una aplicación de OpenSearch interfaz de usuario.

En la página de detalles de la aplicación, puede buscar el nombre de recurso de Amazon (ARN) de una entidad principal de IAM o el nombre de un usuario del Centro de Identidad de IAM.

Para gestionar los administradores de la OpenSearch interfaz de usuario mediante la consola

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, selecciona OpenSearch UI (paneles).
3. En el área de OpenSearch aplicaciones, elija el nombre de una aplicación existente.
4. Elija Editar
5. Para conceder permisos de administrador a otros usuarios, elija una de las siguientes opciones:

- Otorgue el permiso de administrador a usuarios específicos: en el campo de administradores de la OpenSearch aplicación, en la lista emergente Propiedades, seleccione usuarios de IAM o

  AWS IAM Identity Center usuarios y, a continuación, elija los usuarios individuales a los que va a conceder los permisos de administrador.
- Otorgue permisos de administrador a todos los usuarios: todos los usuarios de su organización o cuenta tienen permisos de administrador.

6. Elija Actualizar.

Puede eliminar administradores adicionales, pero cada aplicación de OpenSearch interfaz de usuario debe tener al menos un administrador.

Administrar los administradores OpenSearch de la interfaz de usuario mediante el AWS CLI

Puede crear y actualizar los administradores de aplicaciones de OpenSearch interfaz de usuario mediante el AWS CLI.

Crear administradores OpenSearch de interfaz de usuario mediante el AWS CLI

Los siguientes son ejemplos de cómo agregar directores de IAM y usuarios del Centro de Identidad de IAM como administradores al crear una OpenSearch aplicación de interfaz de usuario.

Ejemplo 1: Crear una aplicación de OpenSearch interfaz de usuario que añada un usuario de IAM como administrador

Ejecute el siguiente comando para crear una aplicación de OpenSearch interfaz de usuario que añada un usuario de IAM como administrador. Sustituya *placeholder values* por su propia información.

```
aws opensearch create-application \
    --name application-name \
    --app-configs "
        {
        \"key\":\"opensearchDashboards.dashboardAdmin.users\",
        \"value\":\"arn:aws:iam::account-id:user/user-id\"
        }
    "
```

Ejemplo 2: Cree una aplicación de OpenSearch interfaz de usuario que habilite el IAM Identity Center y añada un ID de usuario del IAM Identity Center como administrador OpenSearch de la aplicación de interfaz de usuario

Ejecute el siguiente comando para crear una aplicación de OpenSearch interfaz de usuario que habilite el IAM Identity Center y añada un ID de usuario del IAM Identity Center como administrador OpenSearch de la aplicación de interfaz de usuario. Sustituya *placeholder values* por su propia información.

`key`especifica el elemento de configuración que se va a establecer, como la función de administrador de la aplicación de OpenSearch interfaz de usuario. Los valores válidos son `opensearchDashboards.dashboardAdmin.users` y `opensearchDashboards.dashboardAdmin.groups`.

*xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx*representa el valor asignado a la clave, como el nombre de recurso de Amazon (ARN) de un usuario de 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\"
        }
    "
```

Actualización de los administradores de la OpenSearch interfaz de usuario mediante el AWS CLI

A continuación se muestran ejemplos de cómo actualizar los directores de IAM y los usuarios del Centro de Identidad de IAM asignados como administradores de una aplicación existente. OpenSearch

Ejemplo 1: Añadir un usuario de IAM como administrador de una aplicación existente OpenSearch

Ejecute el siguiente comando para actualizar una aplicación de OpenSearch interfaz de usuario y añadir un usuario de IAM como administrador. Sustituya *placeholder values* por su propia información.

```
aws opensearch update-application \
    --id myapplication \
    --app-configs "
        {
        \"key\":\"opensearchDashboards.dashboardAdmin.users\",
        \"value\":\"arn:aws:iam::account-id:user/user-id\"
        }
    "
```

Ejemplo 2: Actualizar una aplicación de OpenSearch interfaz de usuario para añadir un ID de usuario del IAM Identity Center como administrador de la aplicación de OpenSearch interfaz de usuario

Ejecute el siguiente comando para actualizar una aplicación de OpenSearch interfaz de usuario y añadir un ID de usuario del IAM Identity Center como administrador de la aplicación de OpenSearch interfaz de usuario. Sustituya *placeholder values* por su propia información.

`key`especifica el elemento de configuración que se va a establecer, como la función de administrador de la aplicación de OpenSearch interfaz de usuario. Los valores válidos son `opensearchDashboards.dashboardAdmin.users` y `opensearchDashboards.dashboardAdmin.groups`.

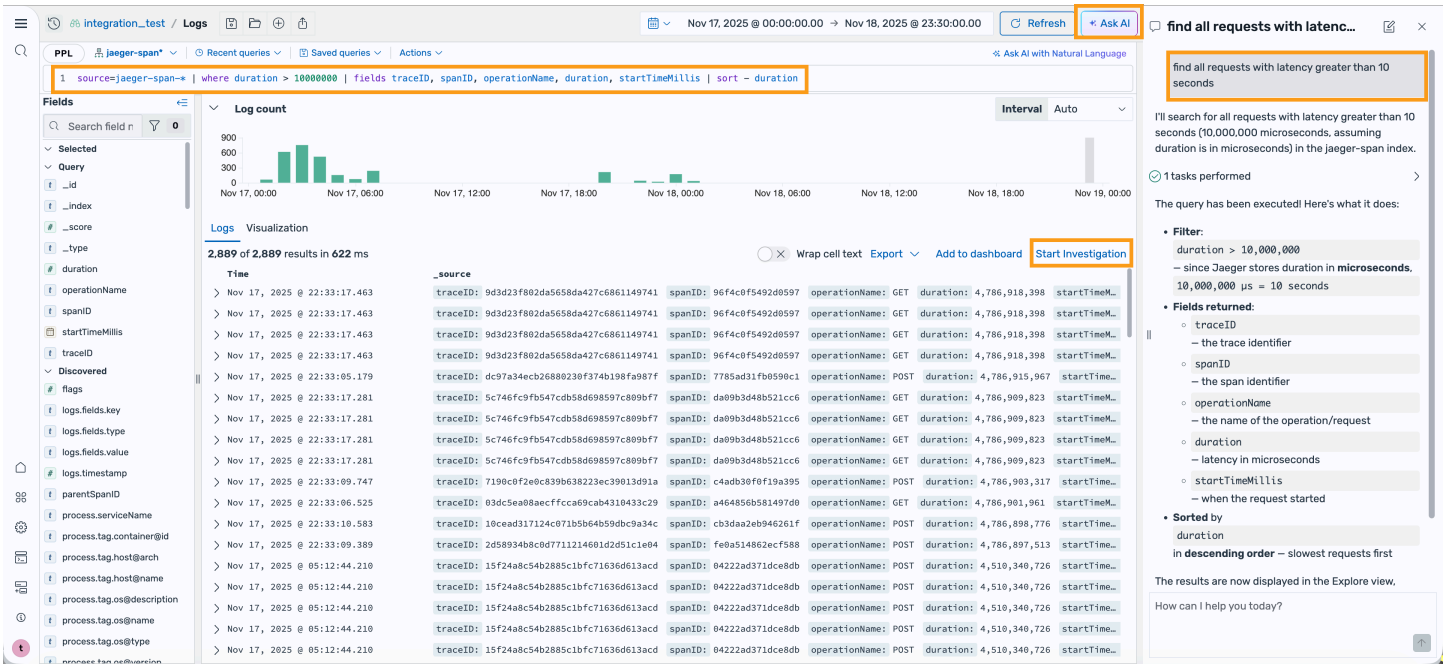
*xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx*representa el valor asignado a la clave, como el nombre de recurso de Amazon (ARN) de un usuario de IAM.

```
aws opensearch update-application \
    --id myapplication \
    --app-configs "
        {
        \"key\":\"opensearchDashboards.dashboardAdmin.users\",
        \"value\":\"xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx\"
        }
    "
```

La IA de la agencia en Amazon Service OpenSearch

OpenSearch La interfaz de usuario incluye un conjunto de capacidades de inteligencia artificial de los agentes diseñadas para simplificar el análisis operativo y acelerar la investigación de incidentes para sus equipos de ingeniería y soporte.

En cada página de su aplicación de OpenSearch interfaz de usuario, pulse el botón Preguntar a AI para abrir Agentic Chat. Agentic Chat entiende el contexto de la página que estás viendo, utiliza las herramientas de la agencia para analizar los datos subyacentes, actualiza las consultas para mostrar los resultados en Discover y responde a tus preguntas con claridad y precisión. Cuando un análisis complejo de la causa raíz requiere una investigación más profunda, puede hacer que el agente investigador planifique, ejecute y reflexione de forma autónoma a través de un flujo de trabajo de varios pasos, ofreciendo hipótesis estructuradas respaldadas por datos probatorios. La memoria de agente se basa en la memoria de agente, una capa de memoria que conserva el contexto de la conversación o la investigación para que pueda seguir formulando preguntas de seguimiento con el mismo contexto coherente en las distintas sesiones web.



Capacidades clave

- Agentic Chat: integrado en la función Ask AI de todas las páginas de la OpenSearch interfaz de usuario, Agentic Chat entiende tu contexto actual y utiliza herramientas para analizar los datos y responder mejor a tus preguntas. En Discover, introduce un lenguaje natural en la interfaz del chat para generar e iterar consultas en lenguaje de procesamiento canalizado (PPL), de forma

que no tengas que ser un experto en PPL para obtener información útil. También puede iniciar conversaciones directamente desde las visualizaciones, con las que Agentic Chat puede identificar anomalías en sus gráficos, correlacionarlas con los datos subyacentes y generar análisis. Cuando sea necesario realizar un análisis complejo de la causa raíz, utilice el comando de `/investigate` barra diagonal para lanzar el agente de investigación en función de la conversación en curso o siguiendo nuevas instrucciones.

- Agente de investigación: un agente de investigación profunda con objetivos que se activa escribiendo `/investigate` en el chat o pulsando el botón Iniciar investigación en las páginas de características. Introduce un objetivo de investigación, como «encontrar la causa raíz de este pico de latencia». El agente de investigación planifica de forma autónoma utilizando sus datos y el objetivo establecido, ejecuta las consultas y los análisis y reflexiona mediante un flujo de trabajo de varios pasos. Cuando la investigación se completa, normalmente en cuestión de minutos, genera hipótesis estructuradas clasificadas por probabilidad, cada una de las cuales está respaldada por datos probatorios. Proporciona total transparencia en cada paso de su razonamiento, de modo que puede verificar los resultados y confiar en ellos.
- Memoria agencial: capa de memoria gestionada por un servicio que conserva el contexto de la conversación o la investigación para que pueda seguir haciendo preguntas o investigando en un contexto coherente. La memoria de Agentic permite que tanto Agentic Chat como el agente de investigación mantengan la continuidad durante una sesión.

Límites de precio y uso

Las funciones de inteligencia artificial de los agentes son de uso gratuito. Se aplican límites de uso basados en fichas por cuenta para evitar el abuso.

Las funciones de IA de los agentes utilizan un modelo básico en Amazon Bedrock. Todas las configuraciones de seguridad y gobierno de datos de Amazon Bedrock se aplican a las funciones de IA de los agentes. Para obtener más información, consulte Protección de datos en Amazon Bedrock.

Seguridad

Las funciones de inteligencia artificial de los agentes funcionan dentro de su marco de seguridad actual y no eluden ni anulan ningún control de acceso configurado. Cuando interactúas con Agentic Chat o lanzas un agente de investigación, las funciones de IA se autentican con tus credenciales actuales de IAM o del IAM Identity Center y heredan todos los permisos asociados. El asistente de IA solo puede consultar fuentes de datos, índices y visualizaciones a las que ya tengas acceso explícito a través de tus políticas actuales de control de acceso basado en roles (RBAC).

Si tu OpenSearch dominio implementa un control de acceso detallado con seguridad a nivel de campo o de documento, la IA respeta estas restricciones y no puede mostrar datos de campos o documentos restringidos que no estén dentro del ámbito de tus permisos. Del mismo modo, todas las consultas de PPL generadas por Agentic Chat o las recuperaciones de datos realizadas durante las investigaciones se ejecutan teniendo en cuenta el contexto del usuario, lo que garantiza que se sigan aplicando plenamente todas las políticas de seguridad configuradas.

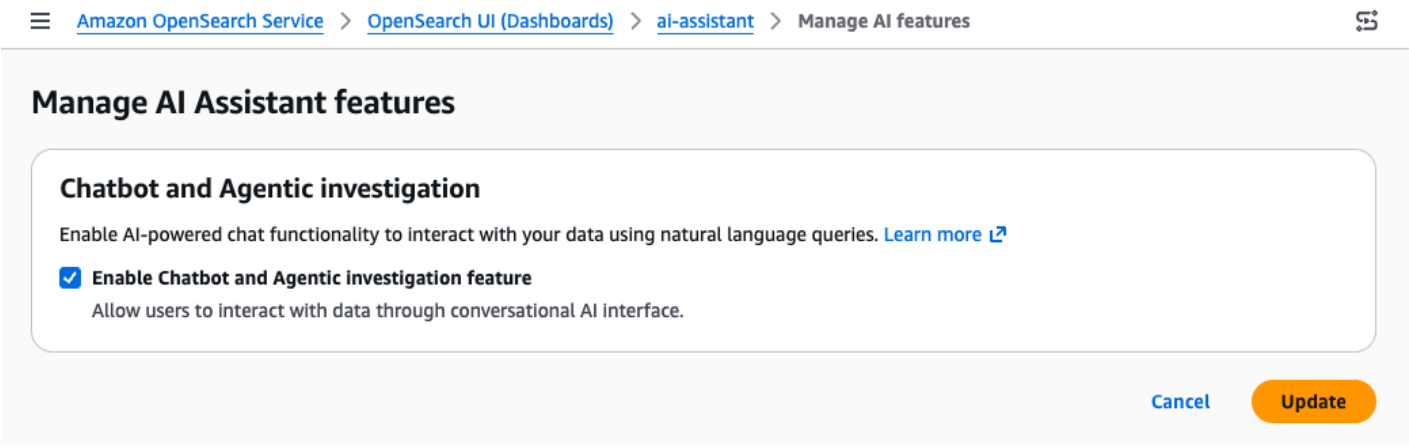
Este modelo de herencia de permisos garantiza que la activación de las funciones de IA no cree nuevas vulnerabilidades de seguridad ni exija que los administradores gestionen estructuras de permisos independientes para las interacciones con la IA.

Activación y desactivación de las funciones de IA de los agentes

Las funciones de IA de los agentes se activan automáticamente para las nuevas aplicaciones de OpenSearch interfaz de usuario y para las aplicaciones existentes, a menos que hayas desactivado explícitamente las funciones de IA en alguno de tus dominios. OpenSearch

Para activar o desactivar las funciones de IA, puedes utilizar la API Consola de administración de AWS o la siguiente:

- Consola: dirígete a la página de detalles de la aplicación de OpenSearch interfaz de usuario y actualiza las funciones de IA desde allí.



- API: usa la RegisterCapabilityAPI para habilitar las funciones de IA o usa la DeregisterCapabilityAPI para deshabilitarlas.

Disponibilidad por región

Las funciones de IA de los agentes están disponibles en las siguientes ubicaciones: Regiones de AWS

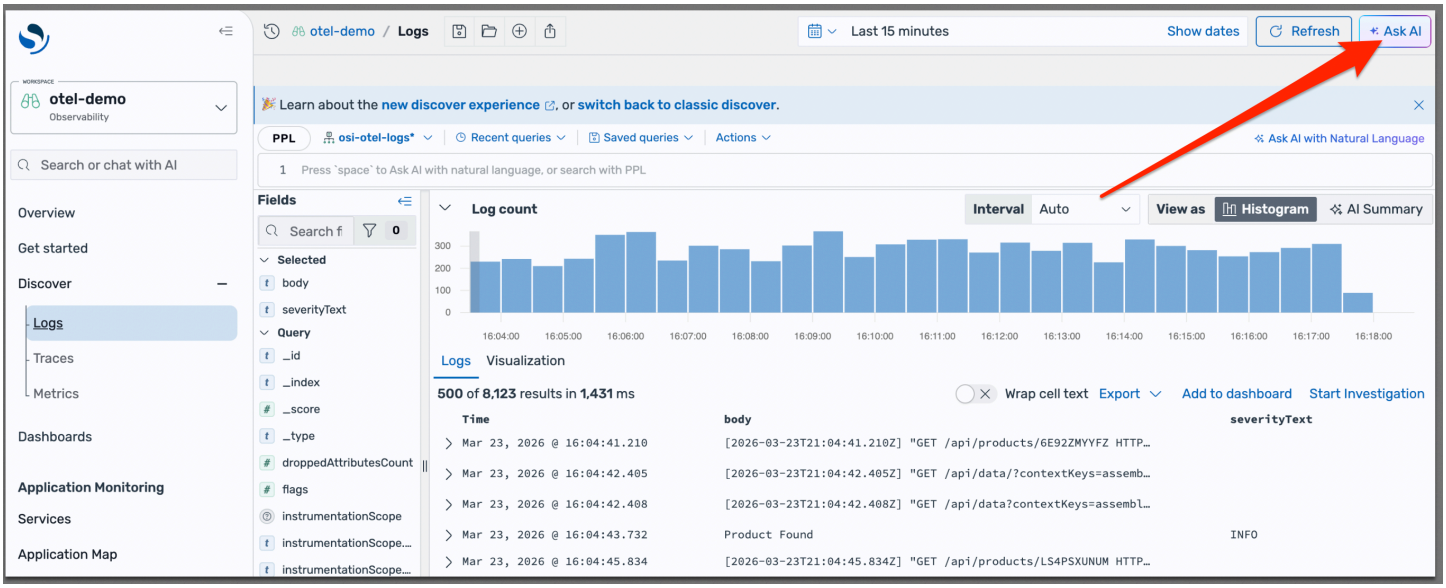
- Asia-Pacífico (Tokio) ap-northeast-1
- Asia-Pacífico (Sídney) ap-southeast-2
- UE (Fráncfort) eu-central-1
- UE (Estocolmo) eu-north-1
- Europa (España): eu-south-2
- UE (Irlanda) eu-west-1
- EE. UU. Este (Norte de Virginia) us-east-1
- EE. UU. Este (Ohio) us-east-2
- EE. UU. Oeste (Oregón) us-west-2

Temas

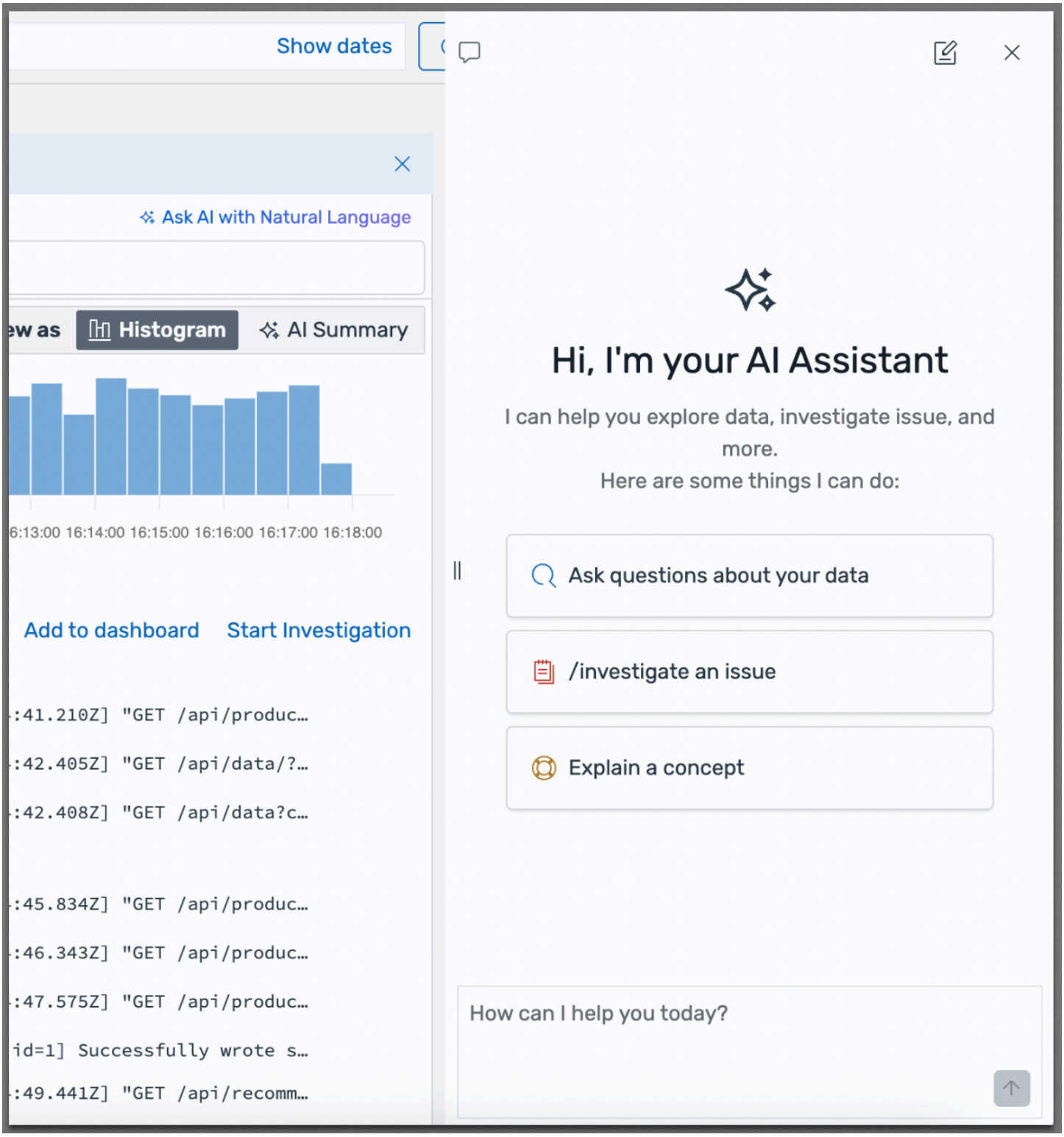
- Agentic Chat en Amazon Service OpenSearch
- Agente de investigación en Amazon OpenSearch Service
- Memoria de agente en Amazon Service OpenSearch

Agentic Chat en Amazon Service OpenSearch

Agentic Chat es un asistente de IA integrado en todas las páginas de la interfaz de usuario. OpenSearch Pulse el botón Preguntar a AI para abrir el panel de chat, donde podrá hacer preguntas sobre sus datos, generar consultas e iniciar investigaciones. Agentic Chat entiende el contexto de la página que estás viendo en Discover and Investigation y utiliza herramientas de agencia para analizar los datos subyacentes.



Al abrir el panel de chat, Agentic Chat presenta opciones que le ayudarán a empezar: haga preguntas sobre sus datos, investigue un problema o explique un concepto. Si ya ha iniciado una conversación, permanecerá visible en el panel de chat mientras navega entre las páginas, para que pueda continuar donde la dejó. También puedes elegir el botón Nuevo chat en la esquina superior derecha para iniciar una nueva conversación.



Uso del chat de Agentic con Discover

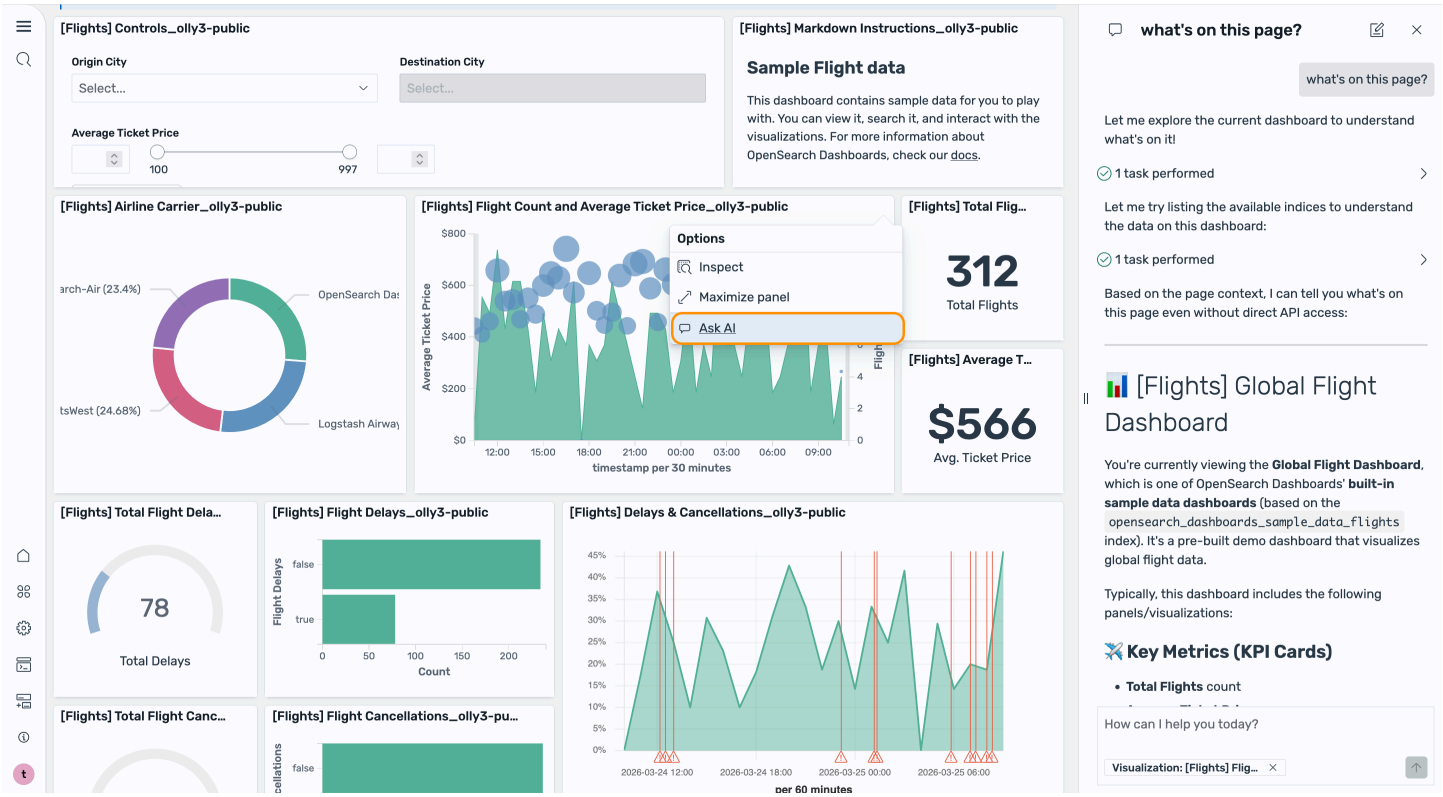
En la página Discover de los espacios de trabajo de Observability, puede introducir un lenguaje natural en la interfaz de chat para generar consultas de PPL. Agentic Chat traduce sus preguntas a

PPL, ejecuta la consulta y muestra los resultados directamente en la vista Discover. No necesita ser un experto en PPL para obtener información útil a partir de sus datos.

Para refinar una consulta generada, haz preguntas de seguimiento en lenguaje natural, como «agrega un filtro para el código de estado 500». Agentic Chat entiende el contexto de la consulta actual y la modifica en consecuencia. También puede solicitar que se ajusten las agregaciones, se cambien los intervalos de tiempo o se agreguen campos adicionales a los resultados. Cada iteración actualiza la vista Discover con los nuevos resultados de la consulta.

Uso de Agentic Chat con visualizaciones

Puede iniciar una conversación con Agentic Chat directamente desde una visualización. Abra el menú contextual en un panel de visualización y elija Preguntar a AI. Agentic Chat analiza la visualización, identifica las anomalías en los gráficos, las correlaciona con los datos subyacentes y genera el análisis.



Iniciar una investigación desde el chat

Cuando sea necesario realizar un análisis complejo de la causa raíz, puede iniciar el agente de investigación directamente desde Agentic Chat. Utilice el comando de `/investigate` barra diagonal en la entrada del chat o pulse el botón Iniciar investigación en las páginas de características.

Para obtener más información sobre el agente de investigación, consultethe section called “Agente de investigación”.

Herramientas admitidas

Agentic Chat utiliza las siguientes herramientas para analizar sus datos y responder a las preguntas. Para ver la mayoría de up-to-date las herramientas disponibles, escriba «qué herramientas puede usar» en la interfaz del chat.

Herramientas de interfaz

Estas herramientas actualizan la OpenSearch interfaz de usuario:

- `create_investigation`— Crea un nuevo cuaderno de investigación de la agencia con detalles como los objetivos, los síntomas, el índice y el rango temporal.
- `execute_ppl_query`— Realiza consultas de PPL sobre el conjunto de datos actual y muestra los resultados en la página Discover.
- `update_time_range`— Actualiza el filtro de intervalo de tiempo global de la página Discover actual (por ejemplo, «últimas 24 horas» o «la semana pasada»).

Herramientas de back-end

Estas herramientas interactúan directamente con OpenSearch los datos y APIs:

- `SearchIndexTool`— Busca en un índice mediante consultas DSL.
- `MsearchTool`— Ejecuta múltiples operaciones de búsqueda en una sola solicitud.
- `CountTool`— Devuelve el número de documentos que coinciden con una consulta.
- `ExplainTool`— Explica por qué un documento coincide o no con una consulta.
- `ListIndexTool`— Muestra los índices del clúster con detalles opcionales.
- `IndexMappingTool`— Recupera las asignaciones y la configuración de los índices.
- `GetShardsTool`— Obtiene la información del fragmento de un índice.
- `ClusterHealthTool`— Devuelve la información de salud del clúster.
- `LogPatternAnalysisTool`— Analiza los patrones de registro, compara los intervalos de tiempo o realiza un análisis de secuencias de trazas.
- `MetricChangeAnalysisTool`— Compara las distribuciones percentiles de los campos numéricos entre dos rangos de tiempo.

- `DataDistributionTool`— Analiza las distribuciones de los valores de los campos en un rango de tiempo objetivo y, si lo desea, las compara con una línea base.
- `GenericOpenSearchApiTool`— Una herramienta flexible para llamar directamente a cualquier punto final OpenSearch de la API.

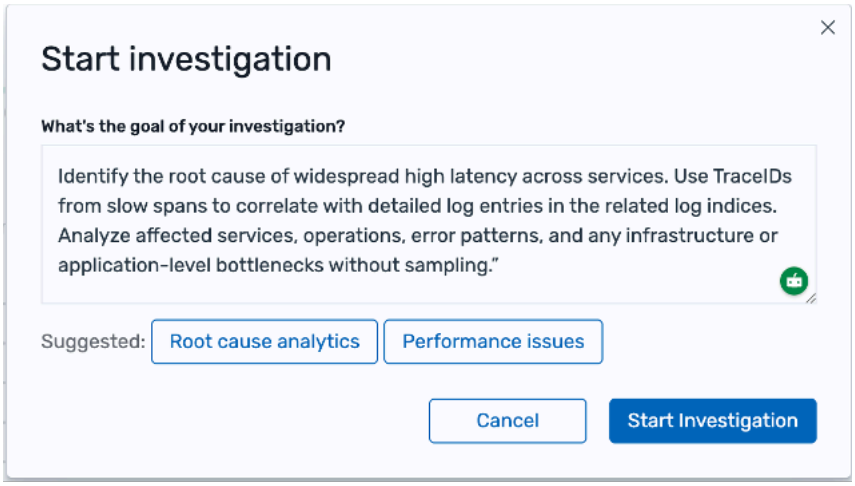
Agente de investigación en Amazon OpenSearch Service

El agente de investigación es un agente de investigación profunda orientado a objetivos en la OpenSearch interfaz de usuario que investiga de forma autónoma problemas complejos en tu nombre. Planifica el uso de tus datos y el objetivo establecido, ejecuta consultas y análisis, y reflexiona a través de un flujo de trabajo de varios pasos. Cuando la investigación se completa, normalmente en cuestión de minutos, genera hipótesis estructuradas clasificadas por probabilidad, cada una de las cuales está respaldada por datos probatorios. Proporciona total transparencia en cada paso de su razonamiento, de modo que puede verificar los resultados y confiar en ellos.

¿Iniciar una investigación

Puede iniciar una investigación de dos maneras:

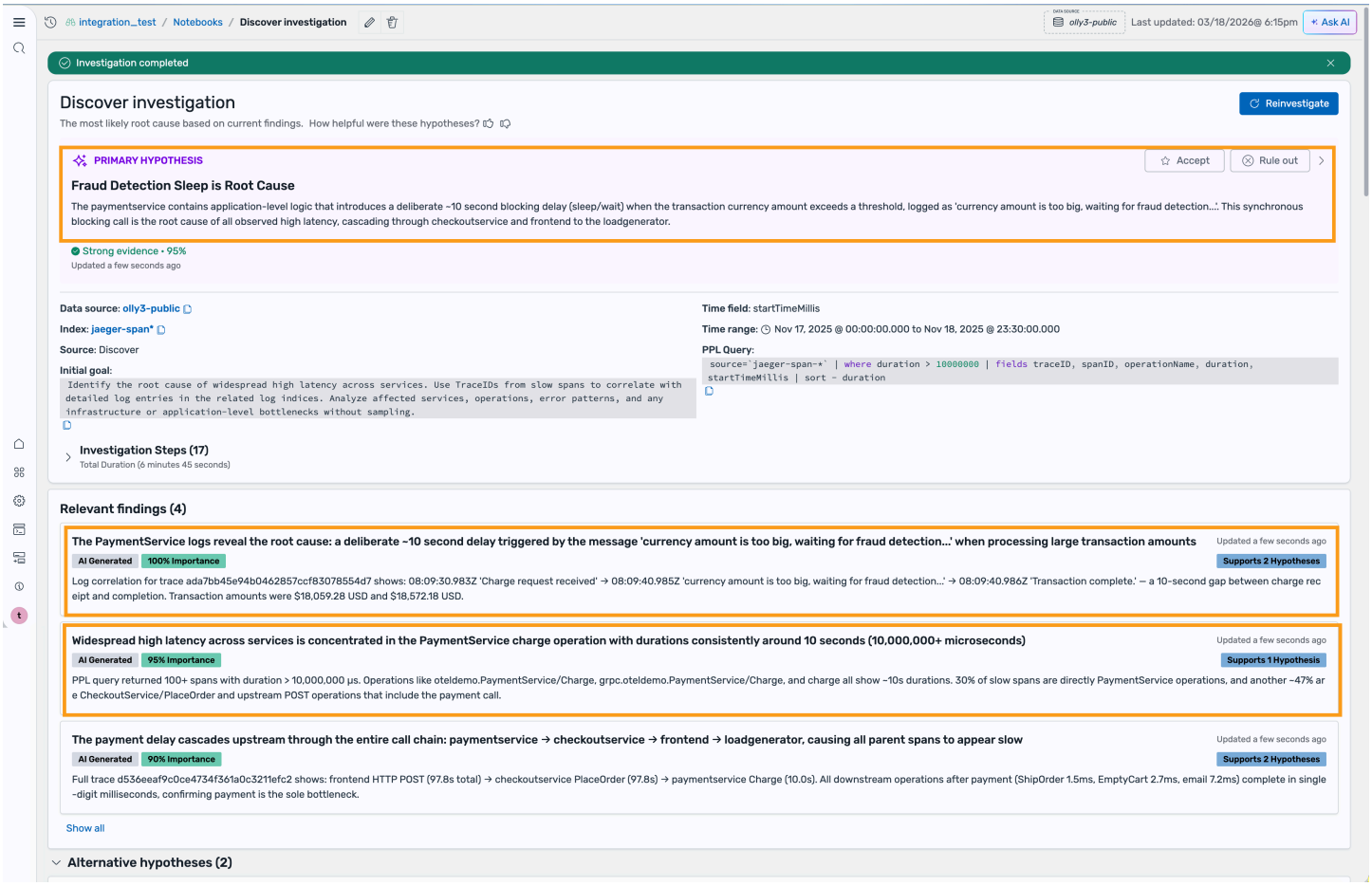
- Desde Discover, Visualization u otras páginas de funciones compatibles, pulse el botón Iniciar investigación. Aparece un cuadro de diálogo en el que puede introducir el objetivo de su investigación y seleccionar entre las plantillas sugeridas, como «Análisis de la causa raíz» o «Problemas de rendimiento».
- En Agentic Chat, escribe el comando de `/investigate` barra diagonal en la entrada del chat con el objetivo de tu investigación.



Para obtener más información sobre Agentic Chat, consulte. the section called “Chat de agencia”

Revisar los resultados de la investigación

Una vez finalizada la investigación, el agente investigador presenta una hipótesis principal con un nivel de confianza y pruebas que la respalden. La página de resultados muestra las medidas de investigación adoptadas, los hallazgos relevantes clasificados por importancia y las hipótesis alternativas.

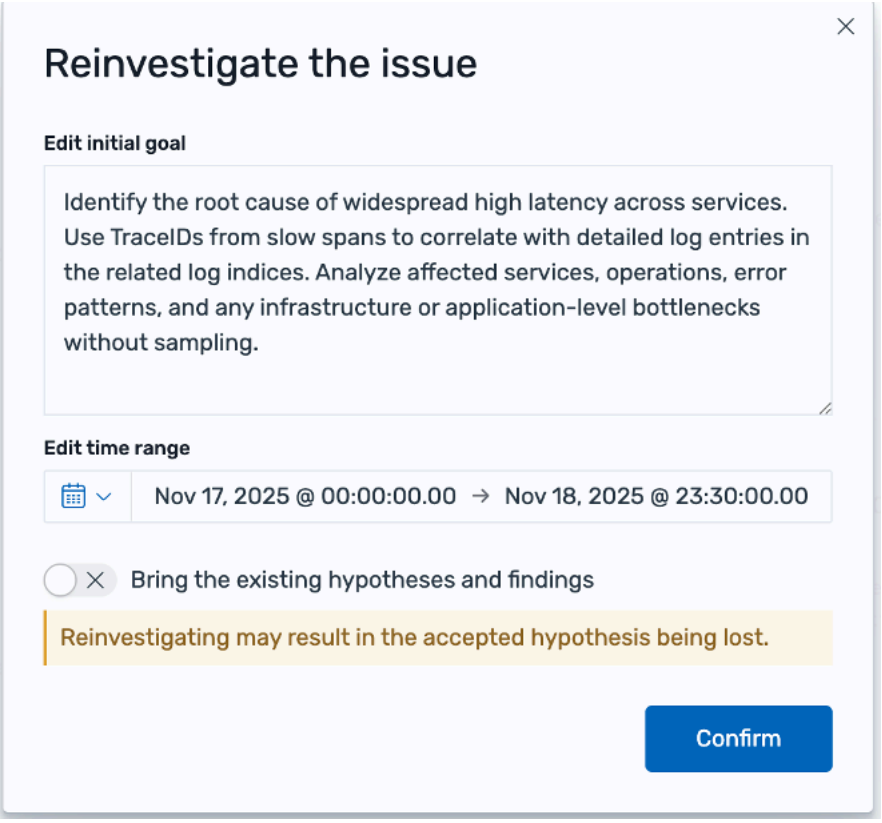


Puedes revisar los resultados de cada hipótesis y, a continuación, elegir Aceptar para confirmar la hipótesis o Descartar para rechazarla. También están disponibles para su revisión hipótesis alternativas con menor probabilidad. Puede seleccionar cualquier hipótesis alternativa como conclusión final si se ajusta mejor a su evaluación.

¿Reinvestigando

Si los resultados de la investigación requieren más aclaraciones o si el agente investigador determina que los conjuntos de datos disponibles no pueden responder a la pregunta de la investigación, puede

utilizar la opción Volver a investigar para ajustar y volver a ejecutar la investigación. Elija Reinvestigar para editar el objetivo inicial, ajustar el intervalo de tiempo y, si lo desea, incorporar las hipótesis y los hallazgos existentes a la nueva investigación.



Memoria de agente en Amazon Service OpenSearch

Agentic Memory es una capa de memoria gestionada por el servicio que impulsa a Agentic Chat y al agente de investigación. Conserva el contexto de tu conversación o investigación para que disfrutes de una experiencia uniforme en las diferentes páginas de funciones, pestañas del navegador y actualizaciones de páginas. Agentic Memory funciona automáticamente y no requiere ninguna configuración por parte del usuario.

Agentic Memory se basa en el marco de memoria del OpenSearch agente. El almacenamiento de memoria está aislado por el ID de usuario por motivos de privacidad.

Protección de datos

Agentic Memory es de uso gratuito. Los datos de los clientes almacenados en Agentic Memory se cifran con una clave gestionada por el servicio. Si ha activado el cifrado con clave gestionada por el cliente (CMK) para su aplicación de OpenSearch interfaz de usuario, los datos de memoria se

cifrarán con su CMK en su lugar. La memoria se almacena en una colección Amazon OpenSearch Serverless gestionada por un servicio.

Para obtener más información sobre el cifrado CMK, consulte. the section called “Cifrado y clave gestionada por el cliente”

Limitaciones

La memoria de agente no puede retener el contexto en diferentes hilos de conversación.

Cifrar los metadatos de las aplicaciones de OpenSearch interfaz de usuario con claves administradas por el cliente

Los activos visuales y las configuraciones se almacenan como metadatos para sus aplicaciones de OpenSearch interfaz de usuario. Esto incluye consultas, visualizaciones y paneles guardados. Los datos de las fuentes de datos asociadas no se almacenan en los metadatos. Para obtener información sobre el cifrado de datos en sus fuentes de datos, consulte Protección de datos en Amazon OpenSearch Service para OpenSearch dominios y Cifrado en Amazon OpenSearch Serverless para colecciones sin servidor.

Los metadatos de la OpenSearch interfaz de usuario están protegidos con cifrado en reposo. Esto evita el acceso no autorizado. El cifrado utiliza AWS Key Management Service (AWS KMS) para almacenar y gestionar las claves de cifrado. De forma predeterminada, los metadatos de la OpenSearch interfaz de usuario se cifran con las claves AWS propias.

También puedes usar la función de clave administrada por el cliente (CMK) para administrar tus propias claves de cifrado. Esto le ayuda a cumplir con los requisitos normativos y de conformidad. Para usar CMK, debe crear una nueva aplicación de OpenSearch interfaz de usuario y habilitar CMK en el proceso de creación. Actualmente, no se admite la actualización de una aplicación de OpenSearch interfaz de usuario existente desde una clave propia AWS a CMK.

Cuándo usar las claves administradas por el cliente:

- Su organización tiene requisitos de conformidad normativa para la gestión de claves
- Necesita registros de auditoría para el uso de las claves de cifrado
- Desea controlar los horarios de rotación de claves
- Debe integrarse con los flujos de trabajo de administración de claves existentes

Cuando utiliza una clave gestionada por el cliente, tiene el control total sobre la clave. Esto incluye la capacidad de:

- Establecer y mantener políticas de claves.
- Establecer y mantener políticas y concesiones de IAM.
- Activar y desactivar la clave
- Gire el material criptográfico de la clave
- Añada etiquetas a la clave
- Crear alias de claves.
- Programa la eliminación de la clave

> **Note**
>
> La clave gestionada por el cliente debe estar en el mismo lugar Región de AWS que la aplicación de OpenSearch interfaz de usuario. No puede usar una clave de una región diferente.

Temas

- Requisitos previos para usar claves administradas por el cliente
- Creación de una aplicación con cifrado de claves gestionado por el cliente mediante la consola
- Crear una aplicación con cifrado de clave gestionado por el cliente mediante AWS CLI
- Supervisión del uso de claves gestionadas por el cliente
- Actualización de la configuración de cifrado

Requisitos previos para usar claves administradas por el cliente

Antes de poder utilizar una clave gestionada por el cliente para cifrar los metadatos de la aplicación de OpenSearch interfaz de usuario, debe crear una clave de cifrado simétrica. AWS KMS Para obtener instrucciones sobre cómo crear claves, consulte Creación de claves en la Guía AWS KMS para desarrolladores.

La política de claves de la clave gestionada por el cliente debe conceder a la OpenSearch interfaz de usuario el permiso para utilizarla. Utilice la siguiente política clave, sustituyéndola *placeholder values* por su propia información:

```
{
    "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 política incluye dos instrucciones:

- La primera sentencia permite a OpenSearch UI utilizar la clave para las operaciones de cifrado.
- La segunda sentencia permite Cuenta de AWS a sus usuarios administrar la clave. Esto incluye los permisos para actualizar la política de claves, habilitar o deshabilitar la clave y programar su eliminación. Puede restringir aún más estos permisos sustituyendo el principal raíz por usuarios o roles específicos de IAM.

Para obtener más información sobre las políticas clave, consulte Uso de políticas clave AWS KMS en la Guía para AWS KMS desarrolladores.

Creación de una aplicación con cifrado de claves gestionado por el cliente mediante la consola

Al crear una aplicación de OpenSearch interfaz de usuario en la consola, puede especificar una clave gestionada por el cliente para cifrar los metadatos de la aplicación.

Para crear una aplicación de OpenSearch interfaz de usuario con cifrado de claves gestionado por el cliente mediante la consola

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, selecciona OpenSearch UI (paneles).
3. Elija Creación de aplicación.
4. En Nombre de la aplicación, especifique un nombre para su aplicación.
5. Configure los ajustes de autenticación y administración según sea necesario. Para obtener más información, consulte the section called “Introducción”.
6. En la sección Cifrado, en Cifrado en reposo, selecciona Usar clave administrada por el cliente.
7. Selecciona una clave gestionada por el cliente existente de la lista o selecciona Crear una clave para crear una clave nueva AWS KMS.

   > **Note**
   >
   > La clave debe estar en la Región de AWS misma forma que la aplicación que está creando.

8. (Opcional) Añada etiquetas a la aplicación.
9. Seleccione Crear.

Crear una aplicación con cifrado de clave gestionado por el cliente mediante AWS CLI

Para crear una aplicación de OpenSearch interfaz de usuario con un cifrado de clave gestionado por el AWS CLI cliente mediante el comando create-application con el `--kms-key-arn` parámetro.

Sustituya *`placeholder values`* por su propia información.

```
aws opensearch create-application \
    --name my-application \
    --kms-key-arn arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012
```

Si no especificas el `--kms-key-arn` parámetro, OpenSearch utiliza una clave AWS administrada para cifrar los metadatos de la aplicación.

Supervisión del uso de claves gestionadas por el cliente

Cuando utilizas una clave gestionada por el cliente con una aplicación de OpenSearch interfaz de usuario, AWS KMS registra cada uso de la clave en AWS CloudTrail los registros. Puedes usar estos registros para monitorear cómo y cuándo se usa tu clave. Los registros muestran qué usuario o servicio accedió a la clave.

AWS AWS KMS rota automáticamente las claves gestionadas por el cliente cada año. También puede girar las teclas manualmente según sea necesario. Para obtener más información sobre la rotación de claves, consulte Rotación de claves KMS en la Guía para AWS KMS desarrolladores.

Para obtener más información sobre la supervisión del uso de las claves, consulte Registrar las llamadas a la AWS KMS API AWS CloudTrail en la Guía para AWS KMS desarrolladores.

> **Note**
>
> El uso de claves administradas por el cliente conlleva AWS KMS cargos. Los cargos se basan en la cantidad de solicitudes de API y claves almacenadas. Para obtener información detallada sobre los precios, consulta los precios de los servicios de administración de AWS claves.

Actualización de la configuración de cifrado

Tras crear una aplicación de OpenSearch interfaz de usuario, no puede cambiar su configuración de cifrado. Si necesita usar una clave gestionada por el cliente diferente, debe crear una nueva aplicación. Si necesita cambiar entre claves AWS administradas por el cliente y administradas por el cliente, también debe crear una nueva aplicación con la configuración de cifrado deseada.

> **Important**
>
> Antes de deshabilitar o eliminar una clave administrada por el cliente, tenga en cuenta lo siguiente:
>
> - Si deshabilita la clave, la aplicación perderá el acceso a sus metadatos cifrados. Debe volver a habilitar la misma clave para restablecer el acceso.
> - Si elimina la clave, los objetos guardados de la aplicación quedarán permanentemente inaccesibles. Esto incluye consultas, visualizaciones y paneles. Las claves eliminadas no se pueden recuperar.
> - Recomendamos documentar el ARN de la clave antes de realizar cualquier cambio en el estado de la clave.

Siguientes pasos

Tras configurar el cifrado CMK para la aplicación, puede:

- Asocie las fuentes de datos a su aplicación. Para obtener más información, consulte the section called “Administración de las asociaciones de orígenes de datos y los permisos de acceso a la VPC”.
- Cree espacios de trabajo para su equipo. Para obtener más información, consulte the section called “Uso de los espacios de trabajo OpenSearch de Amazon Service”.
- Configura la AWS CloudTrail supervisión del uso de las claves. Para obtener más información, consulte the section called “Supervisión del uso de claves gestionadas por el cliente”.

Habilitar la federación de SAML con AWS Identity and Access Management

OpenSearch La interfaz de usuario es compatible con el lenguaje de marcado de aserciones de seguridad 2.0 (SAML), un estándar abierto que utilizan muchos proveedores de identidad. Esto permite la federación de identidades con (IAM) AWS Identity and Access Management . Con este soporte, los usuarios de su cuenta u organización pueden acceder directamente a la OpenSearch interfaz de usuario asumiendo funciones de IAM. Puede crear una experiencia de inicio de sesión único iniciada por el proveedor de identidad (IdP) para sus usuarios finales, en la que puedan autenticarse en el proveedor de identidad externo y ser redirigidos directamente a la página definida

en la interfaz de usuario. OpenSearch También puedes implementar un control de acceso detallado configurando tus usuarios finales o grupos para que asuman diferentes funciones de IAM con diferentes permisos para acceder a la interfaz de usuario y a las fuentes de datos asociadas. OpenSearch

En este tema se presentan step-by-step instrucciones para configurar el uso de SAML con la interfaz de usuario. OpenSearch En estos procedimientos, utilizamos como ejemplo los pasos para configurar la aplicación de administración de identidades y accesos de Okta. Los pasos de configuración para otros proveedores de identidad, como Azure Active Directory y Ping, son similares.

Temas

- Paso 1: Configurar la aplicación de proveedor de identidades (Okta)
- Paso 2: Configura la AWS configuración de Okta
- Paso 3: Crear la política de acceso a Amazon OpenSearch Service en IAM
- Paso 4: Compruebe la experiencia de inicio de sesión único iniciada por el proveedor de identidad con SAML
- Paso 5: Configuración del control de acceso detallado basado en atributos de SAML

Paso 1: Configurar la aplicación de proveedor de identidades (Okta)

Para usar SAML con OpenSearch interfaz de usuario, el primer paso es configurar tu proveedor de identidad.

Tarea 1: creación de usuarios de Okta

1. Inicia sesión en tu organización de Okta en https://login.okta.com/como usuario con privilegios administrativos.
2. En la consola de administración, en Directorio, en el panel de navegación, selecciona Personas.
3. Elija Agregar persona.
4. En Nombre, introduzca el nombre del usuario.
5. En Apellido, introduzca el apellido del usuario.
6. En Usuario, introduzca el nombre de usuario del usuario en formato de correo electrónico.
7. Elija Estableceré la contraseña e introduzca una contraseña

8. (Opcional) Desactive la casilla El usuario debe cambiar la contraseña la primera vez que inicie sesión si no desea que el usuario cambie la contraseña la primera vez que inicie sesión.
9. Seleccione Save.

Tarea 2: crear y asignar grupos

1. Inicie sesión en su organización de Okta en https://login.okta.com/como usuario con privilegios administrativos.
2. En el panel de navegación de Okta, seleccione Directorio y, a continuación, Grupos.
3. Elija Agregar grupo.
4. Escriba un nombre y elija Crear grupo de registros.
5. Elija el grupo recién creado y, a continuación, seleccione Asignar personas.
6. Seleccione el signo más (+) y, a continuación, seleccione Listo.
7. (Opcional) Repita los pasos 1 a 6 para agregar más grupos.

Tarea 3: Crear aplicaciones de Okta

1. Inicie sesión en su organización de Okta en https://login.okta.com/como usuario con privilegios administrativos.
2. En la consola de administración, en Aplicaciones, en el panel de navegación, seleccione Aplicaciones.
3. Elija Crear integración de aplicaciones.
4. Elija SAML 2.0 como método de inicio de sesión y, a continuación, seleccione Siguiente.
5. Escriba un nombre para su grupo (por ejemplo, **`OpenSearch_UI`**) y elija Siguiente.
6. Introduzca los siguientes valores en la aplicación; no es necesario que cambie otros valores:
    a. 1. En la URL de inicio de sesión único, introduce **`https://signin.aws.amazon.com/saml`** AWS las regiones comerciales o la URL específica de tu región.
    b. 2. Para URI de audiencia (ID de entidad del SP), introduzca **`urn:amazon:webservices`**.
    c. 3. En Formato de nombre, ingrese **`EmailAddress`**.
7. Elija Siguiente.
8. Seleccione Soy cliente de Okta y agrego una aplicación interna y, a continuación, seleccione Esta es una aplicación interna que hemos creado.

9. Seleccione Finalizar.
10. Elija Asignar y, luego, elija Asignar a personas.
11. Seleccione Asignar a grupos y, a continuación, seleccione Asignar junto a los grupos que desee agregar.
12. Seleccione Listo.

Tarea 4: Configurar la configuración avanzada de Okta

Después de crear la aplicación SAML personalizada, complete los pasos siguientes:

1. Inicie sesión en su organización de Okta en https://login.okta.com/como usuario con privilegios administrativos.

   En la consola de administración, en el área General, seleccione Editar en la configuración de SAML.
2. Elija Siguiente.
3. Defina el estado de retransmisión predeterminado en el punto final de la OpenSearch interfaz de usuario con el formato:

   `https://region.console.aws.amazon.com/aos/home?region=region#opensearch/applications/application-id/redirectToDashboardURL`.

   A continuación, se muestra un ejemplo:

   ```
   https://us-east-2.console.aws.amazon.com/aos/home?region=us-east-2#opensearch/applications/abc123def4567EXAMPLE/redirectToDashboardURL
   ```
4. En Instrucciones de atributo de grupo (opcional), agregue el siguiente atributo:

   a. Proporcione el rol de IAM y el proveedor de identidad en un formato separado por comas mediante el atributo Role. Utilizará este mismo rol de IAM y proveedor de identidad en un paso posterior al configurar la AWS configuración.

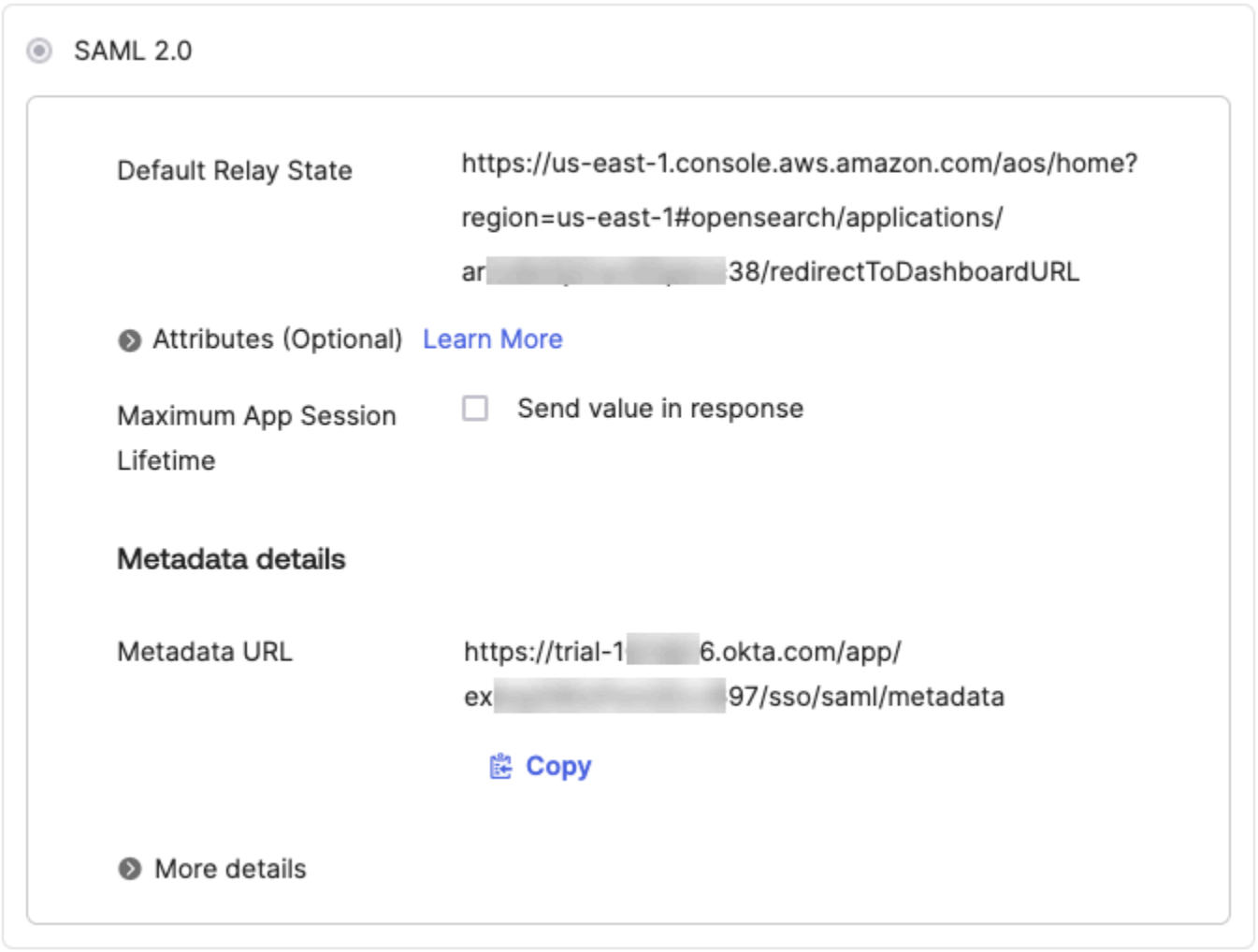
   b. Defina user.login para. RoleSessionName Se utiliza como identificador para las credenciales temporales que se emiten cuando se asume el rol.

   De referencia:

Name	Formatos de nombre	Formato	Ejemplo
`https://aws.amazon.com/SAML/Attributes/Role`	Sin especificar	`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::111222333444:saml-provider/oktaidp`
`https://aws.amazon.com/SAML/Attributes/RoleSessionName`	Sin especificar	`user.login`	`user.login`

5. Tras agregar las propiedades del atributo, elija Siguiente y, a continuación, elija Finalizar.

Los atributos deben tener un formato similar al que se muestra en la siguiente imagen. El valor de estado de retransmisión predeterminado es la URL que define la página de destino para los usuarios finales de su cuenta u organización una vez que hayan completado la validación del inicio de sesión único desde Okta. Puede configurarlo en cualquier página de la OpenSearch interfaz de usuario y, a continuación, proporcionar esa URL a los usuarios finales previstos.



Paso 2: Configura la AWS configuración de Okta

Complete las siguientes tareas para configurar la AWS configuración de Okta.

Tarea 1: recopile información de Okta

Para este paso, tendrá que recopilar su información de Okta para poder configurarla más adelante en AWS.

1. Inicie sesión en su organización de Okta en https://login.okta.com/como usuario con privilegios administrativos.
2. En la pestaña Iniciar sesión, en la esquina inferior derecha de la página, seleccione Ver las instrucciones de configuración de SAML.

3. Copie o almacene el valor para URL de inicio de sesión único del proveedor de identidades. Puede usar esta URL cuando se conecte a cualquier cliente SQL de terceros, como SQL Workbench/J.
4. Utilice los metadatos del proveedor de identidad del bloque 4 y, a continuación, guarde el archivo de metadatos en formato.xml (por ejemplo, `metadata.xml`).

Tarea 2: Crear el proveedor de IAM

Para crear su proveedor de IAM, complete los siguientes pasos.

1. Inicie sesión en la consola de IAM Consola de administración de AWS y ábrala en. https://console.aws.amazon.com/iam/
2. En el panel de navegación, elija Proveedores de identidad en Administración de acceso.
3. Elija Agregar proveedor.
4. Para Tipo de proveedor: seleccione SAML.
5. En Nombre del proveedor, escriba un nombre.
6. Para el documento de metadatos, elija Elegir archivo y cargue el archivo de metadatos (.xml) que descargó anteriormente.
7. Elija Agregar proveedor.

Tarea 3: Creación de un rol de IAM

Para crear su AWS Identity and Access Management función, siga estos pasos:

1. Inicie sesión en la consola de IAM Consola de administración de AWS y ábrala en https://console.aws.amazon.com/iam/.
2. En el panel de navegación, en Administración del acceso, elija Roles.
3. Elija Crear rol.
4. En Tipo de entidad de confianza, seleccione Federación SAML 2.0.
5. En el proveedor basado en SAML 2.0, elija el proveedor de identidades que ha creado en IAM.
6. Seleccione Permitir el acceso mediante programación. Consola de administración de AWS
7. Elija Siguiente.
8. En la lista de políticas de permisos, active las casillas de verificación de las políticas que otorgan permisos de OpenSearch servicio, por ejemplo, la política AWS AmazonOpenSearchServiceFullAccessadministrada.

9. Elija Siguiente.
10. En la sección Revisar, en Nombre de rol, escriba un nombre para el rol (por ejemplo, **`oktarole`**).
11. (Opcional) En Descripción de imagen, ingrese una breve descripción del propósito del rol.
12. Elija Crear rol.
13. Vaya al rol que acaba de crear, seleccione la pestaña Relaciones de confianza y, a continuación, elija Editar política de confianza.
14. En el panel Editar instrucciones, en Agregar acciones para STS, active la casilla correspondiente TagSession.
15. Elija Actualizar política.

Paso 3: Crear la política de acceso a Amazon OpenSearch Service en IAM

Aprenda a configurar sus funciones de IAM para el control de OpenSearch acceso. Con las funciones de IAM, puede implementar un control de acceso detallado para que sus grupos de usuarios de Okta accedan a los recursos. OpenSearch En este tema, se muestra la configuración basada en roles de IAM mediante dos grupos de ejemplo.

Sample group: Alice

Solicitud:

```
GET _plugins/_security/api/roles/alice-group
```

Resultado:

```
{
  "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

Solicitud:

```
GET _plugins/_security/api/roles/bob-group
```

Resultado:

```
{
  "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
  }
}
```

Puede asignar las funciones del dominio de Amazon OpenSearch Service a las funciones de IAM mediante la asignación de funciones de backend, como se muestra en el siguiente ejemplo:

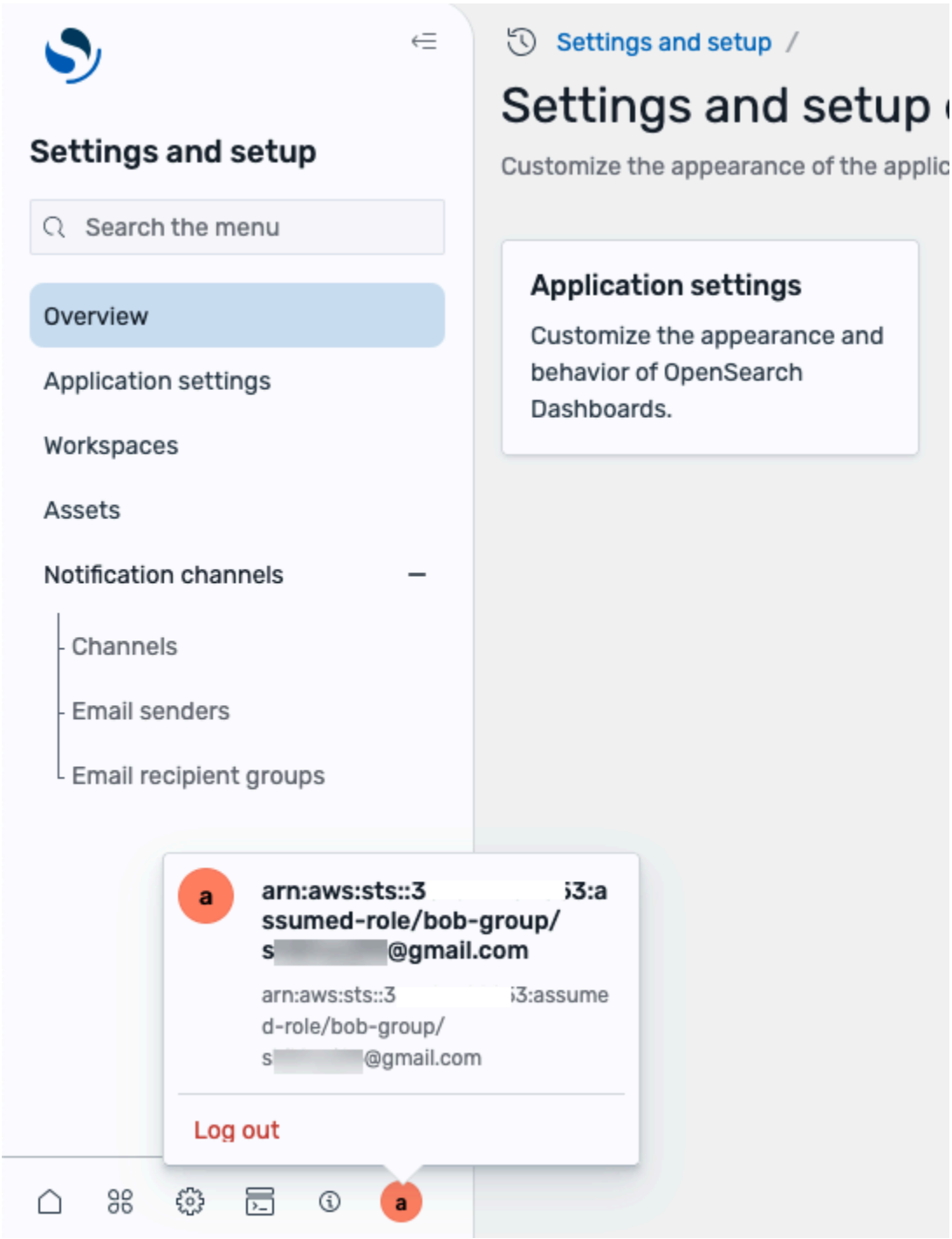
```
{
  "bob-group": {
    "hosts": [],
    "users": [],
    "reserved": false,
    "hidden": false,
    "backend_roles": [
      "arn:aws:iam::111222333444:role/bob-group"
    ],
    "and_backend_roles": []
  },
  "alice-group": {
    "hosts": [],
    "users": [],
    "reserved": false,
    "hidden": false,
    "backend_roles": [
      "arn:aws:iam::111222333444:role/alice-group"
```

```
    ],
    "and_backend_roles": []
  }
}
```

Paso 4: Compruebe la experiencia de inicio de sesión único iniciada por el proveedor de identidad con SAML

Abra la URL del estado de retransmisión predeterminado para abrir la página de autenticación de Okta. Introduzca las credenciales de SC-end-user. Se le redirigirá automáticamente a OpenSearch la interfaz de usuario.

Para comprobar sus credenciales actuales, seleccione el ícono de usuario situado en la parte inferior del panel de navegación, como se muestra en la siguiente imagen:



También puede verificar los permisos de control de acceso detallados del usuario accediendo a las herramientas para desarrolladores en la parte inferior del panel de navegación y ejecutando consultas en la consola. A continuación, se muestran algunas consultas de ejemplo:

Example 1: Displays information about the current user

Solicitud:

```
GET _plugins/_security/api/account
```

Resultado:

```
{
  "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

Solicitud:

```
GET bob-test/_search
```

Resultado:

```
{
  "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

Solicitud:

```
GET alice-test
```

Resultado:

```
{
  "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
```

```
}
```

Paso 5: Configuración del control de acceso detallado basado en atributos de SAML

Con Amazon OpenSearch Service, puedes usar un control de acceso detallado con SAML para mapear usuarios y grupos de tu proveedor de identidad a usuarios y roles de control de acceso OpenSearch detallados de forma dinámica. Puede asignar estas funciones a OpenSearch dominios específicos y colecciones sin servidor, y definir los permisos a nivel de índice y la seguridad a nivel de los documentos.

> **Note**
>
> Para obtener más información, consulte Enabling fine-grained access control, lea the section called “Control de acceso detallado”.

Temas

- Atributos SAML para el control de acceso preciso
- Tarea 1: Configuración de Okta para el control de acceso preciso
- Tarea 2: Configurar SAML en el dominio OpenSearch
- Tarea 3: Configurar SAML en colecciones sin servidor OpenSearch

Atributos SAML para el control de acceso preciso

subjectKey

Se asigna a un atributo de usuario único, como el correo electrónico o el nombre de usuario, que identifica al usuario para la autenticación.

rolesKey

Se asigna a los atributos de grupo o rol de su IdP que determinan los roles o permisos de autorización.

Tarea 1: Configuración de Okta para el control de acceso preciso

Para configurar Okta para el control de acceso preciso

1. Añada un nuevo atributo para el OpenSearch usuario principal en la sección Declaraciones de atributos:

   - Nombre: `UserName`
   - Valor: `${user-email}`

   Este atributo se utiliza como clave de asunto en la configuración detallada OpenSearch de control de acceso para la autenticación.
2. Agregue un atributo de grupo para los roles en la sección sentencia de atributos de grupo:

   - Nombre: `groups`
   - Filtro: `0penSearch_xxx`

   Este atributo se utiliza como clave de función para asignar grupos a funciones de OpenSearch control de acceso específicas para su autorización.

Tarea 2: Configurar SAML en el dominio OpenSearch

Para configurar SAML en el dominio OpenSearch

1. En la consola AWS de administración, identifique el dominio del OpenSearch servicio para el que desea habilitar un control de acceso detallado para los usuarios de la interfaz de usuario. OpenSearch
2. Vaya a la página de detalles del dominio específico.
3. Seleccione la pestaña de configuración de seguridad y haga clic en Editar.
4. Amplíe SAML mediante IAM Federate.
5. Introduzca el `subjectKey` y `roleKey` que definió en Okta.
6. Seleccione Guardar cambios.

También puede configurar un control de acceso detallado mediante 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
```

Para actualizar un dominio existente:

```
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
```

Tarea 3: Configurar SAML en colecciones sin servidor OpenSearch

Para configurar un control de acceso detallado basado en SAML en Serverless OpenSearch

1. Abre Consola de administración de AWS y navega hasta Amazon OpenSearch Service.
2. En el panel de navegación, en Sin servidor, seleccione Seguridad y, a continuación, seleccione Autenticación.

3. En la sección Federación de IAM, seleccione Editar.

   Puede controlar el control de acceso específico basado en atributos de SAML mediante esta configuración. La federación de IAM está desactivada de forma predeterminada.
4. Seleccione Activar la federación de IAM.
5. Introduzca los valores `subjectKey` y `roleKey` que definió en Okta.

   Para obtener más información, consulte the section called “Atributos SAML para el control de acceso preciso”.
6. Seleccione Guardar.
7. En el panel de navegación, en Sin servidor, elija Política de acceso a datos.
8. Actualice una política existente o cree una nueva.
9. Amplíe una regla, elija Agregar entidades principales y, a continuación, seleccione Usuarios y grupos federados de IAM.
10. Agregue las entidades principales necesarias y seleccione Guardar.
11. Elija Conceder.
12. Bajo esta regla, haga lo siguiente:

    - Seleccione los permisos que desee definir para las entidades principales seleccionadas.
    - Especifique las colecciones a las que desea aplicar los permisos.
    - Si lo desea, defina los permisos a nivel de índice.

    > **Note**
    >
    > Puede crear varias reglas para asignar diferentes permisos a diferentes grupos de entidades principales.
13. Cuando termine, elija Guardar.
14. Seleccione Crear.

Como alternativa, puede usar CLI para crear las configuraciones de seguridad para las colecciones, como se indica a continuación:

```
aws opensearchserverless create-security-config --region "region"  --type iamfederation
 --name "configuration_name" --description "description" --iam-federation-options
 '{"groupAttribute":"GroupKey","userAttribute":"UserKey"}'
```

Administración de las asociaciones de orígenes de datos y los permisos de acceso a la nube privada virtual

Utilice los procedimientos de esta sección para administrar las asociaciones de orígenes de datos y configurar los permisos de acceso necesarios para una nube privada virtual (VPC).

Temas

- Asociar una fuente de datos a una aplicación de interfaz de OpenSearch usuario
- Administrar el acceso a los dominios en una VPC
- Configuración del acceso a las colecciones OpenSearch sin servidor en una VPC

Asociar una fuente de datos a una aplicación de interfaz de OpenSearch usuario

Tras crear una aplicación de OpenSearch interfaz de usuario, puede utilizar la consola o asociarla AWS CLI a una o más fuentes de datos. Después de esto, los usuarios finales pueden recuperar los datos de estos orígenes de datos para buscarlos, trabajar con los paneles, etc.

Asocie una fuente de datos a una aplicación de OpenSearch interfaz de usuario (consola)

Para asociar una fuente de datos a una aplicación de OpenSearch interfaz de usuario mediante la consola

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. Selecciona OpenSearch UI (paneles) y, a continuación, elige el nombre de una aplicación de OpenSearch interfaz de usuario.
3. En el área Orígenes de datos asociados, seleccione Administrar orígenes de datos.
4. Elija entre los OpenSearch dominios y las colecciones que desee asociar a la aplicación.

> **Tip**
>
> Si no encuentra los orígenes de datos que busca, póngase en contacto con sus administradores para que le concedan los permisos necesarios. Para obtener más información, consulte the section called “Permisos para crear una aplicación que utilice la autenticación del Centro de identidades de IAM (opcional)”.

5. Elija Siguiente y, a continuación, elija Guardar.

Una vez que haya asociado un origen de datos a la aplicación, se habilitará el botón Lanzar aplicación en la página de detalles de la aplicación. Puede elegir Iniciar aplicación para abrir la OpenSearch página de bienvenida, en la que puede crear y gestionar espacios de trabajo.

Para obtener más información sobre cómo trabajar con WorkSpaces, consulte the section called “Uso de los espacios de trabajo OpenSearch de Amazon Service”.

Administrar el acceso a los dominios en una VPC

Si un OpenSearch dominio de una VPC estaba asociado a la aplicación, el administrador de la VPC debe autorizar el acceso entre la interfaz de usuario OpenSearch y la VPC mediante la consola o. AWS CLI

Administrar el acceso a los dominios en una VPC (consola)

Para configurar el acceso a un dominio de VPC mediante Consola de administración de AWS:

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, seleccione Dominios y, a continuación, elija el nombre del dominio de VPC.

   -o bien-

   Seleccione Crear dominio y, a continuación, configure los detalles del dominio.
3. Elija la pestaña Puntos de conexión de VPC y, a continuación, elija Autorizar entidad principal.
4. En el cuadro de diálogo Autorizar a los responsables, selecciona Autorizar a los responsables de otros AWS servicios y, a continuación, selecciona OpenSearch las aplicaciones (panel de control) de la lista.

5. Seleccione Autorizar.

Administrar el acceso a los dominios en una VPC (AWS CLI)

Para autorizar un dominio de VPC mediante el AWS CLI

Para autorizar el dominio de VPC mediante el AWS CLI, ejecute el siguiente comando. Sustituya *placeholder values* por su propia información.

```
aws opensearch authorize-vpc-endpoint-access \
 --domain-name domain-name \
 --service application.opensearchservice.amazonaws.com \
 --region region-id
```

Para revocar una asociación de dominio de VPC mediante la consola

Cuando ya no se necesita una asociación, el propietario del dominio de VPC puede revocar el acceso mediante el siguiente procedimiento.

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, seleccione Dominios y, a continuación, elija el nombre del dominio de VPC.
3. Elija la pestaña de puntos finales de la VPC y, a continuación, seleccione el botón para la fila de OpenSearch aplicaciones (panel de control).
4. Elija Revocar acceso.

Para revocar una asociación de dominio de VPC mediante el AWS CLI

Para revocar una asociación de dominio de VPC con OpenSearch la aplicación de interfaz de usuario, ejecuta el siguiente comando. Sustituya *placeholder values* por su propia información.

```
aws opensearch revoke-vpc-endpoint-access \
    --domain-name domain-name \
    --service application.opensearchservice.amazonaws.com \
    --region region-id
```

Configuración del acceso a las colecciones OpenSearch sin servidor en una VPC

Si una colección de Amazon OpenSearch Serverless en una VPC estaba asociada a la aplicación, el administrador de la VPC puede autorizar el acceso creando una nueva política de red y adjuntándola a la colección.

Configuración del acceso a las colecciones OpenSearch sin servidor en una VPC (consola)

Para configurar el acceso a las colecciones OpenSearch sin servidor en una VPC mediante la consola

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, seleccione Políticas de red, elija el nombre de la política de red y, a continuación, seleccione Editar.

   -o bien-

   Elija Crear política de red y, a continuación, configure los detalles de la política.
3. En el área Tipo de acceso, seleccione Privado (recomendado) y, a continuación, seleccione Acceso privado al servicio de AWS .
4. En el campo de búsqueda, seleccione Servicio y, a continuación, seleccione `application.opensearchservice.amazonaws.com`.
5. En el área Tipo de recurso, selecciona la casilla Habilitar el acceso al OpenSearch punto final.
6. Para buscar colecciones o introducir términos de prefijo específicos, en el campo de búsqueda, seleccione Nombre de la colección y, a continuación, introduzca o seleccione el nombre de las colecciones para asociarlas a la política de red.
7. Seleccione Crear para una nueva política de red o Actualizar para una política de red existente.

Configuración del acceso a las colecciones OpenSearch sin servidor en una AWS CLI VPC ()

Para configurar el acceso a las colecciones OpenSearch sin servidor en una VPC mediante AWS CLI

1. Cree una política con JSON similar a la siguiente. Sustituya *placeholder values* por su propia información.

```
{
    "Description" : "policy-description",
    "Rules": [{
       "ResourceType" : "collection",
        "Resource" : ["collection/collection_name"]
     }],
    "SourceServices" : [
          "application.opensearchservice.amazonaws.com"
      ],
      "AllowFromPublic" : false
}
```

2. Cree o actualice una política de red para que una colección de una VPC funcione con aplicaciones de OpenSearch interfaz de usuario.

   Create a network policy

   Ejecute el comando siguiente. Sustituya *placeholder values* por su propia información.

```
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
```

   El comando devuelve información similar a la siguiente:

```
{
    "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

Ejecute el comando siguiente. Sustituya *placeholder values* por su propia información.

```
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
```

El comando devuelve información similar a la siguiente:

```
{
    "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"
    }
}
```

Uso de los espacios de trabajo OpenSearch de Amazon Service

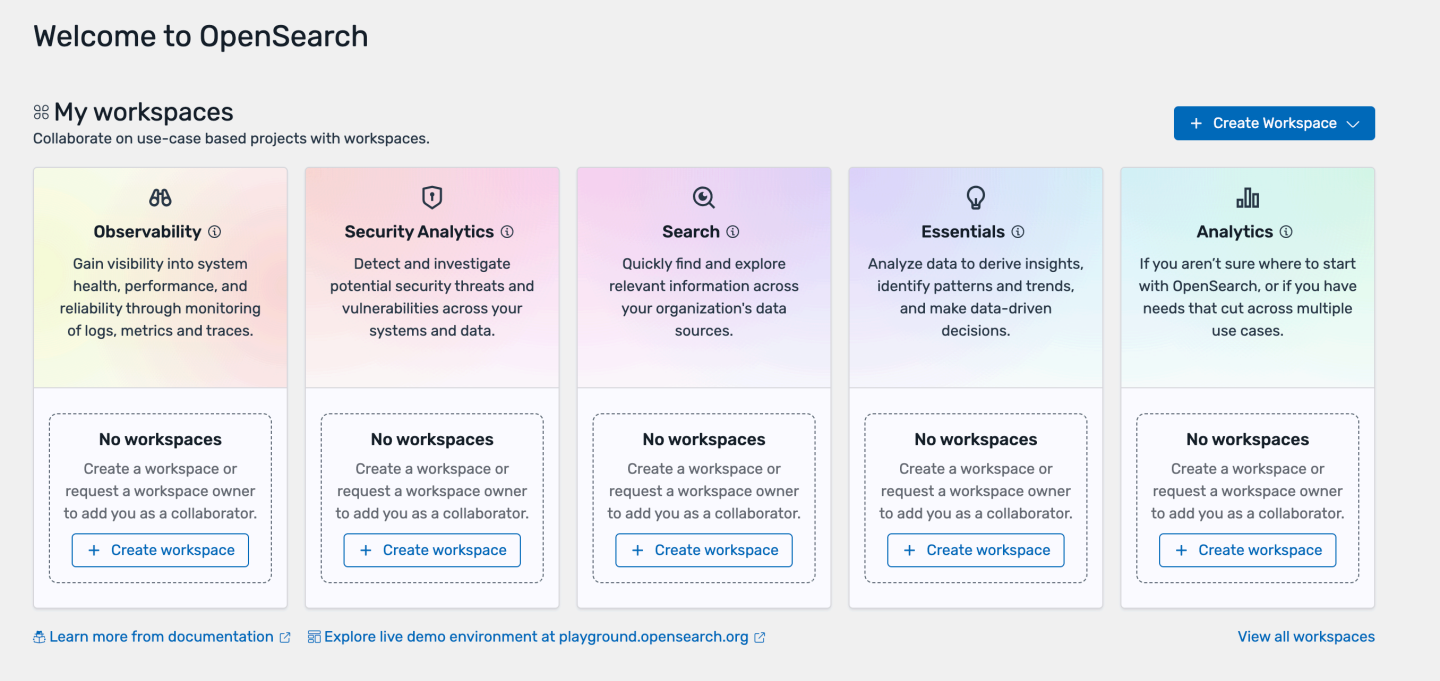
Amazon OpenSearch Service admite la creación de varios espacios de trabajo para casos de uso específicos. Cada espacio de trabajo ofrece una experiencia personalizada para casos de uso populares, como la observabilidad, el análisis de seguridad y la búsqueda. El espacio de trabajo también admite la administración de colaboradores, de modo que puede compartir su espacio de trabajo únicamente con los colaboradores con los que vaya a colaborar y administrar los permisos de cada uno de ellos.

Creación de espacios de trabajo de aplicaciones de OpenSearch interfaz de usuario

Una vez que se haya creado una aplicación de OpenSearch interfaz de usuario y se haya asociado a las fuentes de datos y se hayan configurado los permisos de usuario para la aplicación, puede iniciar la aplicación de OpenSearch interfaz de usuario para crear espacios de trabajo.

Para empezar a crear un espacio de trabajo, puede seleccionar el botón Abrir la aplicación en la página de detalles de la aplicación o utilizar la URL de la aplicación de OpenSearch interfaz de usuario para abrir la página de inicio de la aplicación de OpenSearch interfaz de usuario en una nueva ventana del navegador.

La aplicación de OpenSearch interfaz de usuario ofrece opciones para crear espacios de trabajo y enumera todos los espacios de trabajo existentes en la página de inicio, clasificados por caso de uso.



Para obtener más información sobre los tipos de instancia admitidos, consulte the section called "Tipos de WorkSpace".

La privacidad del espacio de trabajo y los colaboradores

Puede definir una configuración de privacidad para un espacio de trabajo como nivel de permiso predeterminado para todos los usuarios. Puede hacerlo mientras crea un espacio de trabajo o modifica un espacio de trabajo existente (en la pestaña Colaboradores del espacio de trabajo). Existen tres opciones entre las que elegir:

- Privado para los colaboradores: solo los colaboradores que agregue explícitamente al espacio de trabajo pueden acceder al espacio de trabajo. Puede definir niveles de permisos para cada colaborador.

- Cualquiera puede verlo: cualquier persona que tenga acceso a la aplicación de OpenSearch interfaz de usuario puede acceder al espacio de trabajo y ver sus activos, pero no puede realizar ningún cambio en el espacio de trabajo.
- Cualquier persona puede editar: cualquier persona que tenga acceso a la aplicación de OpenSearch interfaz de usuario puede acceder al espacio de trabajo, ver los activos que contiene y realizar cambios en los activos del espacio de trabajo.

En la pestaña Colaboradores del espacio de trabajo, puede añadir usuarios o roles de IAM y AWS IAM Identity Center usuarios como colaboradores en un espacio de trabajo. Existen tres niveles de permisos entre los que elegir los colaboradores:

- Solo lectura: el colaborador solo puede ver los activos del espacio de trabajo. Esta configuración se anula si el espacio de trabajo está configurado para usar la configuración de privacidad Cualquiera puede editar.
- Leer y escribir: el colaborador puede ver y editar los activos del espacio de trabajo. Si el espacio de trabajo está configurado para usar la configuración de privacidad Todos pueden ver, el colaborador podrá seguir editando.
- Administrador: el colaborador puede actualizar la configuración y eliminar el espacio de trabajo. El colaborador también puede cambiar la configuración de privacidad del espacio de trabajo y administrar a los colaboradores. Al usuario que crea el espacio de trabajo se le asigna automáticamente la función de administrador del espacio de trabajo.

Tipos de WorkSpace

Amazon OpenSearch Service ofrece cinco tipos de espacios de trabajo, cada uno con características diferentes para los distintos casos de uso:

- El espacio de trabajo de observabilidad está diseñado para obtener visibilidad del estado, el rendimiento y la fiabilidad del sistema mediante la supervisión de los registros, las métricas y los rastreos.
- El espacio de trabajo de análisis de seguridad está diseñado para detectar e investigar posibles amenazas y vulnerabilidades de seguridad en sus sistemas y datos.
- El espacio de trabajo Search está diseñado para buscar y explorar rápidamente la información relevante en los orígenes de datos de su organización.

- El espacio de trabajo Essentials está diseñado para OpenSearch Serverless como fuente de datos y permite analizar los datos para obtener información, identificar patrones y tendencias y tomar decisiones basadas en los datos con rapidez. Puede buscar y explorar la información relevante de los orígenes de datos de su organización en un espacio de trabajo Essentials.
- El espacio de trabajo de Analytics (todas las funciones) está diseñado para casos de uso polivalentes y admite todas las funciones disponibles en la interfaz de usuario del OpenSearch servicio (paneles).

Acceso a datos entre regiones y cuentas

OpenSearch La interfaz de usuario permite acceder a datos de OpenSearch dominios de diferentes Cuenta de AWS s y Región de AWS s. Puede elegir entre dos enfoques en función de sus necesidades. En la siguiente tabla se comparan los dos enfoques.

> **Note**
>
> Tanto el acceso a los datos entre cuentas como la búsqueda entre clústeres solo funcionan con OpenSearch los dominios. Ninguno de los enfoques admite las recopilaciones OpenSearch sin servidor.

Aspecto	Acceso a los datos entre cuentas	Búsqueda en clústeres
Característica	Asocie dominios de otras cuentas como fuentes de datos directas en OpenSearch la interfaz de usuario	Consulte datos en los dominios conectados mediante conexiones de búsqueda entre clústeres
Mecanismo	Acceso directo: la OpenSearch interfaz de usuario se conecta directamente al dominio de destino de otra cuenta	Acceso indirecto: se requiere un dominio local en la misma cuenta que la OpenSearch interfaz de usuario para retransmitir las solicitudes a dominios remotos
Compatibilidad entre cuentas	Sí	Sí

Aspecto	Acceso a los datos entre cuentas	Búsqueda en clústeres
Compatibilidad entre regiones	No, los dominios de origen y destino deben estar en el mismo Región de AWS	Sí, los dominios de origen y destino pueden estar en Región de AWS s diferentes
Unifique datos entre dominios	No, cada dominio se consulta de forma independiente como una fuente de datos independiente	Sí, una sola consulta puede agregar resultados de varios dominios conectados
Métodos de autenticación	IAM y AWS IAM Identity Center	IAM (con control de acceso detallado)
Complejidad de la configuración	Menor: requiere una función de IAM multicuenta para la validación	Más alto: requiere conexiones entre clústeres, políticas de acceso en ambos dominios y un control de acceso detallado
Visibilidad de la fuente de datos en la interfaz OpenSearch	Cada dominio multicuenta aparece como una fuente de datos independiente	Se accede a los dominios remotos a través de los alias de conexión del dominio de origen local
Acceso de escritura al dominio remoto	Sí, controlado por la política de acceso del dominio de destino	No, la búsqueda entre clústeres proporciona acceso de solo lectura a dominios remotos

Temas

- Acceso a datos multicuenta a dominios OpenSearch
- Búsqueda en clústeres

Acceso a datos multicuenta a dominios OpenSearch

Puede configurar sus aplicaciones de OpenSearch interfaz de usuario en una cuenta para acceder a los OpenSearch dominios de diferentes cuentas. Al crear una aplicación de OpenSearch interfaz de usuario con fuentes de datos multicuenta, se proporciona una `iamRoleForDataSourceArn` que apunta a una función de IAM en la cuenta de destino. OpenSearch La interfaz de usuario valida la

solicitud asumiendo esta función y realizando una llamada `es:DescribeDomain` para verificar la accesibilidad del dominio. La función multicuenta se utiliza únicamente para la validación del plano de control. El acceso al plano de datos se controla por separado mediante la política de acceso del dominio de destino.

> **Código de muestra**
>
> Los ejemplos de código de este tema tienen únicamente fines ilustrativos. Demuestran la funcionalidad básica y es posible que no incluyan la gestión de errores, las mejores prácticas de seguridad ni funciones listas para la producción. Antes de utilizar un código de muestra en producción, revíselo y modifíquelo para que cumpla con sus requisitos específicos y pruébelo exhaustivamente en su entorno.

Conceptos clave

Cuenta de origen

: El Cuenta de AWS que aloja su aplicación de OpenSearch interfaz de usuario.

Cuenta objetivo

: Lugar Cuenta de AWS en el que reside el OpenSearch dominio.

Función multicuenta

: Función de IAM en la cuenta de destino que se utiliza únicamente para la validación del plano de control. Este rol solo requiere el `es:DescribeDomain` permiso.

Función de aplicación de IAM Identity Center

: Función de IAM en la cuenta de origen que se utiliza para acceder al plano de datos de los usuarios del IAM Identity Center.

Requisitos previos

Antes de configurar el acceso a los datos entre cuentas, asegúrese de tener lo siguiente:

- AWS CLI instalado y configurado
- Acceso tanto al origen como al destino Cuenta de AWS
- Para los flujos del IAM Identity Center: una instancia de AWS IAM Identity Center organización

Escenarios

Elija el escenario que coincida con su método de autenticación y configuración de dominio:

- the section called “Escenario 1: un usuario de IAM accede a un dominio público”
- the section called “Escenario 2: un usuario del IAM Identity Center accede a un dominio público”
- the section called “Escenario 3: usuario de IAM que accede a un dominio de VPC”
- the section called “Escenario 4: Un usuario del IAM Identity Center accede a un dominio de VPC”

Escenario 1: un usuario de IAM accede a un dominio público

Paso 1: Crear el rol de IAM multicuenta (cuenta de destino)

Cree un rol de IAM en la cuenta de destino que permita a la cuenta de origen asumirlo para la validación del dominio.

Para crear el rol multicuenta

1. Cree una política de confianza que permita a la cuenta de origen asumir el rol:

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

2. Cree el rol:

```
aws iam create-role \
  --role-name OpenSearchUIAccessRole \
  --assume-role-policy-document file://trust-policy.json
```

3. Crea una política de permisos con solo la siguiente `es:DescribeDomain` acción:

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

```
  "Statement": [{
    "Effect": "Allow",
    "Action": "es:DescribeDomain",
    "Resource": "arn:aws:es:region:target-account-id:domain/*"
  }]
}
```

4. Asocie la política de permisos al rol:

```
aws iam put-role-policy \
  --role-name OpenSearchUIAccessRole \
  --policy-name ValidationOnly \
  --policy-document file://permissions-policy.json
```

Paso 2: Crea el OpenSearch dominio (cuenta de destino)

Cree un OpenSearch dominio en la cuenta de destino con un control de acceso detallado y un cifrado habilitados:

```
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
```

Espere a que el estado del dominio cambie antes de continuar. `Active`

Paso 3: Crear la aplicación de OpenSearch interfaz de usuario (cuenta de origen)

Cree la aplicación en la cuenta de origen con la fuente de datos multicuenta:

```
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":"[\"*
\"]"}]'
```

Paso 4: Verificar y acceder

Recupere los detalles de la aplicación para obtener la URL del punto final:

```
aws opensearch get-application \
  --region region \
  --id application-id
```

- Navegue hasta la URL del punto final de la aplicación desde la respuesta.
- Inicie sesión con las credenciales de IAM.
- El usuario de IAM firma las solicitudes del plano de datos con sus propias credenciales.
- La política de acceso al dominio de destino controla los datos a los que puede acceder el usuario.

Escenario 2: un usuario del IAM Identity Center accede a un dominio público

Paso 1: Crear el rol de IAM multicuenta (cuenta de destino)

Cree un rol de IAM en la cuenta de destino que permita a la cuenta de origen asumirlo para la validación del dominio.

Para crear el rol multicuenta

1. Cree una política de confianza que permita a la cuenta de origen asumir el rol:

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

```
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account-id:root"
    },
    "Action": "sts:AssumeRole"
  }]
}
```

2. Cree el rol:

```
aws iam create-role \
  --role-name OpenSearchUIAccessRole \
  --assume-role-policy-document file://trust-policy.json
```

3. Crea una política de permisos con solo la siguiente `es:DescribeDomain` acción:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": "es:DescribeDomain",
    "Resource": "arn:aws:es:region:target-account-id:domain/*"
  }]
}
```

4. Asocie la política de permisos al rol:

```
aws iam put-role-policy \
  --role-name OpenSearchUIAccessRole \
  --policy-name ValidationOnly \
  --policy-document file://permissions-policy.json
```

Paso 2: Crea el OpenSearch dominio (cuenta de destino)

Crea un OpenSearch dominio en la cuenta de destino. Utilice el mismo comando quethe section called “Paso 2: Crea el OpenSearch dominio (cuenta de destino)”, pero actualice la política de acceso para permitir el rol de aplicación del Centro de Identidad de IAM desde la cuenta de origen:

```
{
  "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/*"
  }]
}
```

Espere a que el estado del dominio cambie `Active` antes de continuar.

Paso 3: Cree la función de IAM para la aplicación IAM Identity Center (cuenta de origen)

Cree una función de IAM en la cuenta de origen que la OpenSearch interfaz de usuario utiliza para acceder al plano de datos de los usuarios del IAM Identity Center.

Para crear el rol de aplicación del Centro de Identidad de IAM

1. Cree una política de confianza:

```
{
  "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. Cree una política de permisos:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "OpenSearchDomain",
    "Effect": "Allow",
    "Action": ["es:ESHttp*"],
    "Resource": "*"
  }]
}
```

3. Cree el rol y adjunte las políticas:

```
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
```

Paso 4: Cree la aplicación de OpenSearch interfaz de usuario con IAM Identity Center (cuenta de origen)

```
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":"[\"*
\"]"}]'
```

Paso 5: Crear y asignar usuarios y grupos del IAM Identity Center

Cree un usuario del Centro de Identidad de IAM

Ejecute el comando siguiente. Sustituya *placeholder values* por su propia información.

```
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
```

Cree un grupo del Centro de identidades de IAM y añada el usuario

Ejecute los siguientes comandos :

```
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
```

Asigne el usuario o el grupo a la aplicación

Use el siguiente comando:

```
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
```

Configure la asignación de roles de backend en el dominio de destino

Asigne el grupo del centro de identidad de IAM a un rol OpenSearch de seguridad en el dominio de destino:

```
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": []
  }'
```

Paso 6: Verificar y acceder

```
aws opensearch get-application \
  --region region \
  --id application-id
```

- Navegue hasta la URL del punto final de la aplicación.
- Inicie sesión con las credenciales de usuario del IAM Identity Center.
- Las solicitudes de datos de los usuarios del IAM Identity Center se firman con el rol de aplicación del Centro de Identidad de IAM, no con el rol multicuenta.
- Las asignaciones de funciones de backend en el dominio controlan los permisos de acceso a los datos.

Escenario 3: usuario de IAM que accede a un dominio de VPC

Paso 1: Crear el rol de IAM multicuenta (cuenta de destino)

Cree un rol de IAM en la cuenta de destino que permita a la cuenta de origen asumirlo para la validación del dominio.

Para crear el rol multicuenta

1. Cree una política de confianza que permita a la cuenta de origen asumir el rol:

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

```
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::source-account-id:root"
    },
    "Action": "sts:AssumeRole"
  }]
}
```

2. Cree el rol:

```
aws iam create-role \
  --role-name OpenSearchUIAccessRole \
  --assume-role-policy-document file://trust-policy.json
```

3. Crea una política de permisos con solo la siguiente `es:DescribeDomain` acción:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": "es:DescribeDomain",
    "Resource": "arn:aws:es:region:target-account-id:domain/*"
  }]
}
```

4. Asocie la política de permisos al rol:

```
aws iam put-role-policy \
  --role-name OpenSearchUIAccessRole \
  --policy-name ValidationOnly \
  --policy-document file://permissions-policy.json
```

Paso 2: Configurar la VPC (cuenta de destino)

Omita este paso si ya existe una VPC en la cuenta de destino.

```
# 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
```

Más información sobre la creación de dominios de VPC.

Paso 3: Crear el dominio de VPC (cuenta de destino)

```
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
```

Espere a que el estado del dominio cambie `Active` antes de continuar.

Paso 4: Autorizar el punto final de la VPC para el principal del servicio de OpenSearch interfaz de usuario (cuenta de destino)

> **Important**
>
> Este es un paso fundamental que es exclusivo de los dominios de VPC. El servicio de OpenSearch interfaz de usuario debe estar autorizado explícitamente para acceder al punto final de la 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
```

Respuesta esperada:

```
{
  "AuthorizedPrincipalList": [
    {
      "PrincipalType": "AWS_SERVICE",
      "Principal": "application.opensearchservice.amazonaws.com"
    }
  ]
}
```

Paso 5: Crear la aplicación de OpenSearch interfaz de usuario (cuenta de origen)

```
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":"[\"*
\"]"}]'
```

Paso 6: Verificar y acceder

Recupere los detalles de la aplicación para obtener la URL del punto final:

```
aws opensearch get-application \
  --region region \
  --id application-id
```

- Navegue hasta la URL del punto final de la aplicación desde la respuesta.
- Inicie sesión con las credenciales de IAM.
- El usuario de IAM firma las solicitudes del plano de datos con sus propias credenciales.
- La política de acceso al dominio de destino controla los datos a los que puede acceder el usuario.

Escenario 4: Un usuario del IAM Identity Center accede a un dominio de VPC

Paso 1: Crear el rol de IAM multicuenta (cuenta de destino)

Cree un rol de IAM en la cuenta de destino que permita a la cuenta de origen asumirlo para la validación del dominio.

Para crear el rol multicuenta

1. Cree una política de confianza que permita a la cuenta de origen asumir el rol:

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

```
}
```

2. Cree el rol:

```
aws iam create-role \
  --role-name OpenSearchUIAccessRole \
  --assume-role-policy-document file://trust-policy.json
```

3. Crea una política de permisos con solo la siguiente `es:DescribeDomain` acción:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": "es:DescribeDomain",
    "Resource": "arn:aws:es:region:target-account-id:domain/*"
  }]
}
```

4. Asocie la política de permisos al rol:

```
aws iam put-role-policy \
  --role-name OpenSearchUIAccessRole \
  --policy-name ValidationOnly \
  --policy-document file://permissions-policy.json
```

Paso 2: Configurar la VPC (cuenta de destino)

Omita este paso si ya existe una VPC en la cuenta de destino.

```
# 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
```

Más información sobre la creación de dominios de VPC.

Paso 3: Crear el dominio de VPC (cuenta de destino)

Utilice el mismo comando quethe section called “Paso 3: Crear el dominio de VPC (cuenta de destino)”, pero actualice la política de acceso para permitir el rol de aplicación de IAM Identity Center desde la cuenta de origen:

```
{
  "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/*"
  }]
}
```

Espere a que el estado del dominio cambie `Active` antes de continuar.

Paso 4: Autorizar el punto final de la VPC para el principal del servicio de OpenSearch interfaz de usuario (cuenta de destino)

> **Important**
>
> Este es un paso fundamental que es exclusivo de los dominios de VPC. El servicio de OpenSearch interfaz de usuario debe estar autorizado explícitamente para acceder al punto final de la 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
```

Respuesta esperada:

```
{
  "AuthorizedPrincipalList": [
    {
      "PrincipalType": "AWS_SERVICE",
      "Principal": "application.opensearchservice.amazonaws.com"
    }
  ]
}
```

Paso 5: Cree la función de IAM para la aplicación IAM Identity Center (cuenta de origen)

Cree una función de IAM en la cuenta de origen que la OpenSearch interfaz de usuario utiliza para acceder al plano de datos de los usuarios del IAM Identity Center.

Para crear el rol de aplicación del Centro de Identidad de IAM

1. Cree una política de confianza:

```
{
  "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. Cree una política de permisos:

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "OpenSearchDomain",
    "Effect": "Allow",
    "Action": ["es:ESHttp*"],
    "Resource": "*"
  }]
}
```

3. Cree el rol y adjunte las políticas:

```
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
```

Paso 6: Cree la aplicación de OpenSearch interfaz de usuario con IAM Identity Center (cuenta de origen)

```
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":"[\"*
\"]"}]'
```

Paso 7: Crear y asignar usuarios y grupos del Centro de Identidad de IAM

Cree un usuario del Centro de Identidad de IAM

Ejecute el comando siguiente. Sustituya *placeholder values* por su propia información.

```
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
```

Cree un grupo del Centro de identidades de IAM y añada el usuario

Ejecute los siguientes comandos :

```
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
```

Asigne el usuario o el grupo a la aplicación

Use el siguiente comando:

```
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
```

Configure la asignación de roles de backend en el dominio de destino

Asigne el grupo del centro de identidad de IAM a un rol OpenSearch de seguridad en el dominio de destino:

```
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": []
  }'
```

Paso 8: Verificar y acceder

```
aws opensearch get-application \
  --region region \
  --id application-id
```

- Navegue hasta la URL del punto final de la aplicación.
- Inicie sesión con las credenciales de usuario del IAM Identity Center.
- Las solicitudes de datos de los usuarios del IAM Identity Center se firman con el rol de aplicación del Centro de Identidad de IAM, no con el rol multicuenta.
- Las asignaciones de funciones de backend en el dominio controlan los permisos de acceso a los datos.

Administración de las aplicaciones de

Actualice una aplicación con fuentes de datos multicuenta

Ejecute el comando siguiente. Sustituya *placeholder values* por su propia información.

```
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**
>
> La operación de actualización reemplaza toda la matriz de fuentes de datos. Incluya todas las fuentes de datos que desee conservar.

Enumerar aplicaciones

Use el siguiente comando:

```
aws opensearch list-applications \
```

```
  --region region
```

Eliminación de una aplicación de

Use el siguiente comando:

```
aws opensearch delete-application \
  --region region \
  --id application-id
```

Revocar el acceso al punto final de la VPC

Use el siguiente comando:

```
aws opensearch revoke-vpc-endpoint-access \
  --domain-name vpc-domain-name \
  --service "application.opensearchservice.amazonaws.com" \
  --region region
```

Referencia rápida

En las siguientes tablas se resumen las principales diferencias entre los tipos de dominio y los métodos de autenticación.

Dominio público comparado con el dominio de VPC

Aspecto	Dominio público	Dominio de VPC
Autorización de puntos finales de VPC	No obligatorio	Obligatorio: debe autorizar `application.opensearchservice.amazonaws.com`
Configuración de la red	Ninguno	VPC, subred y grupo de seguridad con HTTPS (443) entrante
Política de acceso de IAM	Obligatorio	Obligatorio

Aspecto	Dominio público	Dominio de VPC
Función multicuenta	Necesario para el uso de varias cuentas	Necesario para cuentas cruzadas

Un usuario de IAM comparado con un usuario de IAM Identity Center

Aspecto	Usuario de IAM	Usuario de IAM Identity Center
Credenciales del plano de datos	Credenciales de IAM propias del usuario	Función de aplicación del IAM Identity Center
Control de acceso	Política de acceso al dominio	Política de acceso al dominio y asignación de funciones de back-end
Configuración adicional	Ninguno	Rol de aplicación de IAM Identity Center, user/group creación, asignación de aplicaciones, mapeo de roles de back-end
OpenSearch Configuración de aplicaciones de interfaz de usuario	No hay opciones de IAM Identity Center	`--iam-identity-center-options` obligatorio

Notas importantes

- `iamRoleForDataSourceArn`Debe estar en la misma cuenta que. `dataSourceArn`
- Solo `iamRoleForDataSourceArn` es obligatorio para las fuentes de datos entre cuentas. Omita esta opción para las fuentes de datos de la misma cuenta.
- La función multicuenta solo necesita el permiso. `es:DescribeDomain` Nunca se usa para acceder al plano de datos.
- Para los dominios de VPC, se deben configurar tanto la política de IAM como la autorización de puntos de conexión de VPC.
- Versiones de motor compatibles: OpenSearch 1.3 y superiores.

Resolución de problemas

Problema	Resolución
La creación de la aplicación falla y aparece el mensaje «No se puede acceder al dominio»	Compruebe que la función multicuenta tiene el `es:DescribeDomain` permiso y que la política de confianza permite la cuenta de origen.
Se produce un error en la asociación de dominios de VPC	Asegúrese de que el punto final de la VPC esté autorizado para. `application.opensearchservice.amazonaws.com`
Se ha denegado el acceso al plano de datos al usuario de IAM	Compruebe que la política de acceso al dominio de destino permita al usuario o rol principal de IAM.
Se ha denegado el acceso al plano de datos para el usuario del IAM Identity Center	Compruebe que la asignación de funciones de backend incluya el ID de grupo del IAM Identity Center y que la política de dominio permita la función de aplicación del IAM Identity Center.
Error de discordancia de cuentas	Asegúrese de que `iamRoleForDataSourceArn` está en la misma cuenta que el dominio. `dataSourceArn`

Búsqueda en clústeres

Con la búsqueda entre clústeres en Amazon OpenSearch Serverless, puede realizar consultas y agregaciones en varios dominios conectados.

La búsqueda entre clústeres en Amazon OpenSearch Serverless utiliza los conceptos de dominio de origen y dominio de destino. Una petición de búsqueda en clústeres se origina en un dominio de origen. El dominio de destino puede estar en un dominio diferente Cuenta de AWS o Región de AWS (o en ambos) del dominio de origen desde el que se realiza la consulta. Mediante la búsqueda entre clústeres, puedes configurar un dominio de origen para asociarlo a tu OpenSearch interfaz de usuario en la misma cuenta y, a continuación, crear conexiones a los dominios de destino. Como resultado, puedes usar la OpenSearch interfaz de usuario con datos de los dominios de destino aunque estén en una cuenta o región diferente.

Usted paga los cargos de transferencia de AWS datos estándar por los datos transferidos dentro y fuera de Amazon OpenSearch Service. No se le cobrará por los datos transferidos entre los nodos de su dominio OpenSearch de Servicio. Para obtener más información acerca de los cargos de entrada y salida, consulte Transferencia de datos en la página Precios bajo demanda de Amazon EC2.

Puedes usar la búsqueda entre clústeres como mecanismo para que tu OpenSearch interfaz de usuario se asocie a clústeres de una cuenta diferente o de una región diferente. Las solicitudes entre dominios se cifran en tránsito de forma predeterminada como parte del node-to-node cifrado.

> **Note**
>
> La OpenSearch herramienta de código abierto también documenta la búsqueda entre clústeres. Tenga en cuenta que la configuración de la herramienta de código abierto difiere significativamente para los clústeres de código abierto en comparación con los dominios gestionados de Amazon OpenSearch Serverless.
> En particular, en Amazon OpenSearch Serverless, las conexiones entre clústeres se configuran mediante las solicitudes Consola de administración de AWS en lugar de mediante `cURL` solicitudes. Además, el servicio administrado utiliza AWS Identity and Access Management (IAM) para la autenticación entre clústeres, además de un control de acceso detallado.
> Por lo tanto, le recomendamos que utilice el contenido de este tema para configurar la búsqueda entre clústeres para sus dominios en lugar de utilizar la documentación de código abierto. OpenSearch

Diferencias funcionales al utilizar la búsqueda entre clústeres

En comparación con los dominios normales, los dominios de destino creados mediante la búsqueda entre clústeres tienen las siguientes diferencias y requisitos funcionales:

- No puede escribir ni ejecutar comandos `PUT` en el clúster remoto. El acceso al clúster remoto es solo de lectura.
- Tanto el dominio de origen como el de destino deben ser OpenSearch dominios. No puedes conectar un dominio de Elasticsearch ni clústeres de OpenSearch /Elasticsearch autogestionados para la interfaz de usuario. OpenSearch
- Un dominio puede tener un máximo de 20 conexiones a otros dominios. Esto incluye las conexiones entrantes y salientes.

- El dominio de origen debe estar en la misma versión o en una versión superior a la del dominio de destino. OpenSearch Si desea configurar conexiones bidireccionales entre dos dominios, los dos dominios deben estar en la misma versión. Recomendamos actualizar ambos dominios a la última versión antes de realizar la conexión. Si necesita actualizar los dominios después de configurar la conexión bidireccional, primero debe eliminar la conexión y, a continuación, volver a crearla.
- No puede usar diccionarios personalizados o SQL con la búsqueda en clústeres.
- No se puede usar CloudFormation para conectar dominios.
- No se puede utilizar la búsqueda entre clústeres en instancias M3 o bursátiles (T2 y T3).
- La búsqueda entre clústeres no funciona en las colecciones de Amazon OpenSearch Serverless.

Requisitos previos de búsqueda entre clústeres para la interfaz de usuario OpenSearch

Antes de configurar la búsqueda entre clústeres con dos OpenSearch dominios, asegúrate de que tus dominios cumplen los siguientes requisitos:

- Control de acceso detallado habilitado para ambos dominios
- Node-to-node el cifrado está habilitado para ambos dominios

Temas

- Configuración de permisos de acceso para el acceso a datos entre regiones y cuentas mediante la búsqueda entre clústeres
- Crear una conexión entre dominios
- Probar su configuración de seguridad para el acceso a datos entre cuentas y regiones con búsqueda entre clústeres
- Eliminación de una conexión

Configuración de permisos de acceso para el acceso a datos entre regiones y cuentas mediante la búsqueda entre clústeres

Al enviar una petición de búsqueda en clústeres al dominio de origen, el dominio evalúa esa petición en comparación con la política de acceso al dominio. La búsqueda en clústeres requiere un control de acceso preciso. A continuación, se muestra un ejemplo con una política de acceso abierto en el dominio de origen.

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**
>
> Si incluye índices remotos en la ruta, debe codificar la URL del URI en el ARN del dominio. Por ejemplo, use el siguiente formato de ARN:
> `:arn:aws:es:us-east-1:111222333444:domain/my-domain/local_index,dst%3Aremote_index`
> No use el siguiente formato ARN.
> `arn:aws:es:us-east-1:111222333444:domain/my-domain/local_index,dst:remote_index.`

Si elige utilizar una política de acceso restrictivo además de un control de acceso detallado, la política debe permitir el acceso a `es:ESHttpGet` como mínimo. A continuación, se muestra un ejemplo:

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/*"
    }
  ]
}
```

El control de acceso detallado del dominio de origen evalúa la solicitud para determinar si está firmada con credenciales básicas de IAM o HTTP válidas. Si es así, el control de acceso detallado evalúa a continuación si el usuario tiene permiso para realizar la búsqueda y acceder a los datos.

Los siguientes son los requisitos de permiso para las búsquedas:

- Si la solicitud solo busca datos en el dominio de destino (por ejemplo, `dest-alias:dest-index/_search`)), solo necesitará permisos en el dominio de destino.
- Si la solicitud busca datos en ambos dominios (por ejemplo, `source-index,dest-alias:dest-index/_search`)), necesita permisos en ambos dominios.
- En el control de acceso detallado, los usuarios deben tener el permiso `indices:admin/shards/search_shards` además de los permisos de lectura o estándar para los índices pertinentes.

El dominio de origen pasa la solicitud al dominio de destino. El dominio de destino evalúa esta solicitud frente a su política de acceso al dominio. Para admitir todas las funciones de la OpenSearch interfaz de usuario, como la indexación de documentos y la realización de búsquedas estándar, se deben configurar todos los permisos. A continuación, se muestra un ejemplo de nuestra política recomendada en el dominio de destino:

JSON

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

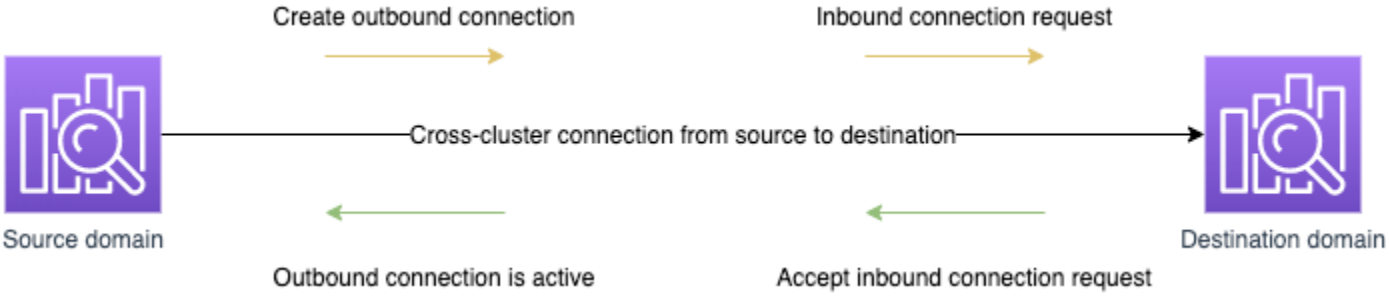
```
 "Statement": [
    {
       "Effect": "Allow",
       "Principal": {
         "AWS": [
           "*"
         ]
      },
      "Action": [
        "es:ESHttp*"
      ],
      "Resource": "arn:aws:es:us-east-2:111222333444:domain/my-destination-
domain/*"
    },
    {
    "Effect": "Allow",
        "Principal": {
    "AWS": "*"
      },
      "Action": "es:ESCrossClusterGet",
      "Resource": "arn:aws:es:us-east-2:111222333444:domain/"
    }
  ]
}
```

Si solo desea realizar búsquedas básicas, el requisito mínimo de la política es que el permiso `es:ESCrossClusterGet` se solicite para el dominio de destino sin el uso de caracteres comodín. Por ejemplo, en la política anterior, especificaría el nombre de dominio como */my-destination-domain* y no*/my-destination-domain/**.

En ese caso, el dominio de destino realiza la búsqueda y devuelve los resultados al dominio de origen. El dominio de origen combina sus propios resultados (si los hay) con los resultados del dominio de destino y los devuelve.

Crear una conexión entre dominios

Una conexión entre clústeres es unidireccional desde el dominio de origen hasta el dominio de destino. Esto significa que los dominios de destino (en una cuenta o región diferente) no pueden consultar el dominio de origen, que es local en la OpenSearch interfaz de usuario. El dominio de origen crea una conexión outbound al dominio de destino. El dominio de destino recibe una solicitud de conexión inbound del dominio de origen.



Para crear una conexión entre dominios

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, seleccione Dominios.
3. Elija el nombre de un dominio que sirva como dominio de origen y, a continuación, elija la pestaña Conexiones.
4. En la sección Conexiones de salida, seleccione Solicitar.
5. En Alias de conexión, ingrese un nombre para la conexión. El alias de conexión se usa en la OpenSearch interfaz de usuario para seleccionar los dominios de destino.
6. En el modo de conexión, elija Directo para realizar búsquedas o replicaciones entre clústeres.
7. Para especificar que la conexión debe omitir los clústeres no disponibles durante una búsqueda, active la casilla Omitir los clústeres no disponibles. Esta configuración garantiza que las consultas entre clústeres devuelvan resultados parciales a pesar de que se produzcan errores en uno o más clústeres remotos.
8. En Clúster de destino, elija entre Conectarse a un clúster de esta Cuenta de AWS o Conectarse a un clúster de otra Cuenta de AWS.
9. En el ARN de dominio remoto, escriba el nombre de recurso de Amazon (ARN) para el clúster. El ARN del dominio se encuentra en el área de información general de la página de detalles del dominio.

   El dominio debe cumplir los siguientes requisitos:

   - El formato del ARN debe ser `arn:`*`partition`*`:es:`*`regionaccount-id`*`:`*`type`*`/`*`domain-id`*. Por ejemplo:

     ```
     arn:aws:es:us-east-2:111222333444:domain/my-domain
     ```
   - El dominio debe configurarse para usar la OpenSearch versión 1.0 (o posterior) o la versión 6.7 de Elasticsearch (o posterior).

- Debe estar habilitado el control de acceso preciso en el dominio.
- El dominio debe estar en ejecución. OpenSearch

10. Seleccione Solicitar.

La búsqueda entre clústeres valida primero la solicitud de conexión para asegurarse de que se cumplen los requisitos previos. Si los dominios son incompatibles, la solicitud de conexión entra en el estado `Validation failed`.

Una vez correctamente validada la solicitud de conexión, se envía al dominio de destino, donde tiene que aprobarse. Mientras no se obtenga la aprobación, la conexión permanecerá en el estado de `Pending acceptance`. Cuando se acepta la solicitud de conexión en el dominio de destino, el estado cambia a `Active` y el dominio de destino se vuelve disponible para consultas.

La página de dominio muestra el estado general del dominio y los detalles del estado de la instancia del dominio de destino. Los propietarios de dominios tienen la flexibilidad de crear, visualizar, eliminar y monitorear conexiones desde o hacia sus dominios.

Una vez establecida la conexión, se cifra todo el tráfico que circule entre los nodos de los dominios conectados. Si conecta un dominio de VPC a un dominio que no es VPC y el dominio que no es VPC es un punto de conexión público que puede recibir tráfico de Internet, el tráfico entre clúster entre los dominios sigue cifrado y seguro.

Probar su configuración de seguridad para el acceso a datos entre cuentas y regiones con búsqueda entre clústeres

Una vez que haya configurado los permisos de acceso a los datos entre regiones y cuentas mediante la búsqueda entre clústeres, le recomendamos que pruebe la configuración con Postman, una plataforma de terceros para el desarrollo colaborativo de las API.

Para configurar la configuración de seguridad mediante Postman

1. En el dominio de destino, indexe un documento: Lo que sigue es un ejemplo de resultado.

```
POST https://dst-domain.us-east-1.es.amazonaws.com/books/_doc/1
{
"Dracula": "Bram Stoker"
}
```

2. Para consultar este índice desde el dominio de origen, incluya el alias de conexión del dominio de destino dentro de la consulta. Puede encontrar el alias de conexión en la pestaña Conexiones del panel de control del dominio. A continuación, se muestra un ejemplo de solicitud y una respuesta truncada:

```
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. (Opcional) Puede crear una configuración que incluya varios dominios en una sola búsqueda. Por ejemplo, supongamos que tiene lo siguiente:

   Una conexión entre `domain-a` a `domain-b`, con un alias de conexión denominado `cluster_b`

   Una conexión entre `domain-a` a `domain-c`, con un alias de conexión denominado `cluster_c`

   En este caso, las búsquedas incluyen el contenido `domain-a`, `domain-b` y `domain-c`. A continuación se muestra un ejemplo de una solicitud y la respuesta:

   Solicitud

```
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"
}
  }
}
```

Respuesta:


```
{
"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
    }
  }
]
  }
}
```

Si no eligió omitir los clústeres no disponibles en la configuración de la conexión, todos los clústeres de destino que busca tienen que estar disponibles para que la petición de búsqueda se ejecute correctamente. De lo contrario, se produce un error en toda la solicitud; incluso si uno de los dominios no está disponible, la búsqueda no devuelve ningún resultado.

Eliminación de una conexión

La eliminación de una conexión detiene cualquier operación de búsqueda entre clústeres en el dominio de destino.

Puede realizar estos pasos en el dominio de origen o de destino para eliminar la conexión. Una vez quitada la conexión, seguirá visible con el estado `Deleted` durante un periodo de 15 días.

No se puede eliminar un dominio con conexiones activas entre clústeres. Para eliminar un dominio, primero quite todas las conexiones entrantes y salientes de ese dominio. Esto es para asegurarse de tener en cuenta a los usuarios del dominio entre clústeres antes de eliminar el dominio.

Para eliminar una conexión

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, seleccione Dominios.
3. Elija el nombre de un dominio y, a continuación, la pestaña Contenido.
4. El nombre de la conexión que se va a eliminar.
5. Elija Eliminar y, a continuación, confirme la eliminación.

Administrar el acceso a la OpenSearch interfaz de usuario desde un punto final de VPC

Puede crear una conexión privada entre la VPC y la OpenSearch interfaz de usuario mediante. AWS PrivateLink Con esta conexión, puede acceder a las aplicaciones de OpenSearch interfaz de usuario como si estuvieran en la misma VPC. De esta forma, no es necesario configurar una puerta de enlace a Internet, un dispositivo NAT o una conexión VPN ni AWS Direct Connect establecer la conexión. Las instancias de su VPC no necesitan direcciones IP públicas para acceder OpenSearch a la interfaz de usuario.

Para establecer esta conexión privada, primero debe crear un punto final de interfaz con tecnología. AWS PrivateLink Se crea una interfaz de red de punto de conexión en cada subred especificada para el punto de conexión de interfaz. Se trata de interfaces de red administradas por el solicitante que sirven como punto de entrada para el tráfico destinado a las aplicaciones de interfaz de usuario OpenSearch .

Crear una conexión privada entre una VPC y una interfaz de usuario OpenSearch

Puede crear una conexión privada para acceder a la OpenSearch interfaz de usuario desde una VPC mediante o. Consola de administración de AWS AWS CLI

Crear una conexión privada entre una VPC y una OpenSearch interfaz de usuario (consola)

Para crear una conexión privada entre una VPC y la OpenSearch interfaz de usuario mediante la consola

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación de la izquierda, amplíe Sin servidor y seleccione Puntos de conexión de VPC.
3. Seleccione Crear punto de conexión de VPC.
4. En Nombre, ingrese un nombre para el punto de conexión.
5. En el caso de la VPC, seleccione la VPC desde la que accederá a OpenSearch las aplicaciones de interfaz de usuario.

6. En el caso de las subredes, selecciona una subred desde la que accederás a OpenSearch las aplicaciones de la interfaz de usuario.

> **Note**
>
> La dirección IP y el tipo de DNS del punto de conexión se basan en el tipo de subred:
>
> - Pila doble: si todas las subredes tienen ambos IPv4 rangos de direcciones. IPv6
> - IPv6: Si todas las subredes son IPv6 solo subredes.
> - IPv4: Si todas las subredes tienen IPv4 rangos de direcciones.

7. En Grupos de seguridad, seleccione el grupo de seguridad que se va a asociar con las interfaces de red de punto de conexión de VPC.

> **Note**
>
> Este es un paso fundamental en el que limita los puertos, protocolos y orígenes para el tráfico entrante que autoriza para el punto de conexión. Asegúrese de que las reglas del grupo de seguridad permitan que los recursos que utilizarán el punto final de la VPC para comunicarse con las aplicaciones de la OpenSearch interfaz de usuario también se comuniquen con la interfaz de red del punto final.

8. 8. Seleccione Crear punto de conexión.

Crear una conexión privada entre una VPC y una OpenSearch interfaz de usuario ()AWS CLI

Para crear una conexión privada entre una VPC y una OpenSearch interfaz de usuario mediante AWS CLI

Ejecute el comando siguiente. Sustituya *placeholder values* por su propia información.

```
aws opensearchserverless create-vpc-endpoint \
    --region region \
    --endpoint endpoint \
    --name vpc_endpoint_name \
    --vpc-id vpc_id \
    --subnet-ids subnet_ids
```

Actualización de la política de puntos finales de la VPC para permitir el acceso a la aplicación de interfaz de usuario OpenSearch

Después de crear la conexión privada, actualice la política de puntos finales de la VPC para permitir el acceso a la aplicación de OpenSearch interfaz de usuario en la política de puntos finales de la VPC especificando el ID de la aplicación.

Para obtener más información sobre cómo actualizar la política de punto de conexión de VPC, consulte Actualización de una política de punto de conexión de VPC en la Guía de AWS PrivateLink .

Asegúrese de que la política de punto de conexión de VPC incluya la siguiente declaración. Sustituya *placeholder value* por su propia información.

```
{
    "Statement": [{
        "Action": ["opensearch:*"],
        "Effect": "Allow",
        "Principal": "*",
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "opensearch:ApplicationId": ["opensearch-ui-application-id"]
            }
        }
    }]
}
```

Revocar el acceso a la OpenSearch interfaz de usuario en una política de puntos finales de VPC

OpenSearch La interfaz de usuario requiere un permiso explícito en la política de puntos finales de la VPC para permitir a los usuarios acceder a la aplicación desde la VPC. Si ya no desea que los usuarios accedan a la OpenSearch interfaz de usuario desde la VPC, puede eliminar el permiso en la política de puntos finales. Después de esto, los usuarios reciben un mensaje de `403 forbidden` error al intentar acceder a la OpenSearch interfaz de usuario.

Para obtener más información sobre cómo actualizar la política de punto de conexión de VPC, consulte Actualización de una política de punto de conexión de VPC en la Guía de AWS PrivateLink .

El siguiente es un ejemplo de una política de punto de conexión de VPC que deniega el acceso a las aplicaciones de la interfaz de usuario desde la VPC:

```
{
    "Statement": [{
        "Action": ["opensearch:*"],
        "Effect": "Allow",
        "Principal": "*",
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "opensearch:ApplicationId": [""]
            }
        }
    }]
}
```

Migración de objetos guardados de los OpenSearch paneles a la interfaz de usuario OpenSearch

Si tienes paneles, visualizaciones, patrones de indexación y otros objetos guardados en los OpenSearch paneles, puedes migrarlos y reutilizarlos en la interfaz de usuario. OpenSearch

Ventajas de migrar a la interfaz de usuario: OpenSearch

- Alta disponibilidad: la OpenSearch interfaz de usuario está alojada en el dominio Nube de AWS y permanece disponible durante las actualizaciones y el mantenimiento, mientras que OpenSearch Dashboards se aloja en el dominio y no estará disponible temporalmente.
- Varias fuentes de datos: la OpenSearch interfaz de usuario puede proporcionar un único panel de control consolidado para múltiples fuentes de datos, incluidos OpenSearch dominios, colecciones sin servidor y conexiones de datos con Amazon S3 y Amazon CloudWatch; mientras que cada OpenSearch panel solo puede funcionar con un dominio o colección.
- En la interfaz de usuario están disponibles funciones adicionales, como el asistente de inteligencia artificial y los espacios de trabajo. OpenSearch Más información:OpenSearch INTERFAZ DE USUARIO.

Temas

- Información general sobre la migración

- Requisitos previos
- Paso 1: Exporte los objetos guardados desde los OpenSearch paneles
- Paso 2: Importa los objetos guardados a la interfaz de usuario OpenSearch

Información general sobre la migración

El proceso de migración consta de los siguientes pasos de alto nivel:

1. Exportación de objetos guardados desde OpenSearch paneles: utilice la interfaz de usuario de administración de objetos guardados de OpenSearch Dashboards o la API de exportación para descargar sus paneles, visualizaciones, patrones de indexación y otros objetos como un archivo JSON (NDJSON) delimitado por líneas nuevas.
2. Cree una aplicación de OpenSearch interfaz de usuario y un espacio de trabajo: si aún no lo ha hecho, cree una aplicación de OpenSearch interfaz de usuario y un espacio de trabajo para recibir los objetos importados.
3. Registre la fuente de datos en la OpenSearch interfaz de usuario: asocie su OpenSearch dominio a la aplicación de OpenSearch interfaz de usuario y regístrela como fuente de datos en el espacio de trabajo. Los patrones de índice de los objetos importados hacen referencia a esta fuente de datos.
4. Importación de objetos guardados a la OpenSearch interfaz de usuario: utilice la OpenSearch interfaz de usuario de administración de objetos guardados o la API de importación para cargar el archivo NDJSON en el espacio de trabajo de destino.
5. Valide los objetos importados: abra los paneles y las visualizaciones en la OpenSearch interfaz de usuario para comprobar que se muestran correctamente y que los datos provienen del dominio o la colección asociados.

Requisitos previos

Antes de la migración, compruebe los siguientes requisitos previos:

- Dispones de los permisos de IAM necesarios para llamar al OpenSearch Servicio de Amazon y OpenSearch APIs. Para obtener más información, consulte the section called “Permisos necesarios para crear aplicaciones OpenSearch de Amazon Service”.
- Puede acceder al dominio o la colección y a los OpenSearch paneles de control que desee migrar.

- Ha creado una aplicación de OpenSearch interfaz de usuario. Para obtener información sobre la creación de una aplicación y un espacio de trabajo, consultethe section called “Introducción”.
- Ha asociado el mismo dominio o colección a la aplicación de OpenSearch interfaz de usuario. Para obtener información sobre la asociación de fuentes de datos, consultethe section called “Administración de las asociaciones de orígenes de datos y los permisos de acceso a la VPC”.

> **Note**
>
> OpenSearch La interfaz de usuario solo es compatible con OpenSearch la versión 1.3 y posteriores. Comprueba que tu OpenSearch dominio ejecute la versión 1.3 o superior antes de intentar migrar los objetos guardados.

Paso 1: Exporte los objetos guardados desde los OpenSearch paneles

Exporte los objetos guardados desde los OpenSearch paneles de control mediante la interfaz de usuario de administración o la API de exportación. La exportación genera un archivo JSON (NDJSON) delimitado por líneas nuevas que contiene todos los tipos de objetos guardados seleccionados y sus dependencias.

Temas

- OpenSearch Exporte manualmente a los paneles
- Exportación mediante API

OpenSearch Exporte manualmente a los paneles

Para exportar objetos guardados mediante la interfaz de usuario de administración de OpenSearch paneles

1. Abre tu instancia de OpenSearch Dashboards.
2. En el panel de navegación izquierdo, selecciona Administración.
3. En Administración de paneles, seleccione Objetos guardados.
4. Seleccione los objetos guardados que desee exportar. Para exportar todos los objetos de un tipo específico, filtre por tipo mediante la barra de búsqueda. Para exportar todos los objetos, active la casilla de verificación del encabezado de la tabla.

5. Seleccione Exportar.
6. En el cuadro de diálogo Exportar objetos guardados, asegúrese de que esté seleccionada la opción Incluir objetos relacionados. Esta opción incluye todos los objetos de los que dependen los objetos guardados seleccionados, como los patrones de índice a los que hacen referencia las visualizaciones. Desactive esta opción solo si pretende gestionar las dependencias manualmente.
7. Seleccione Exportar para descargar el `.ndjson` archivo a su máquina local.

> **Tip**
>
> Si selecciona Incluir objetos relacionados, el archivo NDJSON exportado contiene todos los objetos guardados necesarios para renderizar los cuadros de mando y las visualizaciones seleccionados, incluidos los patrones de índice, las visualizaciones y los objetos de búsqueda que dependen de ellos. Esto simplifica el paso de importación y evita que falten errores de referencia.

Exportación mediante API

Puede utilizar la API de exportación de objetos guardados de OpenSearch Dashboards para exportar los objetos guardados mediante programación. Esto resulta útil para automatizar las migraciones o integrar el paso de exportación en una canalización. CI/CD

> **Note**
>
> Si tu OpenSearch dominio tiene habilitado un control de acceso detallado, debes proporcionar las credenciales de autenticación con tu solicitud de exportación. Usa la autenticación básica HTTP añadiendo la `-u` marca con tu nombre de usuario y contraseña. Para obtener más información acerca de las opciones de autenticación, consulte the section called “Control de acceso detallado”.

En el siguiente ejemplo, se exportan todos los paneles con sus objetos relacionados. Sustituya *`placeholder values`* por su propia información.

```
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
```

Si su dominio no tiene activado el control de acceso detallado, puede omitir la marca. `-u`

Para exportar objetos guardados específicos por ID, usa el parámetro en lugar de`objects`: `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

Para buscar un objeto guardado IDs, puedes usar la API de objetos guardados para enumerar todos los objetos de un tipo específico. En el siguiente ejemplo, se enumeran todos los paneles:

```
curl -X GET \
    "https://dashboards-endpoint/_dashboards/api/saved_objects/_find?
type=dashboard" \
    -u 'master-username:master-password'
```

La respuesta incluye el ID de cada objeto guardado. También puede encontrar el ID en la URL del navegador al ver el objeto en los OpenSearch paneles.

Paso 2: Importa los objetos guardados a la interfaz de usuario OpenSearch

Después de exportar los objetos guardados, puede importar el archivo NDJSON a la OpenSearch interfaz de usuario manualmente o mediante la API.

Temas

- Importe manualmente en la interfaz de usuario OpenSearch
- Importación mediante API

Importe manualmente en la interfaz de usuario OpenSearch

Para importar objetos guardados mediante la OpenSearch interfaz de usuario de administración de la interfaz de usuario

1. Abre la aplicación de OpenSearch interfaz de usuario y navega hasta el espacio de trabajo de destino.
2. En el espacio de trabajo, selecciona Activos en la barra de navegación superior o ve a la página de activos del espacio de trabajo.
3. Seleccione Importar para abrir el cuadro de diálogo Importar activos.
4. Elija Seleccionar archivo y seleccione el `.ndjson` archivo que ha exportado desde los OpenSearch paneles.
5. Para la gestión de conflictos, elija una de las siguientes opciones:
   - Crear nuevos activos con elementos únicos IDs (predeterminado): genera nuevos IDs para todos los objetos importados, lo que evita conflictos con los activos existentes.
   - Comprobar si hay activos existentes: comprueba si hay conflictos con los objetos existentes. Cuando esté seleccionada, elija una de las siguientes subopciones:
     - Sobrescribe automáticamente los conflictos: los activos existentes con el mismo ID se sustituyen automáticamente.
     - Solicita que se tomen medidas en caso de conflicto: se te pedirá que resuelvas cada conflicto de forma individual.
6. Seleccione Importar.
7. Revise el resumen de la importación.

Importación mediante API

Para importar objetos guardados mediante la API con la autenticación AWS Signature versión 4, primero debe obtener el ID de la fuente de datos y, a continuación, utilizarlo en la solicitud de importación. Sustituya *placeholder values* por su propia información.

Paso 1: Obtenga el ID de la fuente de datos de su espacio de trabajo:

```
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 respuesta incluye el ID de la fuente de datos. También puede encontrar el ID de la fuente de datos en la URL del navegador al ver la fuente de datos en la OpenSearch interfaz de usuario.

Paso 2: importe los objetos guardados utilizando el ID de la fuente de datos del paso 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**
>
> Estos ejemplos utilizan la `--aws-sigv4` opción integrada de curl (disponible en curl 7.75 o versiones posteriores) para firmar las solicitudes. Configura tus AWS credenciales como variables de entorno antes de ejecutar los comandos:

`AWS_ACCESS_KEY_IDAWS_SECRET_ACCESS_KEY`, y `AWS_SESSION_TOKEN` (si utilizas credenciales temporales).

OpenSearch Puntos finales y cuotas de la interfaz de usuario

Para conectarse mediante programación a un AWS servicio, se utiliza un punto final. Las cuotas de servicio, que también se denominan límites, establecen el número máximo de recursos u operaciones de servicio que puede haber en una cuenta de AWS .

Amazon OpenSearch UI es la interfaz de OpenSearch usuario de próxima generación para OpenSearch paneles. Proporciona puntos de conexión para acceder a sus OpenSearch paneles de control. Utilice este tema para encontrar los puntos finales del servicio y las cuotas de servicio para la interfaz de usuario. OpenSearch

Para obtener más información sobre otros OpenSearch servicios, consulte Cuotas y puntos finales del servicio.

OpenSearch Puntos finales de interfaz de usuario

OpenSearch La interfaz de usuario está disponible en las siguientes regiones:

Nombre de la región	Región	Punto de conexión	Protocolo
Asia-Pacífico (Mumbai)	ap-south-1	opensearch.ap-south-1.amazonaws.com es.ap-south-1.amazonaws.com	HTTPS HTTPS
Europa (París)	eu-west-3	opensearch.eu-west-3.amazonaws.com es.eu-west-3.amazonaws.com	HTTPS HTTPS
Este de EE. UU. (Ohio)	us-east-2	opensearch.us-east-2.amazonaws.com es.us-east-2.amazonaws.com	HTTPS HTTPS

Nombre de la región	Región	Punto de conexión	Protocolo
Europa (Irlanda)	eu-west-1	opensearch.eu-west-1.amazonaws.com es.eu-west-1.amazonaws.com	HTTPS HTTPS
Europa (Fráncfort)	eu-central-1	opensearch.eu-central-1.amazonaws.com es.eu-central-1.amazonaws.com	HTTPS HTTPS
América del Sur (São Paulo)	sa-east-1	opensearch.sa-east-1.amazonaws.com es.sa-east-1.amazonaws.com	HTTPS HTTPS
Este de EE. UU. (Norte de Virginia)	us-east-1	opensearch.us-east-1.amazonaws.com es.us-east-1.amazonaws.com	HTTPS HTTPS
Europa (Londres)	eu-west-2	opensearch.eu-west-2.amazonaws.com es.eu-west-2.amazonaws.com	HTTPS HTTPS
Asia-Pacífico (Tokio)	ap-northeast-1	zocalo.ap-northeast-1.amazonaws.com es.ap-northeast-1.amazonaws.com	HTTPS HTTPS
Oeste de EE. UU. (Oregón)	us-west-2	opensearch.us-west-2.amazonaws.com es.us-west-2.amazonaws.com	HTTPS HTTPS

Nombre de la región	Región	Punto de conexión	Protocolo
Asia-Pacífico (Singapur)	ap-southeast-1	opensearch.ap-southeast-1.amazonaws.com es.ap-southeast-1.amazonaws.com	HTTPS HTTPS
Asia-Pacífico (Sídney)	ap-southeast-2	opensearch.ap-southeast-2.amazonaws.com es.ap-southeast-2.amazonaws.com	HTTPS HTTPS
Canadá (centro)	ca-central-1	opensearch.ca-central-1.amazonaws.com es.ca-central-1.amazonaws.com	HTTPS HTTPS
Europa (Estocolmo)	eu-north-1	opensearch.eu-north-1.amazonaws.com es.eu-north-1.amazonaws.com	HTTPS HTTPS
Asia-Pacífico (Hong Kong)	ap-east-1	opensearch.ap-east-1.amazonaws.com es.ap-east-1.amazonaws.com	HTTPS HTTPS
Asia-Pacífico (Seúl)	ap-northeast-2	opensearch.ap-northeast-2.amazonaws.com es.ap-northeast-2.amazonaws.com	HTTPS HTTPS
Asia-Pacífico (Osaka)	ap-northeast-3	opensearch.ap-northeast-3.amazonaws.com es.ap-northeast-3.amazonaws.com	HTTPS HTTPS

Nombre de la región	Región	Punto de conexión	Protocolo
Asia-Pacífico (Hyderabad)	ap-south-2	opensearch.ap-south-2.amazonaws.com es.ap-south-2.amazonaws.com	HTTPS HTTPS
Europa (España)	eu-south-2	opensearch.eu-south-2.amazonaws.com es.eu-south-2.amazonaws.com	HTTPS HTTPS
Oeste de EE. UU. (Norte de California)	us-west-1	opensearch.us-west-1.amazonaws.com es.us-west-1.amazonaws.com	HTTPS HTTPS
Europa (Zúrich)	eu-central-2	opensearch.eu-central-2.amazonaws.com es.eu-central-2.amazonaws.com	HTTPS HTTPS
Europa (Milán)	eu-south-1	opensearch.eu-south-1.amazonaws.com es.eu-south-1.amazonaws.com	HTTPS HTTPS

OpenSearch Cuotas de servicio de interfaz

Su AWS cuenta tiene las siguientes cuotas relacionadas con los recursos de la OpenSearch interfaz de usuario.

Name	Predeterminado	Ajustable	Notas
OpenSearch Aplicaciones de interfaz de usuario por cuenta y región	30	Sí	El número máximo de aplicaciones de OpenSearch interfaz de usuario que puede

Name	Predeterminado	Ajustable	Notas
			crear por cuenta y región. Puede aumentar el límite a 50 mediante la cuota de servicio y conseguir que se apruebe automátic amente. Para solicitar un límite superior, envía un ticket de soporte.

Gestión de índices en Amazon Service OpenSearch

Después de añadir datos a Amazon OpenSearch Service, a menudo es necesario volver a indexarlos, trabajar con alias de indexación, mover un índice a un almacenamiento más rentable o eliminarlo por completo. En este capítulo se tratan las soluciones de almacenamiento en niveles, la gestión del estado de los índices y otras operaciones de indexación. Para obtener información sobre el OpenSearch índice APIs, consulte la OpenSearch documentación.

Temas

- OpenSearch Instancias optimizadas para dominios OpenSearch de Amazon Service
- Almacenamiento de varios niveles para Amazon Service OpenSearch
- UltraWarm almacenamiento para Amazon OpenSearch Service
- Almacenamiento en frío para Amazon OpenSearch Service
- Index State Management en Amazon OpenSearch Service
- Resumen de índices en Amazon OpenSearch Service con paquetes acumulativos de índices
- Transformación de los índices en Amazon Service OpenSearch
- Replicación entre clústeres para Amazon Service OpenSearch
- Migración de índices de Amazon OpenSearch Service mediante la reindexación remota
- Gestión de datos de series temporales en Amazon OpenSearch Service con flujos de datos

OpenSearch Instancias optimizadas para dominios OpenSearch de Amazon Service

La familia de instancias OpenSearch optimizada para Amazon OpenSearch Service es una solución rentable para almacenar grandes volúmenes de datos. Un dominio con instancias OpenSearch optimizadas utiliza el almacenamiento local como principal y los datos se copian de forma sincrónica a Amazon S3 a medida que llegan. Esta estructura de almacenamiento proporciona un mayor rendimiento de indexación con una alta durabilidad. OR1, OR2, OM2 utiliza Amazon Elastic Block Store (Amazon `gp3` EBS) `io1` o volúmenes de forma local, OI2 mientras que las instancias utilizan discos locales. NVMe La familia de instancias OpenSearch optimizada también admite la recuperación automática de datos en caso de que se produzca un error. Para obtener información sobre las opciones de tipos de instancias OpenSearch optimizadas, consultethe section called “Tipos de instancias de generación actual”.

Si estás indexando cargas de trabajo de análisis operativo pesadas, como el análisis de registros, la observabilidad o el análisis de seguridad, puedes beneficiarte de la mejora del rendimiento y la eficiencia informática de OpenSearch las instancias optimizadas. Además, la recuperación automática de datos que ofrecen las instancias OpenSearch optimizadas mejora la fiabilidad general del dominio.

OpenSearch El servicio envía métricas de instancias OpenSearch optimizadas relacionadas con el almacenamiento a Amazon. CloudWatch Para ver una lista de las métricas disponibles, consulte ???.

OpenSearch las instancias optimizadas están disponibles bajo demanda o con precios de instancias reservadas, con una tarifa por hora para las instancias y el almacenamiento aprovisionados en Amazon EBS y Amazon S3.

Temas

- Limitaciones
- Ajustes para mejorar el rendimiento de la ingesta
- En qué se diferencian las instancias OpenSearch optimizadas de otras instancias
- En qué se diferencian las instancias OpenSearch optimizadas de UltraWarm las instancias
- Aprovisionamiento de un dominio con instancias OpenSearch optimizadas

Limitaciones

Tenga en cuenta las siguientes limitaciones cuando utilice instancias OpenSearch optimizadas para su dominio.

- Los dominios recién creados deben ejecutar la OpenSearch versión 2.11 o superior.
- Los dominios existentes deben ejecutar la OpenSearch versión 2.15 o superior.
- Su dominio debe tener activado el cifrado en reposo. Para obtener más información, consulte ???.
- Si el dominio usa nodos maestros dedicados, deben usar instancias de Graviton. Para obtener más información sobre nodos maestros dedicados, consulte ???.
- El intervalo de actualización de los índices de las instancias OpenSearch optimizadas debe ser de 10 segundos o más. El intervalo de actualización predeterminado para las instancias OpenSearch optimizadas es de 10 segundos.

Ajustes para mejorar el rendimiento de la ingesta

Para obtener el mejor rendimiento de indexación de sus instancias OpenSearch optimizadas, le recomendamos que haga lo siguiente:

- Use tamaños de carga masivos para mejorar la utilización del búfer. El tamaño recomendado es 10 MB.
- Utilice varios clientes para mejorar el rendimiento del procesamiento paralelo.
- Establezca la cantidad de particiones principales activas para que coincida con la cantidad de nodos de datos para maximizar el uso de los recursos.

En qué se diferencian las instancias OpenSearch optimizadas de otras instancias

OpenSearch Las instancias optimizadas se diferencian de las instancias no optimizadas en los siguientes aspectos:

- En el caso de las instancias OpenSearch optimizadas, la indexación solo se realiza en los fragmentos principales.
- Si las instancias OpenSearch optimizadas se configuran con réplicas, la tasa de indexación puede parecer inferior a la real. Por ejemplo, si hay una partición primario y una partición réplica, la velocidad de indexación podría mostrar una tasa de 1000 cuando la velocidad de indexación real es 2000.
- OpenSearch las instancias optimizadas realizan operaciones de búfer antes de enviarlas a una fuente remota. Esto se traduce en una latencia de ingesta superior.

> **Note**
>
> La métrica `IndexingLatency` no se ve afectada, ya que no incluye el tiempo de sincronización de translog.

- Las particiones de réplica pueden estar unos segundos por detrás de las particiones principales. Puedes monitorizar el retraso con la CloudWatch métrica de `ReplicationLagMaxTime` Amazon

En qué se diferencian las instancias OpenSearch optimizadas de UltraWarm las instancias

OpenSearch El servicio proporciona UltraWarm instancias que son una forma rentable de almacenar grandes cantidades de datos de solo lectura. Tanto las instancias OpenSearch optimizadas como UltraWarm las instancias almacenan los datos de forma local en Amazon EBS y de forma remota en Amazon S3. Sin embargo, las instancias OpenSearch optimizadas y UltraWarm las instancias difieren en varios aspectos importantes:

- OpenSearch las instancias optimizadas guardan una copia de los datos tanto en el almacén local como en el remoto. En UltraWarm algunos casos, los datos se guardan principalmente en un almacén remoto para reducir los costos de almacenamiento. Según sus patrones de uso, los datos se pueden mover al almacenamiento local.
- OpenSearch Las instancias optimizadas están activas y pueden aceptar operaciones de lectura y escritura, mientras que los datos de las UltraWarm instancias son de solo lectura hasta que se devuelvan manualmente al almacenamiento activo.
- UltraWarm se basa en las instantáneas indexadas para garantizar la durabilidad de los datos. OpenSearch En comparación, las instancias optimizadas realizan la replicación y la recuperación entre bastidores. En caso de que aparezca un índice rojo, las instancias OpenSearch optimizadas restaurarán automáticamente los fragmentos que falten del almacenamiento remoto en Amazon S3. El tiempo de recuperación depende del volumen de datos que deban recuperarse.

Para obtener más información sobre el UltraWarm almacenamiento, consulte???.

Aprovisionamiento de un dominio con instancias OpenSearch optimizadas

Puede seleccionar instancias OpenSearch optimizadas para sus nodos de datos al crear un nuevo dominio con Consola de administración de AWS o AWS Command Line Interface (AWS CLI). A continuación, puede indexar y consultar los datos mediante las operaciones y herramientas existentes.

Consola

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Dominios.
3. Elija Crear un dominio.

4. En la sección Número de nodos de datos, expanda el menú de la familia de instancias y elija OpenSearch optimizado.
5. Elija el tipo de instancia y otros ajustes de almacenamiento.
6. En la sección Cifrado, asegúrate de que esté seleccionada la opción Habilitar el cifrado de datos en reposo.
7. Configure el resto del dominio y seleccione Crear.

AWS CLI

Para aprovisionar un dominio que utilice un almacenamiento OpenSearch optimizado mediante el AWS CLI, debe proporcionar el valor del tamaño del tipo de instancia específico (por ejemplo OR1 OR2,, OM2, o OI2) en el`InstanceType`.

En el siguiente ejemplo, se crea un dominio con OR1 instancias de tamaño `2xlarge` y se habilita el cifrado en reposo.

```
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/*"}]}'
```

En el siguiente ejemplo, se crea un dominio con OI2 instancias de tamaño `large` y se habilita el cifrado en reposo. Tenga en cuenta que OI2 las instancias no requieren la configuración de EBS, ya que utilizan NVMe almacenamiento local.

```
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/*"}]}'
```

Almacenamiento de varios niveles para Amazon Service OpenSearch

El almacenamiento en varios niveles para Amazon OpenSearch Service es una solución de gestión de datos inteligente que optimiza el rendimiento y los costes mediante la gestión de los datos en diferentes niveles de almacenamiento. Esta arquitectura permite a las organizaciones equilibrar de manera eficiente el equilibrio entre rendimiento y costo al mantener los datos a los que se accede con frecuencia en un almacenamiento en caliente de alto rendimiento y, al mismo tiempo, transferir los datos a los que se accede con menos frecuencia a un almacenamiento en caliente más rentable.

Amazon OpenSearch Service ofrece dos opciones de arquitectura para el nivel hot/warm de almacenamiento:

- OpenSearch Arquitectura de almacenamiento de varios niveles
  - Combina Amazon S3 con el almacenamiento de instancias local
  - Funciona con instancias OpenSearch optimizadas
  - Soporta operaciones de escritura en niveles cálidos
  - Soporta una migración de datos fluida entre el nivel caliente y el cálido
  - Disponible a partir de la versión OpenSearch 3.3
  - No es compatible con Cold Tier
- UltraWarmArquitectura basada en
  - Combina Amazon S3 con el almacenamiento de instancias local
  - Funciona con UltraWarm instancias

- Optimizado para cargas de trabajo de nivel cálido de solo lectura
- Disponible en las versiones 6.8 y superiores de Elasticsearch y en todas las versiones OpenSearch
- Compatible con Cold Tier

> **Note**
>
> Esta documentación solo se centra en la arquitectura de varios niveles. Para conocer la arquitectura de almacenamiento ultracálido, consulte Almacenamiento ultracálido y en frío

Arquitectura de almacenamiento de varios niveles

Temas

- Ventajas principales
- Requisitos previos
- Limitaciones
- Puntos a tener en cuenta
- Crear un dominio de varios niveles
- Gestión de las migraciones de niveles
- Configuración de seguridad
- Prácticas recomendadas
- Supervisión de métricas

Ventajas principales

- Grabable en caliente: admite operaciones de escritura en índices en caliente
- Migración perfecta: movimiento fluido de datos entre los niveles de almacenamiento
- Optimización de costos: reduzca los costos de almacenamiento al mover los datos menos activos a un almacenamiento en caliente rentable
- Mejora del rendimiento: mantenga un alto rendimiento para los datos a los que se accede con frecuencia en la capa activa

- Administración de datos flexible: elija la arquitectura que mejor se adapte a sus requisitos de carga de trabajo
- Administración automatizada: administración simplificada del ciclo de vida de los datos en todos los niveles de almacenamiento

Requisitos previos

- Versión del motor: OpenSearch 3.3 o posterior
- Familias de instancias:
  - Nodos activos: OR1 OR2, OM2, o OI2
  - Nodos cálidos: OI2
- Seguridad: Node-to-node cifrado, cifrado en reposo, HTTPS reforzado

Limitaciones

- Funciona en todos los dominios con instancias OpenSearch optimizadas que aún no tengan Ultrawarm activado
- No es compatible con el nivel frío

Puntos a tener en cuenta

- Las migraciones de calor a clima cálido no provocan una fusión forzada en una arquitectura de varios niveles. Si es necesario, el usuario aún puede organizar fusiones forzadas mediante la política de gestión de índices.
- Además de la indexación, los nodos cálidos ahora también realizan operaciones de combinación en segundo plano (similares a las calientes).
- Todas las solicitudes de búsqueda en índices calientes se dirigen al fragmento principal y los servidores de réplica solo leen cuando el fragmento principal está inactivo.
- Esta arquitectura también admite instantáneas automatizadas para índices cálidos
- La replicación entre clústeres solo se admite para índices calientes
- Los índices APIs como Shrink, Split y Clone no funcionan con índices cálidos.

Crear un dominio de varios niveles

Paso 1: Crear el 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
```

Paso 2: Verificar los nodos calientes

```
aws opensearch describe-domain-nodes --domain-name my-domain --region us-east-1
```

Ejemplo de respuesta (extracto):

```
{
      "NodeType": "Warm",
      "InstanceType": "oi2.large.search",
      "NodeStatus": "Active"
    }
```

Gestión de las migraciones de niveles

Compatibilidad con dominios de varios niveles:

- Nueva organización en niveles APIs para una experiencia simplificada
- Legacy en materia UltraWarm APIs de compatibilidad

Nueva estratificación APIs

Migrar un índice a cálido:

```
curl -XPOST 'https://localhost:9200/index-name/_tier/warm'
```

Respuesta:

```
{"acknowledged": true}
```

Migrar un índice a caliente:

```
curl -XPOST 'https://localhost:9200/index-name/_tier/hot'
```

Respuesta:

```
{"acknowledged": true}
```

Compruebe el estado de la organización por niveles:

```
curl -XGET 'https://localhost:9200/index-name/_tier'
```

Ejemplo de respuesta:

```
{
      "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 compartida detallada:

```
curl 'https://localhost:9200/index1/_tier?detailed=true'
```

Enumere todas las migraciones en curso (texto):

```
curl 'https://localhost:9200/_tier/all'
```

Enumere todas las migraciones en curso (JSON):

```
curl 'https://localhost:9200/_tier/all?format=json'
```

Filtrar por nivel de destino:

```
curl 'https://localhost:9200/_tier/all?target=_warm'
```

Legado UltraWarm APIs por motivos de compatibilidad

Migre a una zona cálida:

```
curl -XPOST localhost:9200/_ultrawarm/migration/index2/_warm
```

Migrar a calor:

```
curl -XPOST localhost:9200/_ultrawarm/migration/index2/_hot
```

Compruebe el estado:

```
curl -XGET localhost:9200/_ultrawarm/migration/index2/_status
```

Configuración de seguridad

Si habilitas el almacenamiento de varios niveles en un dominio de Amazon OpenSearch Service preexistente, es posible que la `storage_tiering_manager` función no esté definida en el dominio. Los usuarios que no sean administradores deben estar asignados a este rol para poder administrar índices templados en los dominios mediante un control de acceso detallado. Para crear el rol `storage_tiering_manager` de forma manual, siga estos pasos:

1. En los OpenSearch paneles, vaya a Seguridad y elija Permisos.
2. Seleccione Crear grupo de acciones y configure los siguientes grupos:

Nombre de grupo	Permisos
storage_tiering_cluster	índices: admin/_tier/all
storage_tiering_index_read	índices: admin/_ tier/get, indices:admin/get
storage_tiering_index_write	índices: admin/_ _to_hot tier/hot_to_warm, indices:admin/_tier/warm

3. Seleccione Roles y, a continuación, Crear rol.
4. Llame al rol `storage_tiering_manager`.
5. Para Permisos de clúster, seleccione `storage_tiering_cluster` y `cluster_monitor`.
6. Para Índice, escriba `*`.
7. Para los permisos de índice, seleccione y. `storage_tiering_index_read` `storage_tiering_index_write` `indices_monitor`
8. Seleccione Crear.
9. Después de crear el rol, asígnelo a cualquier rol de usuario o backend que administre índices de varios niveles.

Prácticas recomendadas

Al implementar el almacenamiento de varios niveles en sus dominios de Amazon OpenSearch Service, tenga en cuenta las siguientes prácticas recomendadas:

- Revise periódicamente sus patrones de acceso a los datos para optimizar la asignación de niveles
- Supervise las métricas de rendimiento para garantizar una utilización eficiente de los recursos
- Utilice la nueva organización en niveles APIs para un control pormenorizado de la migración de datos

Supervisión de métricas

Los dominios de almacenamiento de varios niveles proporcionan métricas adicionales para monitorear el rendimiento de los niveles intermedios. Estas métricas incluyen tanto las UltraWarm

métricas existentes como las nuevas métricas específicas de la arquitectura de instancias OpenSearch optimizadas:

Nuevas métricas

Nombre de métrica	Estadísticas a nivel de nodo	Estadísticas a nivel de clúster	Granularity (Grado de detalle)
WarmIndexingLatency	Media	Media	1 min
WarmIndexingRate	Media	promedio, máximo, suma	1 min
WarmThreadpoolIndexingQueue	Máximo	Suma, máximo, promedio	1 min
WarmThreadpoolIndexingRejected	Máximo	Sum	1 min
WarmThreadpoolIndexingThreads	Máximo	Suma, promedio	1 min

UltraWarm almacenamiento para Amazon OpenSearch Service

UltraWarm proporciona una forma rentable de almacenar grandes cantidades de datos de solo lectura en Amazon OpenSearch Service. Los nodos de datos estándar utilizan almacenamiento "en caliente", que adopta la forma de almacenes de instancias o volúmenes de Amazon EBS asociados a cada nodo. El almacenamiento en caliente proporciona el rendimiento más rápido posible para indexar y buscar nuevos datos.

En lugar de almacenamiento adjunto, UltraWarm los nodos utilizan Amazon S3 y una sofisticada solución de almacenamiento en caché para mejorar el rendimiento. En el caso de los índices en los que no está escribiendo activamente, que consultan con menos frecuencia y que no necesitan el mismo rendimiento, UltraWarm ofrecen costes significativamente más bajos por GiB de datos. Dado que los índices calientes son de solo lectura, a menos que se devuelvan al almacenamiento activo, UltraWarm es ideal para datos inmutables, como los registros.

En OpenSearch, los índices cálidos se comportan igual que cualquier otro índice. Puede consultarlos de la misma manera APIs o usarlos para crear visualizaciones en OpenSearch los paneles de control.

Temas

- Requisitos previos
- UltraWarm requisitos de almacenamiento y consideraciones de rendimiento
- UltraWarm precios
- Habilitando UltraWarm
- Migración de índices al almacenamiento UltraWarm
- Automatizar migraciones
- Ajuste de migración
- Cancelación de migraciones
- Listado de índices calientes y templados
- Devolución de índices templados al almacenamiento caliente
- Restauración de índices templados a partir de instantáneas
- Instantáneas manuales de índices templados
- Migración de índices templados al almacenamiento frío
- Prácticas recomendadas para índices KNN
- Inhabilitando UltraWarm

Requisitos previos

UltraWarm tiene algunos requisitos previos importantes:

- UltraWarm requiere Elasticsearch 6.8 OpenSearch o superior.
- Para utilizar el almacenamiento "warm", los dominios deben tener nodos maestros dedicados.
- Cuando se utiliza un dominio Multi-AZ con modo de espera, la cantidad de nodos templados debe ser un múltiplo de la cantidad de zonas de disponibilidad que se utilizan.
- Si su dominio utiliza un tipo de instancia T2 o T3 para los nodos de datos, no puede utilizar el almacenamiento templado.
- Si su índice usa k-NN aproximado (`"index.knn":true`), puede pasarlo a almacenamiento en caliente a partir de la versión 2.17 y posteriores. Los dominios de versiones anteriores a la 2.17

pueden actualizarse a la 2.17 para usar esta funcionalidad, pero los índices de KNN creados en versiones anteriores a la 2.x no pueden migrar a ella. UltraWarm

- Si el dominio usa un control de acceso detallado, los usuarios deben estar asignados al rol en los paneles para poder realizar llamadas a la `ultrawarm_manager` API. OpenSearch UltraWarm

> **Note**
>
> Es posible que la `ultrawarm_manager` función no esté definida en algunos dominios de servicio preexistentes. OpenSearch Si no ve el rol en el panel, debe crearlo de forma manual.

Configuración de permisos

Si lo habilita UltraWarm en un dominio de OpenSearch servicio preexistente, es posible que el `ultrawarm_manager` rol no esté definido en el dominio. Los usuarios que no sean administradores deben estar asignados a este rol para poder administrar índices templados en los dominios mediante un control de acceso detallado. Para crear el rol `ultrawarm_manager` de forma manual, siga estos pasos:

1. En los OpenSearch paneles, vaya a Seguridad y elija Permisos.
2. Seleccione Crear grupo de acciones y configure los siguientes grupos:

Nombre del grupo	Permisos
`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. Seleccione Roles y, a continuación, Crear rol.

4. Asigne al rol el nombre ultrawarm_manager.
5. Para Permisos de clúster, seleccione `ultrawarm_cluster` y `cluster_monitor`.
6. Para Índice, escriba `*`.
7. Para Permisos de índice, seleccione `ultrawarm_index_read`, `ultrawarm_index_write`, y `indices_monitor`.
8. Elija Create (Crear).
9. Después de crear el rol, asígnelo a cualquier rol de usuario o backend que UltraWarm administre los índices.

UltraWarm requisitos de almacenamiento y consideraciones de rendimiento

Como se explica en el the section called “Cálculo de requisitos de almacenamiento” apartado anterior, los datos almacenados en caliente suponen una sobrecarga importante: réplicas, espacio reservado en Linux y espacio reservado en el OpenSearch servicio. Por ejemplo, una partición principal de 20 GiB con una partición de réplica requiere aproximadamente 58 GiB de almacenamiento en caliente.

Como utiliza Amazon S3, no UltraWarm incurre en esta sobrecarga. Al calcular los requisitos UltraWarm de almacenamiento, solo se tiene en cuenta el tamaño de las particiones principales. La durabilidad de los datos en S3 elimina la necesidad de mantener réplicas y S3 abstrae (y hace innecesarias) las consideraciones de sistemas operativos o de servicio. Esa misma partición de 20 GiB requiere solo 20 GiB de almacenamiento templado. Si aprovisiona una instancia `ultrawarm1.large.search`, puede usar los 20 TiB de su almacenamiento máximo para particiones principales. Consulte the section called “UltraWarm cuotas de almacenamiento” para obtener un resumen de los tipos de instancia y la cantidad máxima de almacenamiento que puede abordar cada uno.

Con UltraWarm, seguimos recomendando un tamaño de fragmento máximo de 50 GiB. El número de núcleos de CPU y la cantidad de RAM asignados a cada tipo de UltraWarm instancia te dan una idea del número de fragmentos que pueden buscar simultáneamente. Tenga en cuenta que, si bien solo los fragmentos principales cuentan para el UltraWarm almacenamiento en S3, los OpenSearch cuadros de mando indican `_cat/indices` el tamaño del UltraWarm índice como el total de todos los fragmentos principales y de réplica.

Por ejemplo, cada instancia de `ultrawarm1.medium.search` tiene dos núcleos de CPU y puede abordar hasta 1,5 TiB de almacenamiento en S3. Dos de estas instancias tienen una combinación

de 3 TiB de almacenamiento, que funciona en aproximadamente 62 particiones si cada partición es de 50 GiB. Si una solicitud al clúster solo busca cuatro de estas particiones, el rendimiento puede ser excelente. Si la solicitud es amplia y busca los 62, es posible que los cuatro núcleos de CPU tengan dificultades para realizar la operación. Supervise las `WarmJVMMemoryPressure` UltraWarm métricas `WarmCPUUtilization` y las métricas para comprender cómo gestionan las instancias sus cargas de trabajo.

Si realiza búsquedas amplias o frecuentes, considere dejar los índices en el almacenamiento caliente. Al igual que cualquier otra OpenSearch carga de trabajo, el paso más importante para determinar si UltraWarm cumple con tus necesidades es realizar pruebas representativas con los clientes utilizando un conjunto de datos realista.

UltraWarm precios

Con el almacenamiento en caliente, se paga lo que se aprovisiona. Algunas instancias requieren un volumen de Amazon EBS asociado, mientras que otras incluyen un almacén de instancias. Tanto si el almacenamiento está vacío como si está lleno, se paga el mismo precio.

Con el UltraWarm almacenamiento, pagas por lo que usas. Una instancia `ultrawarm1.large.search` puede poner a disposición hasta 20 TiB de almacenamiento en S3. Sin embargo, si únicamente se almacena 1 TiB de datos, solo se le facturará 1 TiB de datos. Como todos los demás tipos de nodos, también pagas una tarifa por hora por cada UltraWarm nodo. Para obtener más información, consulte the section called “Precios”.

Habilitando UltraWarm

La consola es la forma más sencilla de crear un dominio que utiliza almacenamiento templado. Al crear el dominio, elija Habilitar nodos de datos en caliente y la cantidad de nodos en caliente que desee. El mismo proceso básico funciona en dominios existentes, siempre que cumplan los requisitos previos. Incluso después de que el estado del dominio cambie de Procesado a Activo, es UltraWarm posible que no esté disponible para su uso durante varias horas.

Cuando se utiliza un dominio Multi-AZ con modo de espera, la cantidad de nodos templados debe ser un múltiplo de la cantidad de zonas de disponibilidad. Para obtener más información, consulte the section called “Multi-AZ con modo de espera”.

También puedes usar la API de configuración AWS CLIo la API para habilitar UltraWarm`WarmEnabled`, específicamente`WarmCount`, `WarmType` las opciones y`ClusterConfig`.

> **Note**
>
> Los dominios admiten un número máximo de nodos templados. Para más información, consulte the section called “Cuotas”.

Ejemplo de comando de la CLI

El siguiente AWS CLI comando crea un dominio con tres nodos de datos, tres nodos maestros dedicados, seis nodos calientes y un control de acceso detallado activado:

```
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
```

Para obtener más información, consulte la Referencia de comandos de la AWS CLI.

Ejemplo de solicitud a la API de configuración

La siguiente solicitud a la API de configuración crea un dominio con tres nodos de datos, tres nodos maestros dedicados y seis nodos templados con control de acceso detallado habilitado y una política de acceso restrictiva:

```
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/*\"}]}"
}
```

Para obtener información detallada, consulta la referencia de la API OpenSearch de Amazon Service.

Migración de índices al almacenamiento UltraWarm

Si ha terminado de escribir en un índice y ya no necesita el rendimiento de búsqueda más rápido posible, migre el índice a: UltraWarm

```
POST _ultrawarm/migration/my-index/_warm
```

A continuación, verifique el estado de la migración:

```
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
    }
  }
}
```

El estado del índice debe ser verde para realizar una migración. Si migra varios índices en una sucesión rápida, puede obtener un resumen de todas las migraciones en texto sin formato, similar a la API `_cat`:

```
GET _ultrawarm/migration/_status?v

index    migration_type state
my-index HOT_TO_WARM    RUNNING_SHARD_RELOCATION
```

OpenSearch El servicio migra un índice a la vez a UltraWarm. Se pueden tener hasta 200 migraciones en la cola. Cualquier solicitud que supere el límite será rechazada.

Para comprobar el número actual de migraciones en la cola, supervise la métrica de `HotToWarmMigrationQueueSize`. Los índices permanecen disponibles durante todo el proceso de migración, sin tiempo de inactividad.

El proceso de migración tiene los siguientes estados:

```
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
```

Como indican estos estados, se pueden producir errores en las migraciones si se realizan durante instantáneas, reubicaciones de particiones o fusiones forzosas. Los errores durante las instantáneas o las reubicaciones de particiones suelen deberse a errores de nodo o problemas de conectividad de S3. La falta de espacio en el disco suele ser la causa subyacente de los errores en las fusiones forzosas.

Una vez finalizada la migración, la misma solicitud `_status` devuelve un error. Si comprueba el índice en ese momento, verá algunos parámetros de configuración que son exclusivos de los índices templados:

```
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`, en este caso, es el número de réplicas pasivas, que no consumen espacio en disco.
- `routing.allocation.require.box_type` especifica que el índice debe usar nodos templados en lugar de nodos de datos estándar.
- `merge.policy.max_merge_at_once_explicit` especifica el número de segmentos que se van a fusionar simultáneamente durante la migración.

Los índices almacenados en caliente son de solo lectura, a menos que se devuelvan al almacenamiento activo, lo que los hace UltraWarm más adecuados para datos inmutables, como los registros. Puede consultar los índices y eliminarlos, pero no puede agregar, actualizar ni eliminar documentos individuales. Si lo intenta, podría encontrarse con el siguiente error:

```
{
  "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
}
```

Automatizar migraciones

Se recomienda utilizar the section called “Index State Management” para automatizar el proceso de migración después de que un índice alcance cierta antigüedad o cumpla otras condiciones. Consulte la política de ejemplo que muestra este flujo de trabajo.

Ajuste de migración

Las migraciones de índices al almacenamiento requieren una fusión forzada. UltraWarm Cada OpenSearch índice se compone de un número determinado de fragmentos y cada fragmento se compone de un número determinado de segmentos de Lucene. La operación de fusión forzada purga los documentos marcados para su eliminación y conserva espacio en disco. De forma predeterminada, UltraWarm fusiona los índices en un segmento, excepto en los índices kNN, en los que se utiliza un valor predeterminado de 20.

Puede cambiar este valor hasta 1000 segmentos utilizando la configuración de `index.ultrawarm.migration.force_merge.max_num_segments`. Los valores más altos aceleran el proceso de migración, pero aumentan la latencia de consulta para el índice activo una vez finalizada la migración. Para cambiar la configuración, realice la siguiente solicitud:

```
PUT my-index/_settings
{
  "index": {
    "ultrawarm": {
      "migration": {
        "force_merge": {
          "max_num_segments": 1
        }
```

```
      }
    }
  }
}
```

Para comprobar cuánto tiempo tarda esta etapa del proceso de migración, monitoree la métrica de `HotToWarmMigrationForceMergeLatency`.

Cancelación de migraciones

UltraWarm gestiona las migraciones de forma secuencial, en una cola. Si una migración está en la cola, pero aún no se ha iniciado, puede eliminarla de la cola mediante la siguiente solicitud:

```
POST _ultrawarm/migration/_cancel/my-index
```

Si el dominio utiliza un control de acceso detallado, debe tener el permiso de `indices:admin/ultrawarm/migration/cancel` para realizar esta solicitud.

Listado de índices calientes y templados

UltraWarm añade dos opciones adicionales, similares a `_all`, para ayudar a gestionar los índices calientes y cálidos. Para obtener una lista de todos los índices calientes o templados, realice las siguientes solicitudes:

```
GET _warm
GET _hot
```

Puede utilizar estas opciones en otras solicitudes que especifican índices, por ejemplo:

```
_cat/indices/_warm
_cluster/state/_all/_hot
```

Devolución de índices templados al almacenamiento caliente

Si necesita volver a escribir en un índice, vuelva a migrarlo al almacenamiento en caliente:

```
POST _ultrawarm/migration/my-index/_hot
```

Puede tener hasta 10 migraciones en cola de almacenamiento caliente a almacenamiento en caliente a la vez. OpenSearch El servicio procesa las solicitudes de migración de una en una, en

el orden en que se pusieron en cola. Para comprobar el número actual, monitoree la métrica de `WarmToHotMigrationQueueSize`.

Una vez finalizada la migración, compruebe la configuración del índice para asegurarse de que satisfaga sus necesidades. Los índices vuelven al almacenamiento caliente con una réplica.

Restauración de índices templados a partir de instantáneas

Además del repositorio estándar para instantáneas automatizadas, UltraWarm añade un segundo repositorio para índices calientes,. `cs-ultrawarm` Cada instantánea de este repositorio contiene solo un índice. Si elimina un índice caliente, su instantánea permanece en el repositorio de `cs-ultrawarm` durante 14 días, al igual que cualquier otra instantánea automatizada.

Cuando restaura una instantánea desde `cs-ultrawarm`, se restaura en el almacenamiento "warm", no en el almacenamiento en caliente. Las instantáneas de los repositorios `cs-automated` y `cs-automated-enc` restauran el almacenamiento en caliente.

Para restaurar una UltraWarm instantánea en un almacenamiento en caliente

1. Identifique la instantánea más reciente que contiene el índice que desea restaurar:

```
GET _snapshot/cs-ultrawarm/_all?verbose=false

{
  "snapshots": [{
    "snapshot": "snapshot-name",
    "version": "1.0",
    "indices": [
      "my-index"
    ]
  }]
}
```

> **Note**
>
> De forma predeterminada, la `GET _snapshot/<repo>` operación muestra información detallada sobre los datos, como la hora de inicio, la hora de finalización y la duración de cada instantánea de un repositorio. La `GET _snapshot/<repo>` operación recupera información de los archivos de cada instantánea contenidos en un repositorio. Si no necesita la hora de inicio, la hora de finalización y la duración y solo necesita el

nombre y la información de índice de una instantánea, le recomendamos que utilice el `verbose=false` parámetro al enumerar las instantáneas para minimizar el tiempo de procesamiento y evitar que se agote el tiempo de espera.

2. Si el índice ya existe, elimínelo:

```
DELETE my-index
```

Si no quiere eliminar el índice, debe devolverlo al almacenamiento en caliente y reindexarlo.

3. Restaurar la instantánea:

```
POST _snapshot/cs-ultrawarm/snapshot-name/_restore
```

UltraWarm ignora cualquier configuración de índice que especifique en esta solicitud de restauración, pero puede especificar opciones como `rename_pattern` y`rename_replacement`. Para ver un resumen de las opciones de restauración de OpenSearch instantáneas, consulte la OpenSearch documentación.

Instantáneas manuales de índices templados

Usted puede tomar instantáneas manuales de índices templados, pero no lo recomendamos. El repositorio de automatización `cs-ultrawarm` ya contiene una instantánea para cada índice activo, tomada durante la migración, sin cargo adicional.

De forma predeterminada, el OpenSearch servicio no incluye índices calientes en las instantáneas manuales. Por ejemplo, la siguiente llamada solo incluye índices calientes:

```
PUT _snapshot/my-repository/my-snapshot
```

Si elige tomar instantáneas manuales de índices templados, se aplican varias consideraciones importantes.

- No puede mezclar índices calientes y templados. Por ejemplo, la siguiente solicitud devuelve un error:

```
PUT _snapshot/my-repository/my-snapshot
{
```

```
  "indices": "warm-index-1,hot-index-1",
  "include_global_state": false
}
```

Si incluyen una mezcla de índices calientes y cálidos, comodín (*) también devuelven un error.

- Solo se puede incluir un índice templado por instantánea. Por ejemplo, la siguiente solicitud devuelve un error:

```
PUT _snapshot/my-repository/my-snapshot
{
  "indices": "warm-index-1,warm-index-2,other-warm-indices-*",
  "include_global_state": false
}
```

Esta solicitud se realiza correctamente:

```
PUT _snapshot/my-repository/my-snapshot
{
  "indices": "warm-index-1",
  "include_global_state": false
}
```

- Las instantáneas manuales siempre se restauran al almacenamiento en caliente, incluso si originalmente incluían un índice templado.

Migración de índices templados al almacenamiento frío

Si tiene datos UltraWarm que consulta con poca frecuencia, considere la posibilidad de migrarlos a un almacenamiento en frío. El almacenamiento frío está diseñado para los datos a los que solo se accede ocasionalmente o que ya no están en uso activo. No puede leer ni escribir en índices fríos, pero puede migrarlos de nuevo al almacenamiento templado sin costo cada vez que necesite consultarlos. Para obtener instrucciones, consulte Migración de índices a cámaras frigoríficas.

Prácticas recomendadas para índices KNN

- El nivel Ultrawarm/Cold está disponible para todos los tipos de motores de indexación KNN. Lo recomendamos para los índices KNN que utilizan el motor Lucene y la búsqueda vectorial optimizada para discos, ya que no es necesario cargar completamente los datos del gráfico en una memoria externa. Al usarlo con motores integrados en memoria nativos, como FAISS y

NMSLIB, debe tener en cuenta el tamaño del gráfico de los fragmentos en los que se buscará activamente y aprovisionar las UltraWarm instancias, preferiblemente del tipo de instancia, en consecuencia. `uw.large` Por ejemplo, si los clientes tienen 2 instancias `uw.large` configuradas, cada una tendrá aproximadamente GiB de `knn.memory.circuit_breaker.limit * 61` de memoria fuera del montón disponibles. Se obtiene un rendimiento óptimo si todas las consultas en caliente se dirigen a particiones cuyo tamaño de gráfico acumulado no supere la memoria externa disponible. La latencia se ve afectada si la memoria disponible es inferior a la necesaria para cargar el gráfico debido a las expulsiones y a la espera de que se disponga de memoria acumulada. Por eso no recomendamos el uso de instancias de `uw.medium` en los casos de uso en los que se utilicen motores integrados en memoria o para casos de mayor rendimiento de búsqueda, independientemente de los motores.

- Los índices KNN que migren a no se fusionarán forzadamente en un solo segmento. UltraWarm Esto evita cualquier impacto en los nodos en caliente y cálidos que podrían enfrentar problemas de falta de memoria (OOM) debido a que el tamaño del grafo se vuelve demasiado grande para los motores en memoria. Debido al aumento del número de segmentos por partición, esto podría consumir más espacio en la memoria caché local y permitir que menos índices migren a la capa cálida. Puede optar por forzar la fusión de índices en un solo segmento utilizando la configuración existente y anulándola antes de migrar los índices a la capa cálida. Para obtener más información, consulte the section called “Ajuste de migración”.
- Si tiene un caso de uso en el que los índices se buscan con poca frecuencia y no representan una carga de trabajo sensible a la latencia, puede optar por migrar esos índices al nivel. UltraWarm Esto le ayudará a reducir la escala de las instancias informáticas de nivel activo y permitir que la informática de UltraWarm nivel gestione la consulta en esos índices de baja prioridad. Esto también puede ayudar a aislar los recursos consumidos entre las consultas de los índices de alta y baja prioridad para que no se afecten entre sí.

Inhabilitando UltraWarm

La consola es la forma más sencilla de deshabilitar UltraWarm. Seleccione el dominio, Acciones y Editar la configuración del clúster. Anule la selección de Habilitar nodos de datos cálidos y elija Guardar los cambios. También puede usar la opción `WarmEnabled` en la API de configuración y en la AWS CLI .

Antes de deshabilitar UltraWarm, debe eliminar todos los índices calientes o migrarlos de nuevo al almacenamiento activo. Cuando el almacenamiento caliente esté vacío, espere cinco minutos antes de intentar UltraWarm desactivarlo.

Almacenamiento en frío para Amazon OpenSearch Service

El almacenamiento en frío te permite almacenar cualquier cantidad de datos históricos o a los que accedes con poca frecuencia en tu dominio de Amazon OpenSearch Service y analizarlos bajo demanda, a un coste inferior al de otros niveles de almacenamiento. El almacenamiento en frío es apropiado si necesita realizar investigaciones periódicas o análisis forenses sobre datos antiguos. Algunos ejemplos prácticos de datos apropiados para el almacenamiento en frío son los registros a los que se accede con poca frecuencia, los datos que se deben conservar para cumplir con los requisitos de conformidad o los registros que tienen valor histórico.

Al igual que el UltraWarmalmacenamiento, el almacenamiento en frío está respaldado por Amazon S3. Cuando necesite consultar datos inactivos, puede adjuntarlos de forma selectiva a UltraWarm los nodos existentes. Puede administrar la migración y el ciclo de vida de los datos almacenados en frío manualmente o mediante políticas de administración de estados de índice.

Temas

- Requisitos previos
- Requisitos de almacenamiento en frío y consideraciones de rendimiento
- Precio del almacenamiento en frío
- Habilitación del almacenamiento en frío
- Gestión de índices inactivos en los paneles OpenSearch
- Migración de índices al almacenamiento frío
- Automatización de las migraciones al almacenamiento en frío
- Cancelación de migraciones al almacenamiento en frío
- Listado de índices almacenados en frío
- Migración de índices fríos al almacenamiento templado
- Restauración de índices fríos a partir de instantáneas
- Cancelación de migraciones del almacenamiento en frío al templado
- Actualización de metadatos de índices almacenados en frío
- Eliminación de índices almacenados en frío
- Deshabilitar el almacenamiento en frío

Requisitos previos

El almacenamiento en frío tiene los siguientes requisitos previos:

- El almacenamiento en frío requiere la versión 7.9 OpenSearch o posterior de Elasticsearch.
- Para habilitar el almacenamiento en frío en un dominio de OpenSearch servicio, también debe habilitar el almacenamiento en caliente en el mismo dominio.
- Para utilizar el almacenamiento en frío, los dominios deben tener nodos maestros dedicados.
- Si un dominio utiliza un tipo de instancia T2 o T3 para los nodos de datos, no puede utilizar el almacenamiento en frío.
- Si su índice utiliza k-NN aproximado (`"index.knn":true`), puede trasladarlo al almacenamiento en frío a partir de la versión 2.17 en adelante. Los dominios de versiones anteriores a la 2.17 pueden actualizarse a la 2.17 para utilizar esta funcionalidad, pero los índices KNN creados en versiones anteriores a la 2.x no pueden migrar a frío.
- Si el dominio utiliza un control de acceso detallado, los usuarios que no sean administradores deben asignarse a esa `cold_manager` función en los OpenSearch paneles para poder gestionar los índices inactivos.

> **Note**
>
> Es posible que el `cold_manager` rol no exista en algunos dominios de servicio preexistentes. OpenSearch Si no ve el rol en el panel, debe crearlo de forma manual.

Configuración de permisos

Si habilita el almacenamiento en frío en un dominio de OpenSearch servicio preexistente, es posible que la `cold_manager` función no esté definida en el dominio. Si el dominio utiliza el control de acceso detallado, los usuarios que no sean administradores deben estar asignados a este rol para poder administrar índices almacenados en frío. Para crear el rol `cold_manager` de forma manual, siga estos pasos:

1. En los OpenSearch paneles, vaya a Seguridad y elija Permisos.
2. Seleccione Crear grupo de acciones y configure los siguientes grupos:

Nombre del grupo	Permisos
`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. Seleccione Roles y, a continuación, Crear rol.
4. Nombre el rol cold_manager.
5. En Permisos de clúster, seleccione el grupo `cold_cluster` que creó.
6. En Índice, ingrese `*`.
7. En Permisos de índice, seleccione el grupo `cold_index` que creó.
8. Seleccione Crear.
9. Después de crear el rol, asígnelo a cualquier rol de usuario o backend que administre índices almacenados en frío.

Requisitos de almacenamiento en frío y consideraciones de rendimiento

Como el almacenamiento en frío utiliza Amazon S3, no incurre en la sobrecarga del almacenamiento en caliente, como las réplicas, el espacio reservado de Linux y el espacio reservado de OpenSearch servicio. El almacenamiento en frío no tiene tipos de instancias específicos porque no tiene ninguna capacidad informática asociada a él. Puede almacenar cualquier cantidad de datos en el almacenamiento en frío. Supervisa la `ColdStorageSpaceUtilization` métrica en Amazon CloudWatch para ver cuánto espacio de almacenamiento en frío estás utilizando.

Precio del almacenamiento en frío

Al igual que el UltraWarm almacenamiento, con el almacenamiento en frío solo pagas por el almacenamiento de datos. No hay ningún costo informático para los datos en frío y no incurrirá en gastos si no hay datos en el almacenamiento en frío.

No incurrirá en cargos de transferencia al mover datos entre los almacenamientos en frío y templado. Mientras se realiza la migración de los índices entre el almacenamiento templado y el frío, sigue pagando por una sola copia del índice. Una vez finalizada la migración, el índice se factura de acuerdo con el nivel de almacenamiento al que se migró. Para obtener más información sobre los precios del almacenamiento en frío, consulta los precios OpenSearch de Amazon Service.

Habilitación del almacenamiento en frío

La consola es la forma más sencilla de crear un dominio que utiliza el almacenamiento en frío. Mientras crea el dominio, primero elija Habilitar nodos de datos en caliente, porque debe habilitar el almacenamiento en caliente en el mismo dominio. A continuación, elija Habilitar almacenamiento en frío.

El mismo proceso funciona en dominios existentes, siempre que cumplan los requisitos previos. Incluso después de que el estado del dominio cambie de En proceso a Activo, es posible que el almacenamiento en frío no esté disponible para su utilización por varias horas.

También puede utilizar AWS CLI o la API de configuración para habilitar el almacenamiento en frío.

Ejemplo de comando de la CLI

El siguiente AWS CLI comando crea un dominio con tres nodos de datos, tres nodos maestros dedicados, habilitados para el almacenamiento en frío y habilitados para el control de acceso detallado:

```
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
```

Para obtener más información, consulte la Referencia de comandos de la AWS CLI.

Ejemplo de solicitud a la API de configuración

La siguiente solicitud a la API de configuración crea un dominio con tres nodos de datos, tres nodos maestros dedicados, almacenamiento en frío habilitado y control de acceso detallado habilitado:

```
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"
}
```

Para obtener información detallada, consulta la referencia de la API OpenSearch de Amazon Service.

Gestión de índices inactivos en los paneles OpenSearch

Puede administrar los índices de calor, calor y frío con la interfaz de paneles existente en su dominio de servicio. OpenSearch Dashboards permite migrar índices entre el almacenamiento frío y el templado, así como supervisar el estado de la migración de los índices, sin utilizar la CLI ni la API de configuración. Para obtener más información, consulte Administrar índices en los paneles de control. OpenSearch

Migración de índices al almacenamiento frío

Cuando se migran índices al almacenamiento frío, se debe proporcionar un intervalo de tiempo a los datos para facilitar el descubrimiento. Puede elegir un campo de marca temporal basado en los datos del índice, proporcionar de forma manual una marca temporal inicial y final o elegir no especificar una.

Parámetro	Valor admitido	Description (Descripción)
`timestamp_field`	El date/time campo del mapeo de índices.	Los valores mínimo y máximo del campo proporcionado se calculan y almacenan como metadatos `start_time` y `end_time` para el índice almacenado en frío.
`start_time` y `end_time`	Utilice uno de los siguientes formatos:	Los valores proporcionados se almacenan como metadatos

Parámetro	Valor admitido	Description (Descripción)
	• strict_date_optional_time. Por ejemplo, `yyyy-MM-dd'T'HH:mm:ss.SSSZ` o `yyyy-MM-dd` • Fecha de inicio en milisegundos	`start_time` y `end_time` para el índice almacenado en frío.

Si no desea especificar una marca temporal, agregue `?ignore=timestamp` a la solicitud en su lugar.

La siguiente solicitud migra un índice del almacenamiento templado al almacenamiento en frío y proporciona horas de inicio y finalización para los datos de ese índice:

```
POST _ultrawarm/migration/my-index/_cold
  {
    "start_time": "2020-03-09",
    "end_time": "2020-03-09T23:00:00Z"
  }
```

A continuación, verifique el estado de la migración:

```
GET _ultrawarm/migration/my-index/_status

{
  "migration_status": {
    "index": "my-index",
    "state": "RUNNING_METADATA_RELOCATION",
    "migration_type": "WARM_TO_COLD"
  }
}
```

OpenSearch El servicio migra un índice a la vez al almacenamiento en frío. Se pueden tener hasta 100 migraciones en la cola. Cualquier solicitud que supere el límite será rechazada. Para comprobar el número actual de migraciones en la cola, supervise la métrica de `WarmToColdMigrationQueueSize`. El proceso de migración tiene los siguientes estados:

```
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.
```

Automatización de las migraciones al almacenamiento en frío

Se recomienda utilizar la Administración de estados de índice para automatizar el proceso de migración después de que un índice alcance cierta antigüedad o cumpla otras condiciones. Consulte el ejemplo de política, que muestra cómo migrar automáticamente los índices del almacenamiento en caliente UltraWarm al almacenamiento en frío.

> **Note**
>
> Es necesario un `timestamp_field` explícito para mover los índices al almacenamiento frío mediante una política de administración de estados de índice.

Cancelación de migraciones al almacenamiento en frío

Si una migración al almacenamiento en frío está en cola o con un estado de error, puede cancelar la migración mediante la siguiente solicitud:

```
POST _ultrawarm/migration/_cancel/my-index

{
  "acknowledged" : true
}
```

Si el dominio utiliza un control de acceso detallado, necesita el permiso `indices:admin/ultrawarm/migration/cancel` para realizar esta solicitud.

Listado de índices almacenados en frío

Antes de realizar la consulta, puede enumerar los índices almacenados en frío para decidir a cuáles migrar para analizarlos más a fondo UltraWarm . La siguiente solicitud enumera todos los índices almacenados en frío, ordenados por nombre de índice:

```
GET _cold/indices/_search
```

Respuesta de ejemplo

```
{
  "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"
    }
  ]
}
```

Filtrado

Puede filtrar los índices fríos según un patrón de índice basado en prefijos y variaciones del intervalo de tiempo.

La siguiente solicitud presenta una lista de los índices que coinciden con el patrón de prefijo `event-*`:

```
GET _cold/indices/_search
 {
   "filters":{
      "index_pattern": "event-*"
   }
 }
```

Respuesta de ejemplo

```
{
  "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 siguiente solicitud devuelve índices cuyos campos de metadatos `start_time` y `end_time` están comprendidos entre `2019-03-01` y `2020-03-01`:

```
GET _cold/indices/_search
{
  "filters": {
    "time_range": {
      "start_time": "2019-03-01",
      "end_time": "2020-03-01"
```

```
    }
  }
}
```

Respuesta de ejemplo

```
{
  "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"
    }
  ]
}
```

Ordenar

Puede ordenar los índices fríos por campos de metadatos, como el nombre del índice o su tamaño. La siguiente solicitud presenta una lista de todos los índices ordenados por tamaño en orden descendente:

```
GET _cold/indices/_search
 {
 "sort_key": "size:desc"
 }
```

Respuesta de ejemplo

```
{
  "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"
    }
  ]
}
```

Otras claves de ordenamiento válidas son `start_time:asc/desc`, `end_time:asc/desc` y `index_name:asc/desc`.

Paginación

Puede paginar una lista de índices almacenados en frío. Configure el número de índices que se deben devolver por página con el parámetro `page_size` (el valor predeterminado es 10). Cada solicitud `_search` aplicada a los índices fríos devuelve un `pagination_id` que se puede utilizar para llamadas posteriores.

La siguiente solicitud realiza la paginación de los resultados de una solicitud `_search` de los índices fríos y muestra los siguientes 100 resultados:

```
GET _cold/indices/_search?page_size=100
{
"pagination_id": "je7MtGbClwBF/2Zp9Utk/h3yCo8nvbEXAMPLEKEY"
}
```

Migración de índices fríos al almacenamiento templado

Tras reducir la lista de índices fríos con los criterios de filtrado de la sección anterior, vuelva a migrarlos a un UltraWarm lugar donde pueda consultar los datos y utilizarlos para crear visualizaciones.

La siguiente solicitud vuelve a migrar dos índices fríos al almacenamiento templado:

```
POST _cold/migration/_warm
 {
 "indices": "my-index1,my-index2"
 }


{
  "acknowledged" : true
}
```

Para verificar el estado de la migración y recuperar el ID de migración, envíe la siguiente solicitud:

```
GET _cold/migration/_status
```

Respuesta de ejemplo

```
{
  "cold_to_warm_migration_status" : [
    {
      "migration_id" : "tyLjXCA-S76zPQbPVHkOKA",
      "indices" : [
        "my-index1,my-index2"
      ],
      "state" : "RUNNING_INDEX_CREATION"
    }
  ]
}
```

Para obtener información de migración específica del índice, incluya el nombre del índice:

```
GET _cold/migration/my-index/_status
```

En lugar de especificar un índice, puede ver una lista de los índices por su estado de migración actual. Los valores válidos son `_failed`, `_accepted` y `_all`.

El siguiente comando obtiene el estado de todos los índices de una única solicitud de migración:

```
GET _cold/migration/_status?migration_id=my-migration-id
```

Recupere el ID de migración mediante la solicitud de estado. Para obtener información detallada sobre la migración, agregue `&verbose=true`.

Puede migrar índices del almacenamiento de frío al almacenamiento templado en lotes de 10 o menos, con un máximo de 100 solicitudes que se migran en forma simultánea. Cualquier solicitud que supere el límite será rechazada. Para comprobar el número actual de migraciones que se están llevando a cabo actualmente, supervise la métrica de `ColdToWarmMigrationQueueSize`. El proceso de migración tiene los siguientes estados:

```
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.
```

Restauración de índices fríos a partir de instantáneas

Si necesita restaurar un índice en frío eliminado, puede restaurarlo de nuevo en el nivel cálido siguiendo las instrucciones que aparecen en el documento the section called “Restauración de índices templados a partir de instantáneas” y, a continuación, migrando de nuevo el índice al nivel frío. No puede restaurar un índice frío eliminado directamente a la capa inactiva. OpenSearch El servicio conserva los índices inactivos durante 14 días después de haberlos eliminado.

Cancelación de migraciones del almacenamiento en frío al templado

Si se pone en cola una migración de índice del almacenamiento en frío al templado o si presenta un estado de error, puede cancelarla mediante la siguiente solicitud:

```
POST _cold/migration/my-index/_cancel
```

```
{
  "acknowledged" : true
}
```

Para cancelar la migración de un lote de índices (máximo 10 a la vez), especifique el ID de migración:

```
POST _cold/migration/_cancel?migration_id=my-migration-id

{
  "acknowledged" : true
}
```

Recupere el ID de migración mediante la solicitud de estado.

Actualización de metadatos de índices almacenados en frío

Puede actualizar los campos `start_time` y `end_time` de un índice almacenado en frío:

```
PATCH _cold/my-index
 {
 "start_time": "2020-01-01",
 "end_time": "2020-02-01"
 }
```

No puede actualizar el `timestamp_field` de un índice de almacenamiento en frío.

> **Note**
>
> OpenSearch Los paneles no admiten el método PATCH. Utilice curl, Postman o algún otro método para actualizar metadatos almacenados en frío.

Eliminación de índices almacenados en frío

Si no utiliza una política de ISM, puede eliminar los índices fríos de forma manual. La siguiente solicitud elimina un índice almacenado en frío:

```
DELETE _cold/my-index
```

```
{
  "acknowledged" : true
}
```

Deshabilitar el almacenamiento en frío

La consola OpenSearch de servicio es la forma más sencilla de deshabilitar el almacenamiento en frío. Seleccione el dominio y elija Acciones, Editar la configuración del clúster, luego anule la selección de Habilitar el almacenamiento en frío).

Para usar la AWS CLI o la API de configuración, en`ColdStorageOptions`, configure`"Enabled"="false"`.

Antes de desactivar el almacenamiento frío, debe eliminar todos los índices fríos o volver a migrarlos al almacenamiento templado; de lo contrario, la acción de desactivación no funcionará.

Index State Management en Amazon OpenSearch Service

La administración del estado de índices (ISM) de Amazon OpenSearch Service te permite definir políticas de administración personalizadas que automatizan las tareas rutinarias y aplicarlas a los índices y patrones de índices. Ya no es necesario configurar y administrar procesos externos para ejecutar las operaciones de índice.

Una política incluye un estado predeterminado y una lista de estados para que el índice pase de un estado al otro. Dentro de cada estado, se puede definir una lista de acciones para realizar y las condiciones que activan estas transiciones. Un caso de uso típico es eliminar periódicamente los índices antiguos después de un determinado periodo de tiempo. Por ejemplo, puede definir una política que mueva el índice a un estado `read_only` después de 30 días y, a continuación, lo elimine después de 90 días.

Después de adjuntar una política a un índice, ISM crea un trabajo que se ejecuta cada 5 a 8 minutos (o cada 30 a 48 minutos en los clústeres de la versión 1.3 y anteriores) para realizar acciones de políticas, verificar condiciones y pasar el índice a estados diferentes. El tiempo base para que este trabajo se ejecute es cada 5 minutos, además de una fluctuación aleatoria del 0 al 60 % para asegurarse de que no se produzca un aumento de la actividad de todos los índices al mismo tiempo. ISM no ejecuta trabajos si el estado del clúster es rojo.

ISM requiere Elasticsearch 6.8 OpenSearch o una versión posterior.

> **Note**
>
> Esta documentación proporciona una breve información general de ISM y varios ejemplos de políticas. También explica en qué se diferencia el ISM para los dominios OpenSearch de Amazon Service del ISM en los OpenSearch clústeres autogestionados. Para obtener la documentación completa sobre ISM, incluida una referencia exhaustiva de los parámetros, las descripciones de cada configuración y una referencia de la API, consulte Index State Management en la OpenSearch documentación.

> **Important**
>
> Ya no se pueden utilizar plantillas de índice para aplicar políticas de ISM a los índices recién creados. Puede seguir administrando automáticamente los índices recién creados con el campo de plantillas de ISM. Esta actualización introduce un cambio importante que afecta a las CloudFormation plantillas existentes que utilizan esta configuración.

Crear una política de ISM

Para empezar a utilizar la administración de estados de índice

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. Seleccione el dominio para el que desea crear una política de ISM.
3. Desde el panel de control del dominio, navega hasta la URL de OpenSearch Dashboards e inicia sesión con tu nombre de usuario y contraseña maestros. La URL tiene este formato:

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

4. Abre el panel de navegación izquierdo en OpenSearch Dashboards y selecciona Administración de índices y, a continuación, Crear política.
5. Utilice el editor visual o el editor de JSON para crear políticas. Se recomienda utilizar el editor visual, ya que ofrece una forma más estructurada de definir las políticas. Para obtener ayuda para crear políticas, consulte las políticas de muestra a continuación.
6. Después de crear una política, puede adjuntarla a uno o más índices:

   ```
   POST _plugins/_ism/add/my-index
   ```

```
{
  "policy_id": "my-policy-id"
}
```

> **Note**
>
> Si su dominio ejecuta una versión de Elasticsearch heredada, utilice `_opendistro` en lugar de `_plugins`.

También puede seleccionar el índice en los OpenSearch paneles y elegir Aplicar política.

Ejemplos de política

Los siguientes ejemplos de política muestran cómo automatizar los casos de uso comunes de ISM.

Traspaso de almacenamiento en caliente a templado y a frío

Este ejemplo de política mueve un índice del almacenamiento activo al almacenamiento activo y UltraWarm, finalmente, a almacenamiento en frío. A continuación, elimina el índice.

El índice se encuentra inicialmente en el estado `hot`. Transcurridos 10 días, ISM lo mueve al estado `warm`. 80 días más tarde, cuando el índice tiene 90 días, ISM lo mueve al estado `cold`. Luego de un año, el servicio envía una notificación a una sala de Amazon Chime que indica que el índice está siendo eliminado y, a continuación, lo elimina permanentemente.

Tenga en cuenta que los índices fríos requieren la operación `cold_delete`, en lugar de la operación `delete` normal. Tenga en cuenta también que es necesario un `timestamp_field` explícito en los datos para administrar índices fríos con ISM.

```
{
  "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": {}
        }]
      }
    ]
  }
}
```

Reducir el recuento de réplicas

Este ejemplo de política reduce el recuento de réplicas a cero, después de siete días, para conservar espacio en disco y, a continuación, elimina el índice después de 21 días. Esta política supone que su índice no es crítico y ya no recibe solicitudes de escritura; tener cero réplicas conlleva cierto riesgo de pérdida de datos.

```
{
  "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": []
      }
    ]
  }
}
```

Tomar una instantánea de índice

Este ejemplo de política utiliza la operación `snapshot` para tomar una instantánea de un índice tan pronto como contenga al menos un documento. `repository` es el nombre del repositorio de instantáneas manual que registró en Amazon S3. `snapshot` es el nombre de la instantánea. Para obtener más información sobre los requisitos previos de una instantánea y los pasos para registrar un repositorio, consulte the section called “Crear instantáneas de índice”.

```
{
  "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": []
      }
    ]
  }
}
```

Plantillas de ISM

Puede configurar un campo `ism_template` en una política, de modo que cuando se cree un índice que coincida con el patrón de plantilla, la política se adjuntará automáticamente a ese índice. En este ejemplo, cualquier índice que cree con un nombre que comience por "registro" se hace coincidir automáticamente con la política de 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
    }
  }
}
```

Para ver un ejemplo más detallado, consulte Ejemplo de política con plantillas de ISM para reversión automática.

Diferencias

En comparación OpenSearch con Elasticsearch, ISM for Amazon OpenSearch Service presenta varias diferencias.

Operaciones ISM

- OpenSearch El servicio admite tres operaciones de ISM únicas `warm_migration`cold_migration, y: `cold_delete`
  - Si su dominio UltraWarmestá habilitado, la `warm_migration` acción convierte el índice en almacenamiento en caliente.
  - Si el dominio tiene habilitado el almacenamiento frío, la acción `cold_migration` pasa el índice al almacenamiento frío, y la acción `cold_delete` elimina el índice del almacenamiento frío.

  Aunque una de estas acciones no se complete dentro del periodo de tiempo de espera establecido, la migración o eliminación de los índices prosigue. Configuración de un error_notification de una de las acciones anteriores, le notificará que la acción falla por no haberse completado dentro del periodo de tiempo de espera, pero la notificación es solo para su propia referencia. La propia operación no tiene un tiempo de espera inherente, y continúa ejecutándose hasta que finalmente se completa correctamente o falla.
- Si tu dominio ejecuta Elasticsearch 7.4 OpenSearch o una versión posterior, el OpenSearch servicio admite el ISM `open` y `close` las operaciones.
- Si tu dominio ejecuta Elasticsearch 7.7 OpenSearch o una versión posterior, OpenSearch Service admite la operación de ISM. `snapshot`

Operaciones de ISM de almacenamiento en frío

Para los índices fríos, debe especificar un `?type=_cold` parámetro cuando utilice el siguiente ISM: APIs

- Agregar política
- Quitar política
- Actualizar política
- Reintentar índice fallido
- Explicar índice

Los índices fríos presentan las siguientes diferencias adicionales: APIs

- Los operadores comodín no son compatibles excepto cuando se utilizan al final. Por ejemplo, se admite `_plugins/_ism/<add, remove, change_policy, retry, explain>/logstash-*`, pero no `_plugins/_ism/<add, remove, change_policy, retry, explain>/iad-*-prod`.
- No se admiten varios nombres y patrones de índices. Por ejemplo, se admite `_plugins/_ism/<add, remove, change_policy, retry, explain>/app-logs`, pero no `_plugins/_ism/<add, remove, change_policy, retry, explain>/app-logs,sample-data`.

Configuración de ISM

OpenSearch y Elasticsearch te permiten cambiar todos los ajustes de ISM disponibles mediante la API. `_cluster/settings` En Amazon OpenSearch Service, solo puedes cambiar la siguiente configuración de ISM:

- Configuración de nivel de clúster:
  - `plugins.index_state_management.enabled`
  - `plugins.index_state_management.history.enabled`
- Configuración de nivel de índice:
  - `plugins.index_state_management.rollover_alias`

Tutorial: Automatización de los procesos de administración del estado del índice

Este tutorial demuestra cómo implementar una política de ISM que automatiza tareas rutinarias de administración de índices y las aplica a índices y patrones de índices.

La administración del estado de los índices (ISM) de Amazon OpenSearch Service te permite automatizar las actividades recurrentes de administración de índices, de modo que no tendrás que utilizar herramientas adicionales para gestionar los ciclos de vida de los índices. Puedes crear una política que automatice estas operaciones en función de la edad del índice, el tamaño y otras condiciones, todo ello desde tu dominio de Amazon OpenSearch Service.

OpenSearch El servicio admite tres niveles de almacenamiento: el estado «activo» predeterminado para la escritura activa y los análisis de baja latencia, UltraWarm para los datos de solo lectura de hasta tres petabytes, y el almacenamiento en frío para el archivado ilimitado a largo plazo.

Este tutorial presenta un caso de uso de ejemplo del manejo de datos de serie temporal en índices diarios. En este tutorial, configurará una política que tome una instantánea automática de cada índice adjunto después de 24 horas. A continuación, migra el índice del estado activo predeterminado al almacenamiento después de dos días, al UltraWarm almacenamiento en frío después de 30 días y, finalmente, elimina el índice después de 60 días.

Requisitos previos

- Tu dominio OpenSearch de servicio debe ejecutar la versión 6.8 o posterior de Elasticsearch.
- Su dominio debe tener habilitado UltraWarmel almacenamiento en frío.
- Debe registrar un repositorio de instantáneas manuales para su dominio.
- Su función de usuario necesita permisos suficientes para acceder a la consola OpenSearch de servicio. Si es necesario, valide y configure el acceso a su dominio.

Paso 1: configurar la política de ISM

En primer lugar, configure una política de ISM en los OpenSearch paneles.

1. Desde el panel de control de su dominio en la consola de OpenSearch servicio, vaya a la URL de OpenSearch Dashboards e inicie sesión con su nombre de usuario y contraseña maestros. La URL tiene este formato: *domain-endpoint*`/_dashboards/`.
2. En los OpenSearch paneles, selecciona Añadir datos de muestra y añade uno o más índices de muestra a tu dominio.
3. Abra el panel de navegación izquierdo y elija Administración de índices y, luego, Crear política.
4. Llame a la política `ism-policy-example`.
5. Reemplace la política predeterminada con la siguiente política:

```
{
  "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**
>
> El campo `ism_template` adjunta automáticamente la política a cualquier índice creado recientemente que coincida con uno de los valores `index_patterns` especificados. En este caso, todos los índices que comiencen por `index-`. Puede modificar este campo para que coincida con un formato de índice de su entorno. Para más información, consulte Plantillas de ISM.

6. En la sección `snapshot` de la política, sustituya *`snapshot-repo`* por el nombre del repositorio de instantáneas que registró para su dominio. También puede optar por sustituir *`ism-snapshot`*, que será el nombre de la instantánea cuando se cree.
7. Seleccione Crear. La política se puede ver ahora en la página Políticas de administración de estados.

Paso 2: adjuntar la política a uno o más índices

Ya que creó la política, puede adjuntarla a uno o más índices de su clúster.

1. Vaya a la pestaña Índices activos y busque `opensearch_dashboards_sample`, donde se enumeran todos los índices de muestra que agregó en el paso 1.
2. Seleccione todos los índices y elija Aplicar política y, a continuación, elija la ism-policy-examplepolítica que acaba de crear.
3. Seleccione Aplicar.

Puede supervisar los índices a medida que pasan por los distintos estados de la página Índices administrados por políticas.

Resumen de índices en Amazon OpenSearch Service con paquetes acumulativos de índices

Los paquetes acumulativos de índices de Amazon OpenSearch Service te permiten reducir los costes de almacenamiento al acumular periódicamente los datos antiguos en índices resumidos.

Seleccione los campos en los que está interesado y utilice una acumulación de índices para crear un nuevo índice con solo los campos agregados en buckets de tiempo más rigurosos. Puede almacenar meses o años de datos históricos a una fracción del costo con el mismo rendimiento de consulta.

Los paquetes acumulativos de índices requieren Elasticsearch 7.9 OpenSearch o una versión posterior.

> **Note**
>
> Esta documentación te ayuda a empezar a crear un trabajo acumulativo de índices en Amazon OpenSearch Service. Para obtener una documentación completa, que incluye una lista de todas las configuraciones disponibles y una referencia completa de la API, consulte los resúmenes de índices en la documentación. OpenSearch

Crear un trabajo acumulativo de índices

Para empezar, elija Administración de índices en OpenSearch los paneles de control. Seleccione Rollup Jobs (Trabajos de acumulación) y elija Create rollup job (Creación de un trabajo de acumulación).

Paso 1: configuración de índices

Configure los índices de origen y destino. El índice de fuente es el que desea resumir. El índice de destino es donde se guardan los resultados de la acumulación de índices.

Después de crear un trabajo de acumulación de índices, no puede cambiar las selecciones de índices.

Paso 2: definir agregaciones y métricas

Seleccione los atributos con las agregaciones (términos e histogramas) y métricas (promedio, suma, máximo, mínimo y recuento de valores) que desea acumular. Asegúrese de no agregar muchos atributos altamente pormenorizados, ya que no ahorrará mucho espacio.

Paso 3: especificar las programaciones

Especifique una programación para acumular los índices a medida que se ingieren. El trabajo de acumulación de índices está habilitado de forma predeterminada.

Paso 4: revisar y crear

Revise su configuración y seleccione Create (Crear).

Paso 5: buscar en el índice objetivo

Puede utilizar la API `_search` estándar para buscar en el índice de destino. No se puede acceder a la estructura interna de los datos en el índice de destino porque el complemento vuelve a escribir automáticamente la consulta en segundo plano para adaptarse al índice de destino. Esto es para asegurarse de que puede utilizar la misma consulta para el índice de fuente y destino.

Para consultar el índice de destino, establezca `size` en 0:

```
GET target_index/_search
{
  "size": 0,
  "query": {
    "match_all": {}
  },
  "aggs": {
    "avg_cpu": {
      "avg": {
        "field": "cpu_usage"
      }
    }
  }
}
```

> **Note**
>
> OpenSearch las versiones 2.2 y posteriores admiten la búsqueda de varios índices acumulativos en una sola solicitud. OpenSearch las versiones anteriores a la 2.2 y las versiones antiguas de Elasticsearch OSS solo admiten un índice acumulativo por búsqueda.

Transformación de los índices en Amazon Service OpenSearch

Mientras que los trabajos acumulativos de índices le permiten reducir la granularidad de los datos acumulando datos antiguos en índices condensados, los trabajos de transformación le permiten crear una vista resumida diferente de sus datos centrada en ciertos campos, de modo que pueda visualizar o analizar los datos de diferentes maneras.

Las transformaciones de índices tienen una interfaz de usuario de OpenSearch Dashboards y una API REST. La función requiere la OpenSearch versión 1.0 o una versión posterior.

> **Note**
>
> Esta documentación proporciona una breve descripción general de las transformaciones de índice para ayudarte a empezar a utilizarlas en un dominio de Amazon OpenSearch Service. Para obtener una documentación completa y una referencia sobre la API REST, consulte las transformaciones de índice en la OpenSearch documentación de código abierto.

Creación de un trabajo de transformación de índice

Si no tiene ningún dato en su clúster, utilice los datos de vuelo de muestra de los OpenSearch paneles para probar las tareas de transformación. Tras añadir los datos, inicie los OpenSearch paneles. A continuación, elija Administración de índices, Trabajos de transformación, y Crear trabajo de transformación.

Paso 1: elección de índices

En la sección Índices, seleccione el índice de origen y destino. Puede seleccionar un índice de destino existente o crear uno nuevo introduciendo un nombre para él.

Si solo quiere transformar un subconjunto del índice de origen, seleccione Añadir filtro de datos y utilice la OpenSearch consulta DSL para especificar un subconjunto del índice de origen.

Paso 2: Elegir los campos

Después de elegir los índices, elija los campos que desea utilizar en el trabajo de transformación, así como si desea utilizar agrupaciones o agregaciones.

- Puede utilizar agrupaciones para colocar los datos en depósitos separados en el índice transformado. Por ejemplo, si desea agrupar todos los destinos de los aeropuertos incluidos

en los datos de vuelo de muestra, agrupe el `DestAirportID` campo en un campo de `DestAirportID_terms` campo objetivo y podrá encontrar el aeropuerto agrupado IDs en el índice transformado una vez finalizado el trabajo de transformación.

- Por otro lado, las agregaciones le permiten realizar cálculos simples. Por ejemplo, puede incluir una agregación en el trabajo de transformación para definir un nuevo campo de `sum_of_total_ticket_price` que calcula la suma de todos los billetes de avión. A continuación, puede analizar los nuevos datos en su índice transformado.

Paso 3: Especificar un programa

Los trabajos de transformación están habilitados de forma predeterminada y se ejecutan en programaciones. Para transform execution interval (intervalo de ejecución de transformación), especifique un intervalo en minutos, horas o días.

Paso 4: Revisar y supervisar

Revise su configuración y seleccione Crear. A continuación, monitoree la columna Estado del trabajo de transformación.

Paso 5: Buscar en el índice objetivo

Una vez finalizado el trabajo, puede usar la API `_search` para buscar en el índice de destino.

Por ejemplo, después de ejecutar un trabajo de transformación que transforma los datos de vuelo basados en el campo `DestAirportID`, puede ejecutar la siguiente solicitud para devolver todos los campos que tienen un valor de `SFO`:

```
GET target_index/_search
{
  "query": {
    "match": {
      "DestAirportID_terms" : "SFO"
    }
  }
}
```

Replicación entre clústeres para Amazon Service OpenSearch

Con la replicación entre clústeres en Amazon OpenSearch Service, puede replicar índices de usuarios, asignaciones y metadatos de un dominio de OpenSearch servicio a otro. Utilizar la replicación entre clústeres contribuye a garantizar la posibilidad de recuperación de desastres en caso de interrupción, y permite replicar datos entre centros de datos lejanos geográficamente para reducir la latencia. Usted paga los cargos de transferencia AWS de datos estándar por los datos transferidos entre dominios.

La replicación entre clústeres sigue un modelo de replicación activa-pasiva en el que el índice local o seguidor extrae datos del índice remoto o principal. El índice líder hace referencia al origen de los datos o al índice desde el que desea replicar los datos. El índice seguidor hace referencia al destino de los datos o al índice en donde desea replicar los datos.

La replicación entre clústeres está disponible en los dominios que ejecutan Elasticsearch 7.10 o 1.1 o OpenSearch versiones posteriores.

> **Note**
>
> Esta documentación describe cómo configurar la replicación entre clústeres desde la perspectiva de Amazon OpenSearch Service. Esto incluye su uso Consola de administración de AWS para configurar conexiones entre clústeres, lo que no es posible en un clúster OpenSearch autogestionado. Para obtener la documentación completa, que incluye una referencia de configuración y una referencia completa de la API, consulte la replicación entre clústeres en la documentación. OpenSearch

Temas

- Limitaciones
- Requisitos previos
- Requisitos de los permisos
- Configuración de una conexión entre clústeres
- Inicio de la replicación
- Confirmación de replicación
- Pausa y reanudación de la replicación
- Detención de la replicación

- Seguimiento automático
- Actualización de los dominios conectados

Limitaciones

La replicación entre clústeres tiene las siguientes limitaciones:

- No puedes replicar datos entre dominios de Amazon OpenSearch Service y clústeres autogestionados OpenSearch o de Elasticsearch.
- No puede replicar un índice de un dominio seguidor a otro dominio seguidor. Si desea replicar un índice en varios dominios seguidores, solo puede replicarlo desde el dominio líder único.
- Se puede conectar un dominio, mediante una combinación de conexiones de entrada y salida, a un máximo de 20 dominios más.
- Al configurar inicialmente una conexión entre clústeres, el dominio principal debe tener la misma versión o una versión superior a la del dominio seguidor.
- No se puede utilizar CloudFormation para conectar dominios.
- No se puede utilizar la replicación entre clústeres en instancias M3 o bursátiles (T2 y T3).
- No se pueden replicar datos entre índices UltraWarm ni inutilizar índices. Ambos índices deben estar almacenados en caliente.
- Al eliminar un índice del dominio principal, no se elimina automáticamente el índice correspondiente en el dominio seguidor.
- No se admite la replicación entre clústeres entre los modos predeterminado y opcional. Ambos dominios deben estar en las regiones predeterminadas o en las regiones opcionales.

Requisitos previos

Antes de configurar la replicación entre clústeres, asegúrese de que los dominios cumplan los siguientes requisitos:

- Elasticsearch 7.10 o 1.1 o posterior OpenSearch
- Control de acceso detallado habilitado
- Node-to-node cifrado activado
- Los índices líderes deben estar `index.soft_deletes.enabled` configurados en`true`. Los índices creados en Elasticsearch 7.0 o OpenSearch 1.0 y versiones posteriores tienen

esta configuración habilitada de forma predeterminada. Sin embargo, los índices creados en Elasticsearch 6.x y luego actualizados se conservan. `soft_deletes=false` Para replicar dichos índices, primero debes volver a indexarlos.

Para comprobar si un índice tiene habilitadas las eliminaciones temporales:

```
GET <index-name>/_settings?
include_defaults=true&flat_settings=true&filter_path=*.settings.index.soft_deletes.enabled
```

Si `soft_deletes` es así`false`, vuelva a indexar los datos en un índice nuevo antes de iniciar la replicación.

Requisitos de los permisos

Para iniciar la replicación, debe incluir el permiso `es:ESCrossClusterGet` en el dominio remoto (líder). Recomendamos la siguiente política de IAM en el dominio remoto. Esta política también permite realizar otras operaciones, como indexar documentos y efectuar búsquedas estándar:

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"
    }
  ]
}
```

Asegúrese de que el permiso `es:ESCrossClusterGet` se aplica para `/leader-domain` y no `/leader-domain/*`.

Para que los usuarios que no son administradores puedan realizar actividades de replicación, también se les deben asignar los permisos adecuados. La mayoría de los permisos corresponden a operaciones de API REST específicas. Por ejemplo, el permiso `indices:admin/plugins/replication/index/_resume` permite reanudar la replicación de un índice. Para obtener una lista completa de los permisos, consulte los permisos de replicación en la OpenSearch documentación.

> **Note**
>
> Los comandos para iniciar la replicación y crear una regla de replicación son casos especiales. Como invocan procesos en segundo plano en los dominios líder y seguidor, debe aprobar la solicitud con una `leader_cluster_role` `follower_cluster_role` autorización. OpenSearch El servicio utiliza estas funciones en todas las tareas de replicación de back-end. Para obtener información sobre la asignación y el uso de estas funciones, consulte Asignar las funciones del clúster de líderes y seguidores en la OpenSearch documentación.

Configuración de una conexión entre clústeres

Para replicar índices de un dominio a otro, se debe configurar una conexión entre clústeres entre los dominios. La forma más sencilla de conectar dominios es a través de la pestaña Connections (Conexiones) del panel del dominio. También puede usar la API de configuración o la AWS CLI. Dado que la replicación entre clústeres sigue un modelo de “extracción”, las conexiones se inician desde el dominio seguidor.

> **Note**
>
> Si se han conectado anteriormente dos dominios para realizar búsquedas entre clústeres, no se puede usar la misma conexión para la replicación. La conexión está marcada como `SEARCH_ONLY` en la consola. Para realizar la replicación entre dos dominios conectados previamente, se debe eliminar la conexión y volver a crearla. Una vez que se haya hecho esto, la conexión queda disponible tanto para la búsqueda entre clústeres como para la replicación entre clústeres.

Para configurar una conexión

1. En la consola de Amazon OpenSearch Service, selecciona el dominio seguidor, ve a la pestaña Conexiones y selecciona Solicitar.
2. En Alias de conexión, ingrese un nombre para la conexión.
3. Elige entre conectarte a un dominio de tu Cuenta de AWS región o a otra cuenta o región.
   - Para conectarte a un dominio de tu Cuenta de AWS región, selecciona el dominio y selecciona Solicitar.
   - Para conectarse a un dominio de otro dominio Cuenta de AWS o región, especifique el ARN del dominio remoto y elija Solicitar.

OpenSearch El servicio valida la solicitud de conexión. Si los dominios son incompatibles, se produce un error en la conexión. Si la validación se realiza correctamente, se envía al dominio de destino para su aprobación. Cuando el dominio de destino aprueba la solicitud, puede iniciar la replicación.

La replicación entre clústeres admite replicación bidireccional. Esto significa que puede crear una conexión saliente del dominio A al dominio B y otra conexión saliente del dominio B al dominio A. A continuación, puede configurar la replicación para que el dominio A siga un índice en el dominio B y el dominio B siga un índice en el dominio A.

Inicio de la replicación

Después de establecer una conexión entre clústeres, puede empezar a replicar datos. En primer lugar, cree un índice en el dominio líder que se va a replicar:

```
PUT leader-01
```

Para replicar ese índice, envíe este comando al dominio seguidor:

```
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"
   }
}
```

Puede encontrar el alias de conexión en la pestaña Conexiones del panel de control del dominio.

Para simplificar, en este ejemplo se supone que un administrador emite la solicitud y utiliza `all_access` para los elementos `leader_cluster_role` y `follower_cluster_role`. No obstante, en entornos de producción, se recomienda crear usuarios de replicación en los índices líder y seguidor, y asignarlos como corresponda. Los nombres de usuario deben ser idénticos. Para obtener información sobre estas funciones y cómo asignarlas, consulte Asignar las funciones del clúster de líderes y seguidores en la OpenSearch documentación.

Confirmación de replicación

Para confirmar que se está produciendo la replicación, obtenga su estado:

```
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
  }
}
```

Los valores de los puntos de control líder y seguidor comienzan como enteros negativos y reflejan el número de particiones que se tienen (-1 para una partición, -5 para cinco particiones, etc.). Los

valores se incrementan en enteros positivos con cada cambio que se realiza. Si los valores son los mismos, eso significa que los índices están totalmente sincronizados. Puede utilizar estos valores de punto de control para medir la latencia de replicación en todos los dominios.

Para validar aún más la replicación, agregue un documento al índice líder:

```
PUT leader-01/_doc/1
{
   "Doctor Sleep":"Stephen King"
}
```

Y confirme que aparece en el índice seguidor:

```
GET follower-01/_search

{
    ...
    "max_score" : 1.0,
    "hits" : [
      {
        "_index" : "follower-01",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 1.0,
        "_source" : {
          "Doctor Sleep" : "Stephen King"
        }
      }
    ]
  }
}
```

Pausa y reanudación de la replicación

Puede pausar temporalmente la replicación si necesita corregir problemas o reducir la carga en el dominio líder. Envíe esta solicitud al dominio seguidor. Asegúrese de incluir un cuerpo de la solicitud vacío:

```
POST _plugins/_replication/follower-01/_pause
{}
```

Después, obtenga el estado para asegurarse de que la replicación esté en pausa:

```
GET _plugins/_replication/follower-01/_status

{
  "status" : "PAUSED",
  "reason" : "User initiated",
  "leader_alias" : "connection-alias",
  "leader_index" : "leader-01",
  "follower_index" : "follower-01"
}
```

Cuando termine de realizar los cambios, reanude la replicación. Envíe esta solicitud al dominio seguidor. Asegúrese de incluir un cuerpo de la solicitud vacío:

```
POST _plugins/_replication/follower-01/_resume
{}
```

No puede reanudar la replicación después de haber estado en pausa durante más de 12 horas. Debe detener la replicación, eliminar el índice de seguidores y reiniciar la replicación del líder.

Detención de la replicación

Cuando se detiene la replicación por completo, el índice seguidor deja de seguir al líder y se convierte en un índice estándar. No se puede reiniciar la replicación después de detenerla.

Detenga la replicación desde el dominio seguidor. Asegúrese de incluir un cuerpo de la solicitud vacío:

```
POST _plugins/_replication/follower-01/_stop
{}
```

Seguimiento automático

Puede definir un conjunto de reglas de replicación en un único dominio líder que replique automáticamente los índices que coincidan con un patrón especificado. Cuando un índice del dominio líder coincide con uno de los patrones (por ejemplo,`books*`), se crea un índice de seguidores coincidente en el dominio de seguidores. OpenSearch El servicio replica todos los índices existentes que coincidan con el patrón, así como los nuevos índices que cree. No replica los índices que ya existan en el dominio seguidor.

Para replicar todos los índices (a excepción de los índices creados por el sistema y los que ya existan en el dominio seguidor), utilice un patrón comodín (*).

Creación de una regla de replicación

Cree una regla de replicación en el dominio seguidor y especifique el nombre de la conexión entre clústeres:

```
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"
   }
}
```

Puede encontrar el alias de conexión en la pestaña Conexiones del panel de control del dominio.

Para simplificar, en este ejemplo se supone que un administrador emite la solicitud, que utiliza `all_access` como reglas de los dominios líder y seguidor. No obstante, en entornos de producción se recomienda crear usuarios de replicación en los índices líder y seguidor, y asignarlos como corresponda. Los nombres de usuario deben ser idénticos. Para obtener información sobre estas funciones y cómo asignarlas, consulte Asignar las funciones del clúster de líderes y seguidores en la documentación. OpenSearch

Para recuperar una lista de reglas de replicación existentes en un dominio, utilice la operación de la API de estadísticas de seguimiento automático.

Para probar la regla, cree un índice que coincida con el patrón del dominio líder:

```
PUT books-are-fun
```

Y verifique que su réplica aparece en el dominio seguidor:

```
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
```

Eliminación de una regla de replicación

Al eliminar una regla de replicación, OpenSearch Service deja de replicar los índices nuevos que coincidan con el patrón, pero continúa con la actividad de replicación existente hasta que se detenga la replicación de esos índices.

Detenga las reglas de replicación desde el dominio seguidor.

```
DELETE _plugins/_replication/_autofollow
{
   "leader_alias" : "connection-alias",
   "name": "rule-name"
}
```

Actualización de los dominios conectados

Para actualizar la versión del motor de dos dominios que tienen una conexión entre clústeres, actualice primero el dominio seguidor y, después, el dominio principal. No elimine la conexión entre ellos; de lo contrario, la replicación se detendrá y no podrá reanudarla.

Migración de índices de Amazon OpenSearch Service mediante la reindexación remota

La reindexación remota te permite copiar índices de un dominio de Amazon OpenSearch Service a otro. Puede migrar índices desde cualquier dominio de OpenSearch servicio o desde clústeres autogestionados OpenSearch y de Elasticsearch.

Un dominio y un índice remotos hacen referencia al origen de los datos o al dominio y el índice desde los que desea copiar los datos. Un dominio y un índice locales hacen referencia al destino de los datos o al dominio y el índice en donde desea copiar los datos.

La reindexación remota requiere la OpenSearch versión 1.0 o una versión posterior, o Elasticsearch 6.7 o una versión posterior, en el dominio local. El dominio remoto debe ser inferior o la misma versión principal que el dominio local. Las versiones de Elasticsearch se consideran inferiores a las versiones, lo que significa que OpenSearch puedes volver a indexar datos de dominios de Elasticsearch a dominios. OpenSearch Dentro de la misma versión principal, el dominio remoto

puede ser cualquier versión secundaria. Por ejemplo, se admite la reindexación remota de Elasticsearch 7.10.x a 7.9, pero no se admite la versión 1.0 a Elasticsearch 7.10.x. OpenSearch

> **Note**
>
> En esta documentación se describe cómo volver a indexar los datos entre los dominios de Amazon OpenSearch Service. Para obtener la documentación completa de la `reindex` operación, incluidos los pasos detallados y las opciones compatibles, consulte el documento de reindexación en la OpenSearch documentación.

Temas

- Requisitos previos
- Reindexe los datos entre los dominios de Internet OpenSearch del Servicio
- Vuelva a indexar los datos entre los dominios de OpenSearch servicio cuando el control remoto esté en una VPC
- Vuelva a indexar los datos entre dominios que no son de servicio OpenSearch
- Reindexar conjuntos de datos grandes
- Configuración remota de reindexación

Requisitos previos

La reindexación remota requiere lo siguiente:

- El dominio remoto debe ser accesible desde el dominio local. Para un dominio remoto que reside en una VPC, el dominio local debe tener acceso a la VPC. Este proceso depende de la configuración de la red, pero probablemente implica conectarse a una VPN o a una red administrada, o utilizar la conexión de punto de conexión de VPC nativa. Para obtener más información, consulte the section called “Compatibilidad con VPC”.
- La solicitud debe autorizarse por el dominio remoto como cualquier otra solicitud REST. Si el dominio remoto tiene habilitado el control de acceso detallado, debe tener permiso para realizar la reindexación en el dominio remoto y leer el índice en el dominio local. Para obtener más consideraciones de seguridad, consulte the section called “Control de acceso detallado”.
- Recomendamos que cree un índice con la configuración deseada en el dominio local antes de comenzar con el proceso de reindexación.

- Si un dominio utiliza un tipo de instancia T2 o T3 para los nodos de datos, no se puede utilizar la reindexación en remoto.

Reindexe los datos entre los dominios de Internet OpenSearch del Servicio

El escenario más básico es que el índice remoto esté en el mismo Región de AWS lugar que su dominio local con un punto final de acceso público y que tenga las credenciales de IAM firmadas.

Desde el dominio remoto, especifique el índice remoto desde el que se va a reindexar y el índice local para reindexar:

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443"
    },
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

Debe agregar 443 al final del punto de conexión del dominio remoto para una comprobación de validación.

Para comprobar que el índice se ha copiado en el dominio local, envíe esta solicitud al dominio local:

```
GET local_index/_search
```

Si el índice remoto se encuentra en una región diferente del dominio local, pase su nombre de región, como en esta solicitud de ejemplo:

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443",
      "region": "eu-west-1"
```

```
    },
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

En el caso de regiones aisladas, como AWS GovCloud (US) las regiones de China, es posible que no se pueda acceder al punto final porque su usuario de IAM no está reconocido en esas regiones.

Si el dominio remoto está protegido con autenticación básica, especifique el nombre de usuario y la contraseña:

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443",
      "username": "username",
      "password": "password"
    },
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

Vuelva a indexar los datos entre los dominios de OpenSearch servicio cuando el control remoto esté en una VPC

Cada dominio de OpenSearch servicio está compuesto por su propia infraestructura interna de nube privada virtual (VPC). Al crear un dominio nuevo en una VPC de OpenSearch servicio existente, se crea una interface de red elástica para cada nodo de datos de la VPC.

Como la operación de reindexación remota se realiza desde el dominio de OpenSearch servicio remoto y, por lo tanto, dentro de su propia VPC privada, necesita una forma de acceder a la VPC del dominio local. Puede hacerlo mediante la función de conexión de punto final de la VPC integrada para establecer una conexión a través AWS PrivateLink de ella o configurando un proxy.

Si tu dominio local usa la OpenSearch versión 1.0 o posterior, puedes usar la consola o la AWS CLI para crear una AWS PrivateLink conexión. Una AWS PrivateLink conexión permite que los recursos de la VPC local se conecten de forma privada a los recursos de la VPC remota dentro de la misma. Región de AWS

Para crear una conexión de punto de conexión de VPC, el dominio de origen que se va a reindexar debe estar en una VPC local y tanto el dominio de origen como el de destino deben estar en la misma Región de AWS.

Vuelva a indexar los datos con la Consola de administración de AWS

Puede usar la reindexación remota con la consola para copiar índices entre dos dominios que comparten una conexión de punto de conexión de VPC.

1. Dirígete a la consola OpenSearch de Amazon Service enhttps://console.aws.amazon.com/aos/.
2. En el panel de navegación izquierdo, seleccione Dominios.
3. Seleccione el dominio local, que es el dominio en el que desea copiar los datos. Se abrirá la página de detalles del dominio. Seleccione la pestaña Conexiones situada debajo de la información general, y luego Solicitar.
4. En la página Solicitar conexión, seleccione Conexión de punto de conexión de VPC como modo de conexión e ingrese otros detalles relevantes. Estos detalles incluyen el dominio remoto, que es el dominio desde el que desea copiar los datos. A continuación, seleccione Solicitar.
5. Vaya a la página de detalles del dominio remoto, seleccione la pestaña Conexiones y busque la tabla Conexiones entrantes. Seleccione la casilla de verificación situada junto al nombre del dominio desde el que acaba de crear la conexión (el dominio local). Seleccione Aprobar.
6. Vuelva al dominio local, seleccione la pestaña Conexiones y busque la tabla Conexiones salientes. Una vez que la conexión entre los dos dominios esté activa, aparecerá un punto de conexión disponible en la columna Punto de conexión de la tabla. Copie el punto de conexión.
7. Abra el panel de control del dominio local y seleccione Herramientas de desarrollo en el panel de navegación de la izquierda. Para confirmar que el índice del dominio remoto aún no existe en el dominio local, ejecute la siguiente solicitud GET. *remote-domain-index-name*Sustitúyalo por tu propio nombre de índice.

```
GET remote-domain-index-name/_search
{
   "query":{
      "match_all":{}
   }
```

```
}
```

En el resultado, debería aparecer un error que indica que no se ha encontrado el índice.

8. Debajo de la solicitud GET, cree una solicitud POST y utilice el punto de conexión como host remoto, tal y como sigue.

```
POST _reindex
{
   "source":{
      "remote":{
         "host":"connection-endpoint",
         "username":"username",
         "password":"password"
      },
      "index":"remote-domain-index-name"
   },
   "dest":{
      "index":"local-domain-index-name"
   }
}
```

Ejecute esta solicitud.

9. Vuelva a ejecutar la solicitud GET. Ahora el resultado debería indicar que el índice local existe. Puede consultar este índice para comprobar que OpenSearch copió todos los datos del índice remoto.

Vuelva a indexar los datos con las operaciones OpenSearch de la API de servicio

Puede usar la reindexación remota con la API para copiar índices entre dos dominios que comparten una conexión de punto de conexión de VPC.

1. Usa la operación de la CreateOutboundConnectionAPI para solicitar una nueva conexión desde tu dominio local a tu 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"
      }
   }
}
```

Recibirá un `ConnectionId` en la respuesta. Guarde este ID para utilizarlo en el siguiente paso.

2. Usa la operación de AcceptInboundConnectionAPI con tu ID de conexión para aprobar la solicitud del dominio local.

```
PUT https://es.region.amazonaws.com/2021-01-01/opensearch/cc/
inboundConnection/ConnectionId/accept
```

3. Usa la operación de DescribeOutboundConnectionsAPI para recuperar el punto final de tu dominio remoto.

```
{
    "Connections": [
        {
            "ConnectionAlias": "remote-reindex-example",
            "ConnectionId": "connection-id",
            "ConnectionMode": "VPC_ENDPOINT",
            "ConnectionProperties": {
                "Endpoint": "connection-endpoint"
            },
            ...
        }
    ]
}
```

Guarda el archivo *connection-endpoint* para usarlo en el paso 5.

4. Para confirmar que el índice del dominio remoto aún no existe en el dominio local, ejecute la siguiente solicitud GET. *remote-domain-index-name*Sustitúyalo por su propio nombre de índice.

```
GET local-domain-endpoint/remote-domain-index-name/_search
{
   "query":{
      "match_all":{}
   }
}
```

   En el resultado, debería aparecer un error que indica que no se ha encontrado el índice.

5. Cree una solicitud POST y utilice el punto de conexión como host remoto, tal y como se muestra a continuación.

```
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"
   }
}
```

   Ejecute esta solicitud.

6. Vuelva a ejecutar la solicitud GET. Ahora el resultado debería indicar que el índice local existe. Puede consultar este índice para comprobar que OpenSearch copió todos los datos del índice remoto.

Si el dominio remoto está alojado dentro de una VPC y no desea utilizar la característica de conexión de punto de conexión de VPC, debe configurar un proxy con un punto de conexión accesible públicamente. En este caso, el OpenSearch servicio requiere un punto final público porque no tiene la capacidad de enviar tráfico a su VPC.

Cuando ejecuta un dominio en modo VPC, se colocan uno o más puntos de conexión en su VPC. Sin embargo, estos puntos de conexión son únicamente para el tráfico que entra en el dominio de la VPC y no permiten que el tráfico entre en la VPC misma.

El comando reindexación remota se ejecuta desde el dominio local, por lo que el tráfico de origen no puede usar esos puntos de conexión para acceder al dominio remoto. Por eso se requiere un proxy en este caso de uso. El dominio proxy debe tener un certificado firmado por una autoridad de certificación (CA) pública. No se admiten los certificados autofirmados o firmados por una entidad de certificación privada.

Vuelva a indexar los datos entre dominios que no son de servicio OpenSearch

Si el índice remoto está alojado fuera del OpenSearch servicio, como en una instancia EC2 autogestionada, defina el `external` parámetro en: `true`

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443",
      "username": "username",
      "password": "password",
      "external": true
    },
    "index": "remote_index"
  },
  "dest": {
    "index": "local_index"
  }
}
```

En este caso, solo se admite la autenticación básica con un nombre de usuario y contraseña. El dominio remoto debe tener un punto final de acceso público (incluso si está en la misma VPC que el dominio de OpenSearch servicio local) y un certificado firmado por una CA pública. No se admiten los certificados autofirmados o firmados por una CA privada.

Reindexar conjuntos de datos grandes

La reindexación remota envía una solicitud de desplazamiento al dominio remoto con los siguientes valores predeterminados:

- Contexto de búsqueda de 5 minutos
- Tiempo de espera del socket de 30 segundos
- Tamaño del lote de 1000

Recomendamos ajustar estos parámetros para adaptarlos a sus datos. En el caso de documentos de gran tamaño, considere la posibilidad de utilizar un lote más pequeño, con un tiempo de espera and/or más largo. Para obtener más información, consulte Paginación de resultados.

```
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"
  }
}
```

También recomendamos agregar la siguiente configuración al índice local para un mejor rendimiento:

```
PUT local_index
{
  "settings": {
    "refresh_interval": -1,
    "number_of_replicas": 0
  }
}
```

Una vez finalizado el proceso de reindexación, puede establecer el recuento de réplicas deseado y eliminar la configuración del intervalo de actualización.

Para reindexar únicamente un subconjunto de documentos que seleccione a través de una consulta, envíe esta solicitud al dominio local:

```
POST _reindex
{
  "source": {
    "remote": {
      "host": "https://remote-domain-endpoint:443"
    },
    "index": "remote_index",
    "query": {
      "match": {
        "field_name": "text"
      }
    }
  },
  "dest": {
    "index": "local_index"
  }
}
```

La reindexación remota no admite la división, por lo que no puede realizar varias operaciones de desplazamiento para la misma solicitud en paralelo.

Configuración remota de reindexación

Además de las opciones de reindexación estándar, OpenSearch Service admite las siguientes opciones:

Opciones	Valores válidos	Description (Descripción)	Obligatorio
externo	Booleano	Si el dominio remoto no es un dominio de OpenSearch servicio o si vas a volver a indexar entre dos dominios de VPC, especifica como. `true`	No

Opciones	Valores válidos	Description (Descripción)	Obligatorio
region	Cadena	Si el dominio remoto se encuentra en otra región, especifique el nombre de la región.	No

Gestión de datos de series temporales en Amazon OpenSearch Service con flujos de datos

Un flujo de trabajo típico para administrar datos de serie temporal consta de varios pasos, como la creación de un alias de índice de reversión, la definición de un índice de escritura y la definición de mapeos y configuraciones comunes para los índices de respaldo.

Los flujos de datos de Amazon OpenSearch Service ayudan a simplificar este proceso de configuración inicial. Data Streams funciona de inmediato para los datos basados en el tiempo, como los registros de aplicaciones, que normalmente son anexos.

Los flujos de datos requieren OpenSearch la versión 1.0 o posterior.

> **Note**
>
> Esta documentación proporciona los pasos básicos para ayudarte a empezar con las transmisiones de datos en un dominio de Amazon OpenSearch Service. Para obtener una documentación completa, consulte Flujos de datos en la OpenSearch documentación.

Introducción a Data Streams

Un flujo de datos se compone internamente de múltiples índices de respaldo. Las peticiones de búsqueda se dirigen a todos los índices de respaldo, mientras que las solicitudes de indexación se dirigen al índice de escritura más reciente.

Paso 1: crear una plantilla de índices

Para crear un flujo de datos, primero debe crear una plantilla de índice que configure un conjunto de índices como un flujo de datos. El objeto `data_stream` indica que se trata de un flujo de datos y no de una plantilla de índice normal. El patrón de índice coincide con el nombre del flujo de datos:

```
PUT _index_template/logs-template
{
  "index_patterns": [
    "my-data-stream",
    "logs-*"
  ],
  "data_stream": {},
  "priority": 100
}
```

En este caso, cada documento capturado debe tener un campo `@timestamp`. También puede definir su propio campo de marca temporal personalizado como una propiedad en el objeto `data_stream`.

```
PUT _index_template/logs-template
{
  "index_patterns": "my-data-stream",
  "data_stream": {
    "timestamp_field": {
      "name": "request_time"
    }
  }
}
```

Paso 2: crear un flujo de datos

Después de crear una plantilla de índice, puede comenzar directamente a incorporar datos sin crear un flujo de datos.

Como tenemos una plantilla de índice que coincide con un `data_stream` objeto, crea OpenSearch automáticamente el flujo de datos:

```
POST logs-staging/_doc
{
  "message": "login attempt failed",
```

```
  "@timestamp": "2013-03-01T00:00:00"
}
```

Paso 3: incorporar datos en el flujo de datos

Para incorporar datos a un flujo de datos, puede utilizar la indexación APIs normal. Asegúrese de que cada documento que indexe tenga un campo temporal. Si intenta incorporar un documento que no tiene un campo temporal, aparecerá un error.

```
POST logs-redis/_doc
{
  "message": "login attempt",
  "@timestamp": "2013-03-01T00:00:00"
}
```

Paso 4: búsqueda de un flujo de datos

Puede buscar en un flujo de datos igual que en un índice normal o en un alias de índice. La operación de búsqueda se aplica a todos los índices de respaldo (todos los datos presentes en el flujo).

```
GET logs-redis/_search
{
  "query": {
    "match": {
      "message": "login"
    }
  }
}
```

Paso 5: revertir un flujo de datos

Puede configurar una política de Administración de estados de índice (ISM) para automatizar el proceso de reversión del flujo de datos. La política de ISM se aplica a los índices de respaldo en el momento de su creación. Cuando se asocia una política a un flujo de datos, solo afecta a los índices de respaldo futuros de ese flujo de datos. Tampoco es necesario proporcionar la configuración `rollover_alias`, ya que la política de ISM deduce esta información del índice de respaldo.

> **Note**
>
> Si migra un índice secundario a un almacenamiento en frío, OpenSearch elimina este índice del flujo de datos. Incluso si vuelve a mover el índice a UltraWarm, el índice permanece independiente y no forma parte del flujo de datos original. Una vez que se haya eliminado un índice del flujo de datos, al buscar en el flujo no se devolverá ningún dato del índice.

> **Warning**
>
> El índice de escritura de un flujo de datos no se puede migrar a un almacenamiento en frío. Si desea migrar los datos de su flujo de datos a un almacenamiento en frío, debe transferir el flujo de datos antes de la migración.

Paso 6: Gestione los flujos de datos en los OpenSearch paneles

Para gestionar los flujos de datos desde los OpenSearch paneles, abra los OpenSearchpaneles, elija Gestión de índices, seleccione Índices o Índices gestionados por políticas.

Paso 7: eliminar un flujo de datos

La operación de eliminación elimina primero los índices de respaldo de un flujo de datos y, a continuación, elimina el propio flujo de datos.

Para eliminar un flujo de datos y todos sus índices de respaldo ocultos:

```
DELETE _data_stream/name_of_data_stream
```

Monitorización de datos en Amazon OpenSearch Service

Supervisa de forma proactiva tus datos en Amazon OpenSearch Service con alertas y detección de anomalías. Configure alertas para recibir notificaciones cuando los datos superan umbrales determinados. La detección de anomalías utiliza el machine learning para detectar automáticamente cualquier valor atípico en los datos de streaming. Puede emparejar la detección de anomalías con las alertas para asegurarse de que se le notifique tan pronto como se detecte una anomalía.

Temas

- Configuración de alertas en Amazon OpenSearch Service
- Detección de anomalías en Amazon Service OpenSearch

Configuración de alertas en Amazon OpenSearch Service

Configura alertas en Amazon OpenSearch Service para recibir notificaciones cuando los datos de uno o más índices cumplan determinadas condiciones. Por ejemplo, es posible que desee recibir un email si la aplicación registra más de cinco errores HTTP 503 en una hora o tal vez desee avisar a un desarrollador si no se indexaron nuevos documentos en los últimos 20 minutos.

Las alertas requieren Elasticsearch OpenSearch 6.2 o una versión posterior.

> **Note**
>
> Esta documentación proporciona una breve descripción general de las alertas y destaca en qué se diferencian las alertas en un dominio de Amazon OpenSearch Service de las alertas en un clúster de código abierto. OpenSearch Para obtener la documentación completa sobre las alertas, que incluye una referencia exhaustiva de la API, una lista de los campos de solicitud disponibles para los monitores compuestos y una descripción de las variables de activación y acción disponibles, consulte la sección Alertas en la documentación. OpenSearch

Temas

- Permisos de alertas
- Introducción a las alertas

- Notificaciones
- Diferencias

Permisos de alertas

Las alertas admiten el control de acceso detallado. Para obtener más información sobre cómo combinar y combinar permisos para adaptarlos a su caso de uso, consulte la seguridad de las alertas en la documentación. OpenSearch

Para acceder a la página de alertas en los OpenSearch paneles de control, debe estar asignado como mínimo a la función `alerting_read_access` predefinida o tener permisos equivalentes. Este rol concede permisos para ver alertas, destinos y monitores, pero no para reconocer alertas ni modificar destinos o monitores.

Introducción a las alertas

Para crear una alerta, debe configurar un monitor, que es un trabajo que se ejecuta según un cronograma definido y consulta los índices. OpenSearch También puede configurar disparadores, que definen las condiciones que generan los eventos. Por último, se configuran las acciones, lo que ocurre después de que se activa una alerta.

Para comenzar a trabajar con alertas

1. Elija Alertas en el menú principal del OpenSearch panel de mandos y elija Crear monitor.
2. Cree un monitor por consulta, por bucket, por métricas de clúster o por documento. Para obtener instrucciones, consulte Create a monitor (Crear un monitor).
3. En Triggers (Disparadores), cree uno o más disparadores. Para obtener instrucciones, consulte Create triggers (Creación de disparadores).
4. En Actions (Acciones), configure un notification channel (canal de notificación) para la alerta. Seleccione entre Slack, Amazon Chime, un webhook personalizado o Amazon SNS. Como puede imaginar, las notificaciones requieren conectividad con el canal. Por ejemplo, para notificar a un canal de Slack o enviar un webhook personalizado a un servidor de terceros, tu dominio de OpenSearch servicio necesita una conectividad de red adecuada. El webhook personalizado debe tener una dirección IP pública para que un dominio de OpenSearch servicio pueda enviarle alertas.

> **Tip**
>
> Después de que una acción envíe correctamente un mensaje, tiene la responsabilidad de asegurar el acceso a ese mensaje (por ejemplo, el acceso a un canal de Slack). Si el dominio contiene información confidencial, considere utilizar desencadenadores sin acciones y verificar periódicamente si hay alertas en Dashboards.

Notificaciones

Las alertas se integran con las notificaciones, que es un sistema unificado de OpenSearch notificaciones. Las notificaciones le permiten configurar el servicio de comunicación que desea utilizar y ver las estadísticas pertinentes y la información para la solución de problemas. Para obtener una documentación completa, consulte Notificaciones en la OpenSearch documentación.

Tu dominio debe tener la OpenSearch versión 2.3 o posterior para poder usar las notificaciones.

> **Note**
>
> OpenSearch las notificaciones son independientes de las notificaciones de OpenSearch servicio, que proporcionan detalles sobre las actualizaciones del software del servicio, las mejoras de Auto-Tune y otra información importante a nivel de dominio. OpenSearch las notificaciones son específicas del complemento.

Los canales de notificación sustituyeron a los destinos de alerta a partir de la versión 2.0. OpenSearch Los destinos quedaron obsoletos de forma definitiva y, a partir de ahora, todas las notificaciones de alertas se gestionarán a través de los canales.

Cuando actualizas tus dominios a la versión 2.3 o posterior (dado que el soporte de OpenSearch servicio para la versión 2.x comienza con la 2.3), tus destinos actuales se migran automáticamente a los canales de notificación. Si un destino no se migra, el monitor seguirá usándolo hasta que el monitor se migre a un canal de notificación. Para obtener más información, consulta la sección Preguntas sobre los destinos en la OpenSearch documentación.

Para empezar con las notificaciones, inicia sesión en los OpenSearch paneles de control y selecciona Notificaciones, Canales y Crear canal.

Amazon Simple Notification Service (Amazon SNS) es un tipo de canal compatible para las notificaciones. Para autenticar a los usuarios, debe proporcionar al usuario acceso completo a Amazon SNS o dejar que asuma un rol de IAM con permisos para acceder a Amazon SNS. Para obtener instrucciones, consulte Amazon SNS as a channel type (Amazon SNS como un tipo de canal).

Diferencias

En comparación con la versión de código abierto de OpenSearch, las alertas en Amazon OpenSearch Service presentan algunas diferencias notables.

Configuración de alertas

OpenSearch El servicio te permite modificar la siguiente configuración de alertas:

- `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`

Todas las demás configuraciones utilizan los valores predeterminados que no se pueden cambiar.

Para deshabilitar las alertas, envíe la siguiente solicitud:

```
PUT _cluster/settings
{
  "persistent" : {
    "plugins.scheduled_jobs.enabled" : false
  }
}
```

La siguiente solicitud configura las alertas para eliminar automáticamente los índices del historial después de siete días, en lugar de los 30 días predeterminados:

```
PUT _cluster/settings
{
```

```
  "persistent": {
    "plugins.alerting.alert_history_retention_period": "7d"
  }
}
```

Si previamente creó monitores y desea detener la creación de índices de alerta diaria, elimine todos los índices del historial de alertas:

```
DELETE .plugins-alerting-alert-history-*
```

Para reducir el recuento de particiones de los índices del historial, cree una plantilla de índice. La siguiente solicitud establece índices de historial para las alertas y administración de estados de índice para una partición y una réplica:

```
PUT _index_template/template-name
{
  "index_patterns": [".opendistro-alerting-alert-history-*"],
  "template": {
    "settings": {
      "number_of_shards": 1,
      "number_of_replicas": 1
    }
  }
}
```

En función de la tolerancia a la pérdida de datos, incluso podría considerar no utilizar ninguna réplica. Para obtener más información sobre la creación y administración de plantillas de índice, consulte las plantillas de índice en la OpenSearch documentación.

Detección de anomalías en Amazon Service OpenSearch

La detección de anomalías en Amazon OpenSearch Service detecta automáticamente las anomalías en los OpenSearch datos casi en tiempo real mediante el algoritmo Random Cut Forest (RCF). RCF es un algoritmo de machine learning no supervisado que modela un esquema del flujo de datos entrante. El algoritmo calcula un `anomaly grade` y valor `confidence score` para cada punto de datos entrante. La detección de anomalías utiliza estos valores para diferenciar una anomalía de las variaciones habituales de los datos.

Puede emparejar el complemento de detección de anomalías con el complemento Alertas para recibir una notificación en cuanto se detecte una anomalía.

La detección de anomalías está disponible en los dominios que ejecutan cualquier OpenSearch versión de Elasticsearch 7.4 o posterior. Todos los tipos de instancias admiten la detección de anomalías, excepto `t2.micro` y `t2.small`.

> **Note**
>
> Esta documentación proporciona una breve descripción de la detección de anomalías en el contexto de Amazon OpenSearch Service. Para obtener una documentación completa, que incluye los pasos detallados, una referencia de la API, una referencia de todos los ajustes disponibles y los pasos para crear visualizaciones y paneles, consulte la detección de anomalías en la documentación de código abierto. OpenSearch

Requisitos previos

La detección de anomalías posee los siguientes requisitos previos:

- La detección de anomalías requiere Elasticsearch 7.4 OpenSearch o una versión posterior.
- La detección de anomalías solo admite un control de acceso detallado en las versiones 7.9 y posteriores de Elasticsearch y en todas las versiones de. OpenSearch En las versiones anteriores a Elasticsearch 7.9, solo los usuarios administradores pueden crear, ver y administrar detectores.
- Si su dominio utiliza un control de acceso detallado, los usuarios que no sean administradores deben asignarse a ese `anomaly_read_access` rol en los OpenSearch paneles para poder ver los detectores o para crearlos y administrarlos. `anomaly_full_access`

Introducción a la detección de anomalías

Para empezar, elija Detección de anomalías en los paneles. OpenSearch

Paso 1: crear un detector

Un detector es una tarea individual de detección de anomalías. Puede crear varios detectores y todos los detectores se pueden ejecutar simultáneamente; cada uno podrá analizar datos de distintos orígenes.

Paso 2: agregar características a su detector

Una característica es el campo del índice que se verifica para ver si hay anomalías. Un detector puede detectar anomalías en una o varias características. Debe elegir una de las siguientes agrupaciones para cada característica: `average()`, `sum()`, `count()`, `min()` o `max()`.

> **Note**
>
> El método de `count()` agregación solo está disponible en OpenSearch Elasticsearch 7.7 o versiones posteriores. Para Elasticsearch 7.4, utilice una expresión personalizada como la siguiente:
>
> ```
> {
>   "aggregation_name": {
>      "value_count": {
>         "field": "field_name"
>      }
>   }
> }
> ```

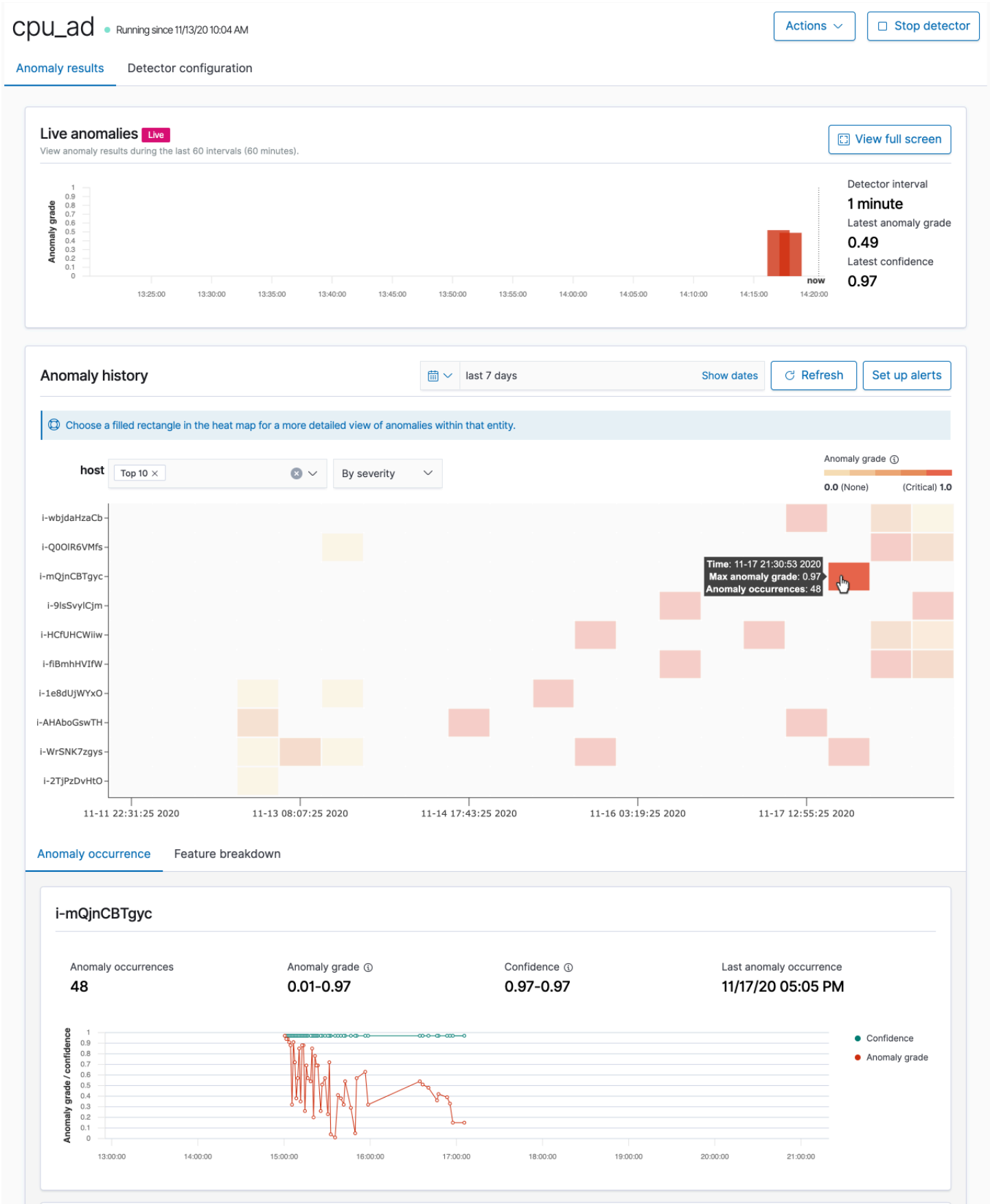
El método de agrupación determina qué constituye una anomalía. Por ejemplo, si elige `min()`, el detector se centra en buscar anomalías basadas en los valores mínimos de la característica. Si elige `average()`, el detector busca anomalías basadas en los valores medios de la característica. Puede agregar un máximo de cinco características por detector.

Puede configurar los siguientes parámetros opcionales (disponibles en Elasticsearch 7.7 y posterior):

- Category field (Categoría): clasifique o corte los datos con una dimensión como dirección IP, ID del producto, código de país, etc.
- Window size (Tamaño de ventana): establezca el número de intervalos de suma del flujo de datos que se tendrán en cuenta en una ventana de detección.

Después de configurar las características, obtenga una previsualización de las muestras de anomalías y ajuste la configuración de las características, si es necesario.

Paso 3: observar los resultados



- Live anomalies (Anomalías activas): muestra los resultados de anomalías activas de los últimos 60 intervalos. Por ejemplo, si el intervalo se establece en 10, muestra los resultados de los últimos 600 minutos. Este gráfico se actualiza cada 30 segundos.
- Anomaly history (Historial de anomalías): traza el grado de anomalía con la medida de confianza correspondiente.
- Feature breakdown (Desglose de características): traza las características según el método de agrupación. Puede variar el rango de fecha y hora del detector.
- Anomaly occurrence (Ocurrencia de anomalías): muestra los valores `Start time`, `End time`, `Data confidence` y `Anomaly grade` de cada anomalía detectada.

  Si establece el campo de categoría, verá un grafico Heat map (Mapa térmico) que correlaciona los resultados de entidades anómalas. Seleccione un rectángulo relleno para ver una vista más detallada de la anomalía.

Paso 4: configurar alertas

Para crear un monitor que envíe notificaciones cuando se detectan anomalías, elija Set up alerts (Configurar alertas). El complemento lo redirige a la página Add monitor (Agregar monitor) donde puede configurar una alerta.

Tutorial: Detectar un uso elevado de la CPU con la detección de anomalías

En este tutorial se muestra cómo crear un detector de anomalías en Amazon OpenSearch Service para detectar un uso elevado de la CPU. Utilizará los OpenSearch paneles de control para configurar un detector que supervise el uso de la CPU y generará una alerta cuando el uso de la CPU supere un umbral específico.

> **Note**
>
> Estos pasos se aplican a la versión más reciente OpenSearch y pueden variar ligeramente en las versiones anteriores.

Requisitos previos

- Debes tener un dominio de OpenSearch servicio que ejecute Elasticsearch 7.4 o una versión posterior, o cualquier OpenSearch versión.

- Debe ingerir archivos de registro de aplicaciones en el clúster que contengan datos de uso de la CPU.

Paso 1: crear un detector

Primero, cree un detector que identifique las anomalías en los datos de uso de la CPU.

1. Abre el menú del panel izquierdo en OpenSearch Dashboards y selecciona Detección de anomalías y, a continuación, selecciona Crear detector.
2. Asigne el nombre **`high-cpu-usage`** al detector.
3. Para el origen de datos, elija el índice que contenga los archivos de registro de uso de la CPU en los que desee identificar anomalías.
4. Seleccione el icono Timestamp field (Campo de marca temporal) de los datos. Opcionalmente, puede agregar un filtro de datos. Este filtro de datos analiza únicamente un subconjunto del origen de datos y reduce el ruido de los datos que no son relevantes.
5. Establezca Detector interval (Intervalo de detector) en 2 minutos. Este intervalo define el tiempo (por intervalo de minutos) que tendrá el detector para recopilar los datos.
6. En Window delay (Demora del plazo), agregue una demora de 1 minute (1 minuto). Esta demora agrega tiempo de procesamiento adicional para garantizar que todos los datos dentro del plazo estén presentes.
7. Seleccione Next (Siguiente). En el panel de detección de anomalías, en el nombre del detector, seleccione Configure model (Configurar modelo).
8. En Feature name (Nombre de la característica), ingrese **`max_cpu_usage`**. En Feature state (Estado de la característica), seleccione Enable feature (Habilitar característica).
9. En Find anomalies based on (Encontrar anomalías en función de), elija Field value (Valor de campo).
10. En Aggregation method (Método de agregación), elija **`max()`**.
11. En Field (Campo), seleccione el campo de los datos para ver si hay anomalías. Por ejemplo, podría llamarse `cpu_usage_percentage`.
12. Mantenga todas las demás configuraciones en sus valores predeterminados y elija Next (Siguiente).
13. Ignore la configuración de trabajos del detector y elija Next (Siguiente).
14. En la ventana emergente, elija cuándo desea iniciar el detector (automática o manualmente) y, a continuación, elija Confirm (Confirmar).

Ya que se configuró el detector, después de inicializarse, podrá ver los resultados en tiempo real del uso de la CPU en la sección Real-time results (Resultados en tiempo real) del panel del detector. La sección Live anomalies (Anomalías en vivo) muestra cualquier anomalía que se produce cuando los datos se ingieren en tiempo real.

Paso 2: configurar una alerta

Ya que se creó un detector, cree un monitor que invoque una alerta para enviar un mensaje a Slack cuando se detecte un uso de la CPU que cumpla con las condiciones especificadas en la configuración del detector. Recibirá notificaciones de Slack cuando los datos de uno o más índices cumplan con las condiciones que invocan la alerta.

1. Abre el menú del panel izquierdo en los OpenSearch paneles de control, selecciona Alertas y, a continuación, selecciona Crear monitor.
2. Proporcione un nombre para el monitor.
3. En Monitor type (Tipo de monitor), elija Per-query monitor (Monitor por consulta). Un monitor por consulta ejecuta una consulta específica y define los activadores.
4. En Monitor defining method (Método de definición de monitor), elija Anomaly detector (Detector de anomalías) y, a continuación, seleccione el detector que ha creado en la sección anterior en el menú desplegable Detector.
5. En Schedule (Programación), elija la frecuencia con la que el monitor recopilará datos y recibirá alertas. Para los fines de este tutorial, establezca la programación para que se ejecute cada 7 minutos.
6. En la sección Triggers (Desencadenadores), elija Add trigger (Agregar desencadenador). En Trigger name (Nombre del desencadenador), ingrese **`High CPU usage`**. En este tutorial, en Severity level (Nivel de gravedad), elija 1, que es el nivel de gravedad más alto.
7. En Anomaly grade threshold (Umbral de clasificación de anomalías), elija IS ABOVE (ES SUPERIOR). En el menú que hay debajo, elija el umbral de clasificación que desea aplicar. En este tutorial, establezca Clasificación de anomalías en 0,7.
8. En Anomaly confidence threshold (Umbral de confianza de anomalías), elija IS ABOVE (ES SUPERIOR). En el menú que hay debajo, ingrese el mismo número que su clasificación de anomalías. En este tutorial, establezca Anomaly confidence threshold (Umbral de confianza de anomalías) en 0,7.
9. En la sección Actions (Acciones), elija Destination (Destino). En el campo Name (Nombre), elija el nombre del destino. En el menú Type (Tipo), elija Slack. En el campo Webhook URL (URL del webhook), ingrese una URL de webhook donde quiera recibir alertas. Para obtener más

información, consulte Sending messages using incoming webhooks (Envío de mensajes mediante webhooks entrantes).

10. Seleccione Create (Crear).

Recursos relacionados

- the section called “Alertas”

- the section called “Detección de anomalías”
- API de detección de anomalías

Asistente de IA para Amazon OpenSearch Service

AI Assistant es un asistente generativo basado en inteligencia artificial que ayuda a los desarrolladores y profesionales de TI a realizar diversas tareas a lo largo del ciclo de vida del desarrollo del software. AI Assistant ayuda a AWS los clientes a codificar, probar, implementar, solucionar problemas y optimizar sus aplicaciones en ejecución. AWS

La integración de AI Assistant con Amazon OpenSearch Service ofrece las siguientes capacidades generativas:

- the section called “Genere visualizaciones utilizando un lenguaje natural”
- the section called “Ver resúmenes e información sobre alertas”
- the section called “Consulta los resúmenes de los resultados de las consultas generados por AI Assistant en la página Discover”
- the section called “Vea los detectores de anomalías recomendados”
- the section called “Accede al chat de AI Assistant si tienes preguntas sobre el OpenSearch servicio”

Para acceder a las funciones de AI Assistant de OpenSearch Service que sean relevantes para su tarea, busque el siguiente icono contextual en las páginas de visualización Alerts, Discover y Create. Haga clic en el ícono para la petición de asistencia y utilizar las características descritas en esta sección.

También puedes chatear directamente con AI Assistant haciendo clic en el icono situado en la esquina superior derecha de la consola de OpenSearch servicio. El chat de AI Assistant permite responder a preguntas sobre las características y la funcionalidad del OpenSearch Servicio.

Compatible Regiones de AWS

El soporte de AI Assistant para el OpenSearch servicio está disponible en los siguientes lugares Regiones de AWS:

- Este de EE. UU. (Norte de Virginia)
- Oeste de EE. UU. (Oregón)
- Asia-Pacífico (Mumbai)

- Asia-Pacífico (Sídney)
- Asia-Pacífico (Tokio)
- Canadá (centro)
- Europa (Fráncfort)
- Europa (Londres)
- Europa (París)
- América del Sur (São Paulo)

Configuración de AI Assistant para el OpenSearch servicio

Complete los siguientes pasos para configurar AI Assistant for OpenSearch Service.

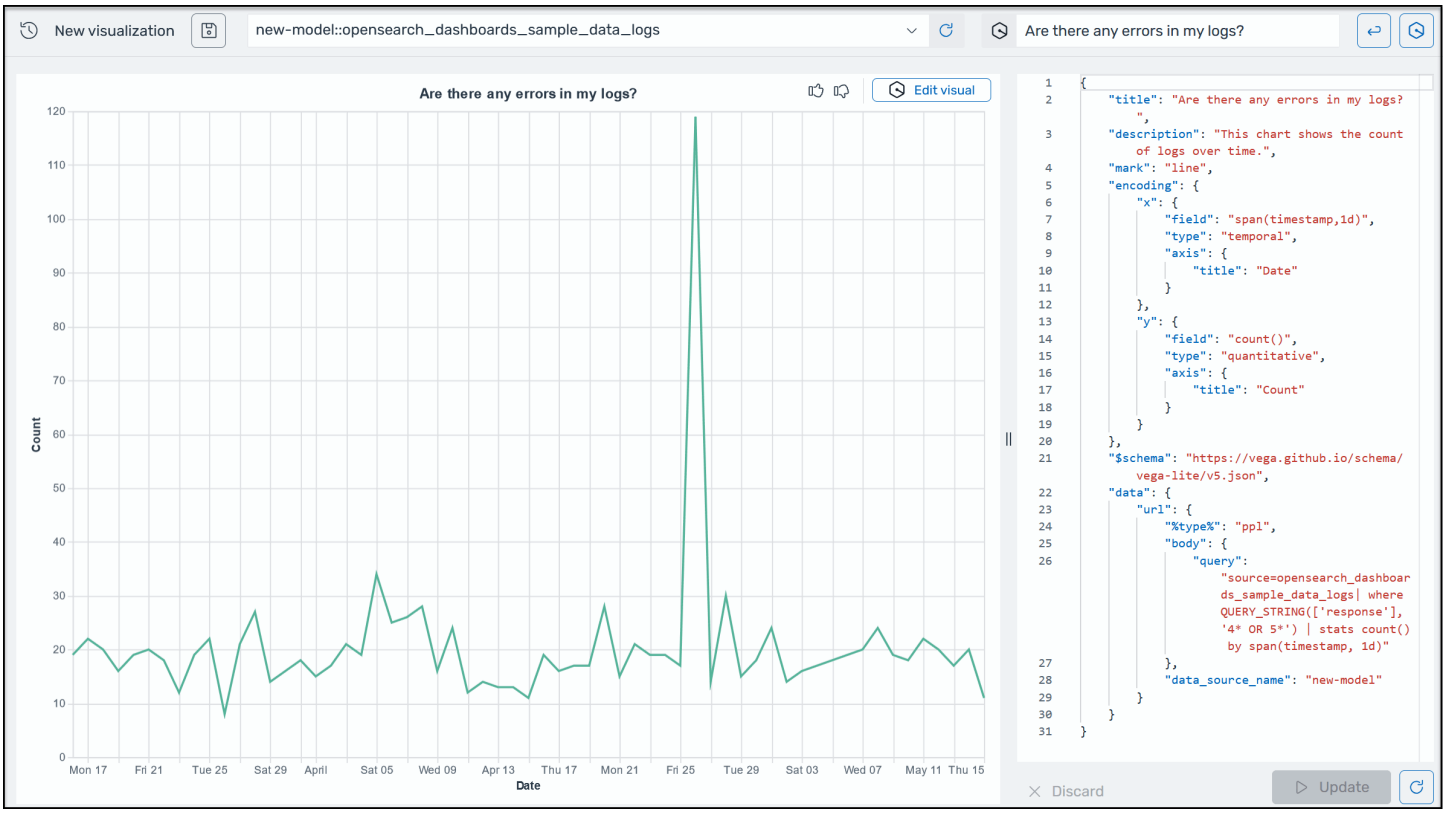
1. Compruebe que ha configurado el control OpenSearch de acceso detallado del Servicio. Para obtener más información, consulte Control de acceso detallado en Amazon Service OpenSearch.
2. Compruebe que la fuente de datos esté en la OpenSearch versión 2.17 o posterior.
3. Verifique haber seleccionado la casilla Habilitar generación de consulta en lenguaje natural en la sección Inteligencia artificial (IA) y machine learning (ML) durante la creación del dominio o editando la configuración del clúster.

> **Note**
>
> OpenSearch Las funciones de Amazon Q for Developer están disponibles con la capa gratuita Q. Para obtener más información, consulte Comprensión de los niveles de servicio en la Guía del usuario de Amazon Q.

Genere visualizaciones utilizando un lenguaje natural

Para ayudarlo a obtener más información sobre sus datos operativos, AI Assistant for OpenSearch Service admite el uso de indicaciones en lenguaje natural para crear visualizaciones. Antes de empezar, compruebe que ha configurado AI Assistant for OpenSearch Service. Cuando lo confirme, inicie la aplicación de interfaz de usuario del OpenSearch servicio, vaya a la página de visualizaciones, seleccione el botón + Crear visualización y elija un lenguaje natural. Ahora

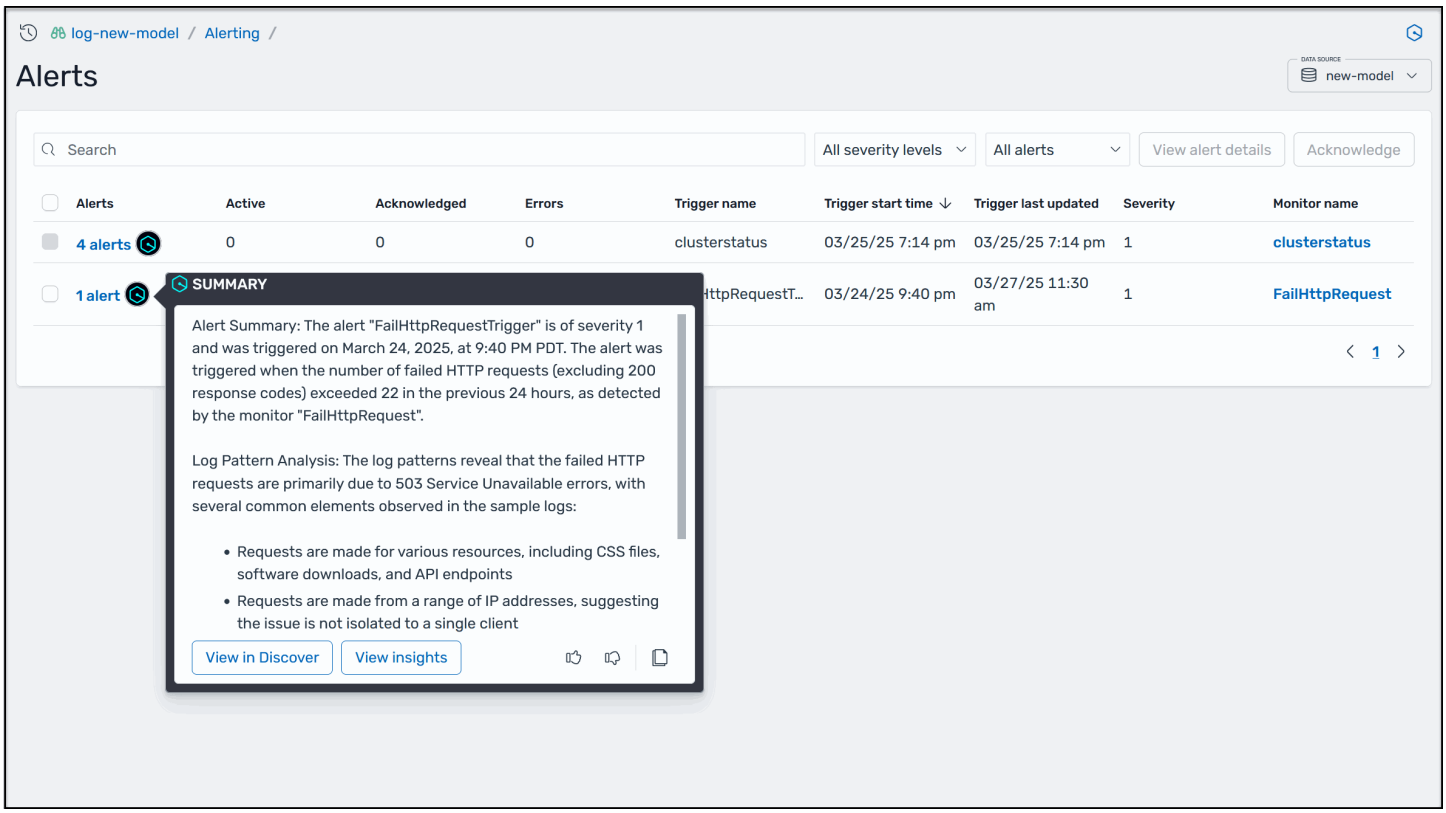
puede especificar una fuente de datos e introducir un requisito en lenguaje natural para generar las visualizaciones. A continuación se muestra un ejemplo.



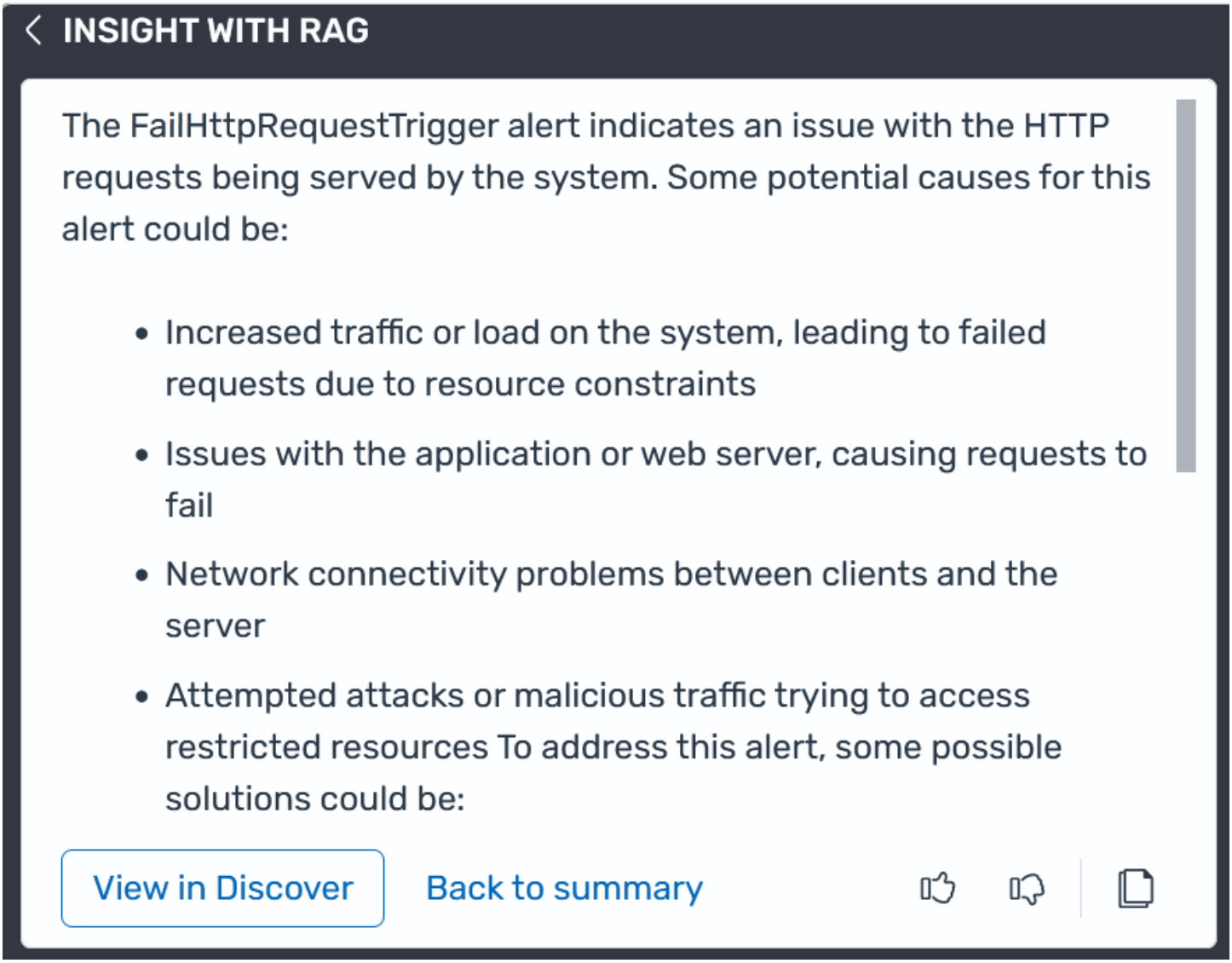
Las visualizaciones pueden agilizar la resolución de problemas al desglosar el análisis de errores según diferentes dimensiones. También pueden ayudarlo a identificar patrones y tendencias, acelerar la toma de decisiones, revelar relaciones y simplificar datos complejos. Como alternativa, también puedes crear visualizaciones en la página Discover como parte de los resultados de la consulta.

Ver resúmenes e información sobre alertas

Puede configurar el OpenSearch servicio para crear un monitor de alertas cuando los datos de uno o más índices cumplan determinadas condiciones. Para ayudarte a entender y solucionar rápidamente una alerta, puedes ver un resumen de la alerta haciendo clic en el icono de AI Assistant situado junto a la alerta. Un resumen proporciona detalles sobre el problema subyacente que activó la alerta y, cuando está disponible, un análisis adicional para ayudarlo a localizar la causa raíz del problema. En la siguiente captura de pantalla se muestra un ejemplo de un resumen de alertas creado por AI Assistant.



Si conectas una base de conocimientos para proporcionar un contexto adicional sobre tu entorno, como se describe más adelante en este tema, AI Assistant crea información sobre una alerta. Los conocimientos proporcionan detalles y opciones de solución de problemas para ayudarlo a corregir la causa raíz de una alerta. Para la alerta mostrada anteriormente, AI Assistant también obtuvo la siguiente información.



> **Note**
>
> En función de la naturaleza de la alerta y de la información disponible, AI Assistant puede ofrecerte la opción de ver los datos de la alerta en la página Descubrir de los OpenSearch paneles de control. Si ves el botón Ver en Discover en la parte inferior de un resumen de alertas de AI Assistant, haz clic en el botón para abrir el conjunto de datos correspondiente en Discover con un filtro activo para los datos de las alertas.

Temas

- Antes de empezar
- Visualización de resúmenes de alertas y conocimientos

Antes de empezar

Complete los siguientes pasos para configurar una base de conocimiento de Amazon Bedrock para que AI Assistant pueda crear información para las alertas de OpenSearch servicio.

Paso 1: Cree el LambdaInvokeOpenSearch MLCommons rol (rol de IAM)

Cree un nuevo rol con el nombre `LambdaInvokeOpenSearchMLCommonsRole` de AWS Identity and Access Management (IAM). OpenSearch El servicio utiliza esta función para crear un conector de IA OpenSearch que ayude a generar información basada en los artículos configurados de la base de conocimientos. Debe asignar esta función a la función de OpenSearch servicio`ml_full_access`, tal y como se describe en el paso 2.

Cuando cree el rol nuevo, para Tipo de entidad de confianza, seleccione cuenta de AWS . No necesita especificar una política de permisos. En la página Agregar permisos, elija Siguiente. Para obtener más información sobre la creación de una nueva función, consulte Creación de una función para un AWS servicio (consola).

Paso 2: Asigne el LambdaInvokeOpenSearch MLCommons rol del rol al rol ml_full_access del OpenSearch servicio

Utilice el siguiente procedimiento para asignar la función a la `LambdaInvokeOpenSearchMLCommonsRole` función de servicio. OpenSearch `ml_full_access` Este mapeo también ayuda a OpenSearch Service a crear el conector AI.

Para asignar la función de IAM requerida a la función OpenSearch ml_full_access del servicio

1. Abra la página de administración de datos de OpenSearch Service Dashboard.
2. En Acceso a datos y usuario, elija Roles.
3. Utilice el cuadro de búsqueda para localizar el rol `ml_full_access`.
4. En la página ml_full_access, seleccione la pestaña Usuarios asignados.
5. Elija Asignar usuarios.
6. En el campo Roles de backend, pegue el Nombre de recurso de Amazon (ARN) del rol `LambdaInvokeOpenSearchMLCommonsRole` y luego seleccione Asignar.

Paso 3: Configure una base OpenSearch de conocimientos del servicio mediante CloudFormation

Utilice el siguiente procedimiento para configurar una base de conocimientos del OpenSearch servicio de AWS CloudFormation forma que AI Assistant pueda generar información.

Para configurar una base de conocimientos para generar conocimientos:

1. Inicia sesión en la página de https://console.aws.amazon.com/aos/inicio de la consola de Amazon OpenSearch Service de forma compatible Región de AWS. Para obtener más información, consulte Compatible Regiones de AWS.
2. En el panel de navegación, elija Integraciones.
3. En la sección Plantillas de integración, elija la plantilla Integrar con base de conocimientos mediante Amazon Bedrock. Si no ve esta plantilla, verifique que se encuentre en una Región compatible.
4. En el mosaico Integrar con la base de conocimientos mediante Amazon Bedrock, elija Configurar dominio y, a continuación, elija una de las opciones disponibles. OpenSearch El servicio abre la plantilla de CloudFormation pila con los campos obligatorios rellenados previamente. La CloudFormation pila admite la integración de dominios públicos y de VPC.
5. Seleccione Creación de pila. Una CloudFormation vez creados los recursos, el servicio muestra el agente AgentIdde Amazon Bedrock y ModelId. ConnectorId

Cuando corresponde, AI Assistant ahora crea información para las alertas OpenSearch del servicio.

Visualización de resúmenes de alertas y conocimientos

Utilice el siguiente procedimiento para ver los resúmenes de las alertas y la información valiosa en OpenSearch Service.

Visualización de resúmenes de alertas y conocimientos

1. Comprueba que has configurado AI Assistant for OpenSearch Service.
2. Comprueba que has configurado las alertas para el OpenSearch servicio.
3. En el menú principal del OpenSearch panel de mandos, elija Alertas y, a continuación, Alertas.
4. Selecciona el icono del Asistente de IA situado junto a la alerta. El Asistente de IA puede tardar hasta 10 segundos en generar el resumen.

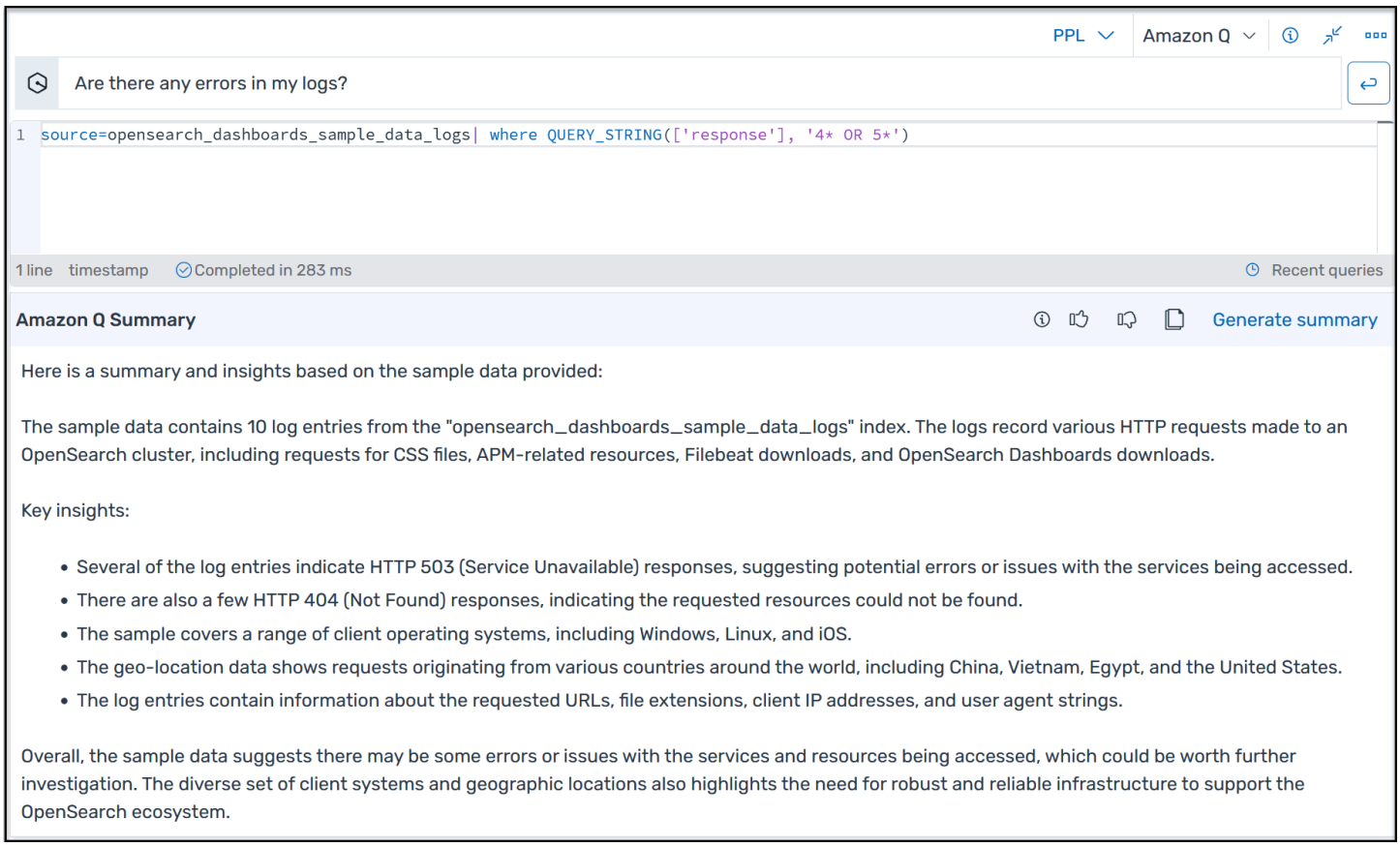
5. Si aparece en el resumen de alertas, elija Ver información para ver más detalles sobre la alerta según su base de conocimiento configurada.
6. Si aparecen en el resumen de la alerta, seleccione Ver en Discover para ver los datos de la alerta en la página Discover de los OpenSearch paneles de control.

Consulta los resúmenes de los resultados de las consultas generados por AI Assistant en la página Discover

OpenSearch El servicio le permite consultar sus datos con indicaciones en lenguaje natural utilizando el lenguaje de procesamiento canalizado (PPL) de la página Discover. Por ejemplo, usted puede escribir consultas como las siguientes:

- ¿Hay algún error en mis registros de errores?
- ¿Cuál es el tamaño promedio de las solicitudes por semana?
- ¿Cuántas solicitudes hubo, agrupadas por código de respuesta, la semana pasada?

En respuesta, AI Assistant genera resúmenes en lenguaje natural de los resultados de las consultas en función de los diez primeros registros, como los siguientes:



La combinación de la generación de consultas en lenguaje natural y los resúmenes de consultas puede agregar un nivel adicional de análisis al solucionar una alerta, o brindar una manera más accesible de comprender sus datos sin tener que escribir consultas complejas.

Para ver los resúmenes de los resultados de las consultas generadas por AI Assistant en la página Discover

1. Comprueba que has configurado AI Assistant for OpenSearch Service.
2. En el menú principal del OpenSearch panel de mandos, selecciona Discover.
3. En la lista desplegable del lenguaje de consultas, elija PPL.
4. En el cuadro de texto AI Assistant, introduzca un mensaje y, a continuación, haga clic en el botón situado junto al cuadro de texto para ejecutar la consulta. AI Assistant puede tardar hasta 10 segundos en devolver un resumen. Después de su consulta inicial, usted debe elegir Generar resumen para obtener resúmenes posteriores.

> **Note**
>
> Puedes desactivar la generación de resúmenes en la lista desplegable del Asistente de IA en Discover.

Vea los detectores de anomalías recomendados

La detección de anomalías en Amazon OpenSearch Service detecta automáticamente las anomalías en los OpenSearch datos casi en tiempo real mediante el algoritmo Random Cut Forest (RCF). RCF es un algoritmo de machine learning no supervisado que modela un esquema del flujo de datos entrante. El algoritmo calcula un `anomaly grade` y valor `confidence score` para cada punto de datos entrante. La detección de anomalías utiliza estos valores para diferenciar una anomalía de las variaciones habituales de los datos.

Para simplificar el proceso de creación de detectores de anomalías, AI Assistant puede generar sugerencias de detectores en función de la fuente de datos seleccionada en la página Descubrir. AI Assistant admite los detectores de anomalías sugeridos para cualquier idioma.

Para ver los detectores de anomalías recomendados por AI Assistant

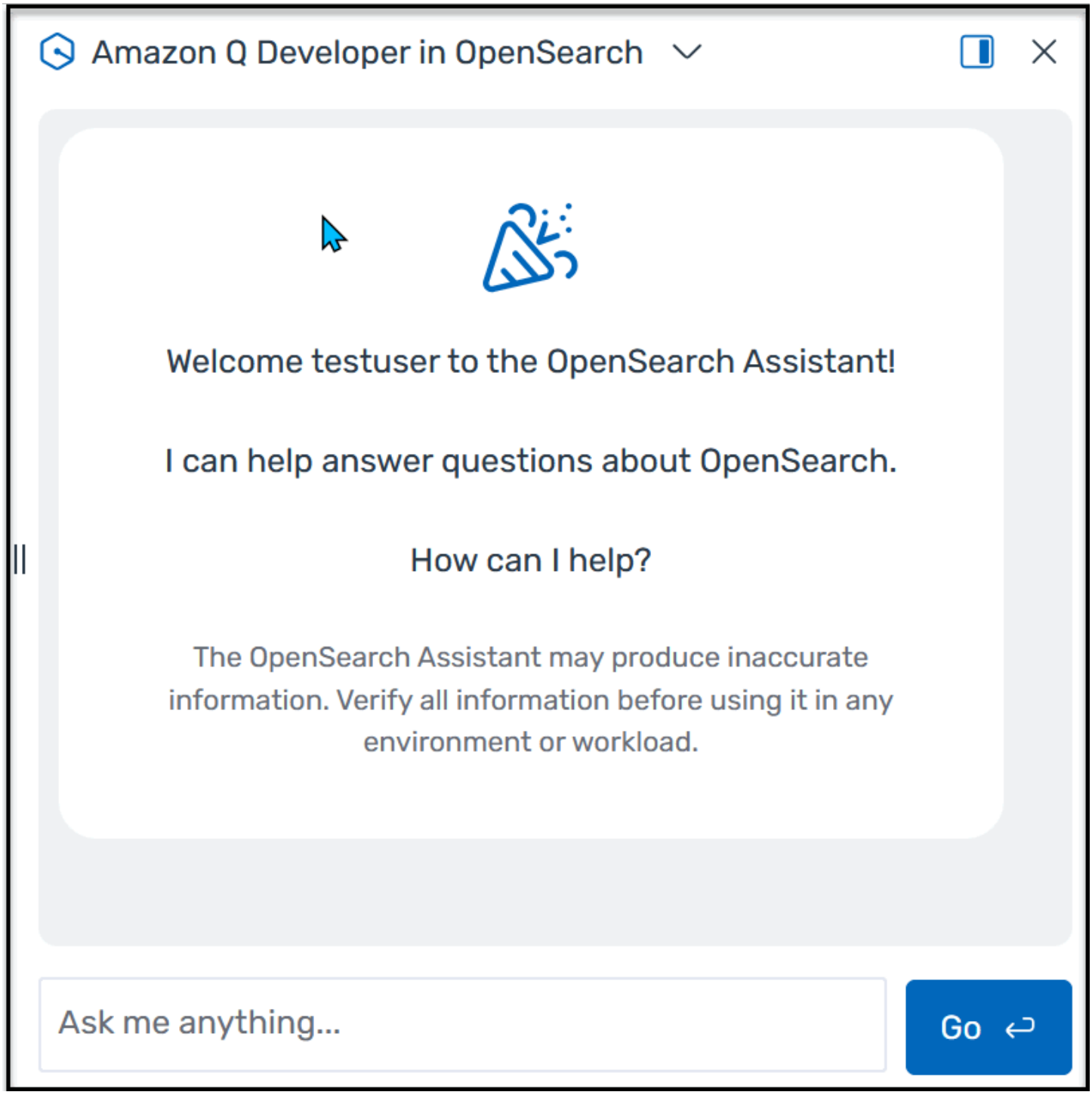
1. Comprueba que has configurado AI Assistant for OpenSearch Service.
2. En el menú principal de la OpenSearch interfaz de usuario, selecciona la página Descubrir y, a continuación, elige una fuente de datos.
3. En el menú del Asistente de IA, selecciona Sugerir detector de anomalías, como se muestra en la siguiente captura de pantalla.



   AI Assistant puede tardar unos segundos en generar las funciones del detector.
4. Seleccione Crear detector.

Accede al chat de AI Assistant si tienes preguntas OpenSearch relacionadas

Si tienes preguntas sobre Amazon OpenSearch Service, incluidas cuestiones conceptuales o de procedimiento relacionadas con las características o la funcionalidad del OpenSearch Servicio, puedes solicitar información a AI Assistant. El chatbot admite un chat conversacional que conserva el contexto de su conversación. También guarda un historial de sus conversaciones para consultarlo más adelante. Para acceder al chatbot, haz clic en el icono de AI Assistant situado en la esquina superior derecha de cualquier página de OpenSearch servicio. Escriba una pregunta en el cuadro de texto y haga clic en Ir. El Asistente de IA puede tardar hasta 30 segundos en generar una respuesta.



Este es un ejemplo de respuesta de AI Assistant.

Can you suggest some node management best practices?

Based on your context of being in the OpenSearch Dashboard in the us-west-2 region, here are some node management best practices for Amazon OpenSearch Service: [1]

1. Implement Dedicated Manager Nodes:
    - Use dedicated manager nodes to improve cluster stability. [2]
    - Configure three dedicated manager nodes for optimal domain stability across three Availability Zones.
    - This offloads cluster management tasks from data nodes, enhancing overall performance.
2. Multi-AZ Deployment:
    - Deploy your domain across multiple Availability Zones for increased availability and resilience.
    - Consider using the Multi-AZ with Standby option for 99.99% availability and consistent performance.
3. Proper Sizing:

Aprendizaje automático para Amazon OpenSearch Service

ML Commons es un OpenSearch complemento que proporciona un conjunto de algoritmos comunes de aprendizaje automático (ML) mediante llamadas a la API de transporte y REST. Esas llamadas eligen los nodos y los recursos correctos para cada solicitud de ML y supervisan las tareas de ML para garantizar el tiempo de actividad. Esto le permite aprovechar los algoritmos de ML de código abierto existentes y reducir el esfuerzo necesario para desarrollar nuevas características de ML. Para obtener más información sobre el complemento, consulte Aprendizaje automático en la OpenSearch documentación. En este capítulo se explica cómo utilizar el complemento con Amazon OpenSearch Service.

Temas

- Conectores Amazon OpenSearch Service ML para Servicios de AWS
- Conectores Amazon OpenSearch Service ML para plataformas de terceros
- Se utiliza CloudFormation para configurar la inferencia remota para la búsqueda semántica
- Configuración de ML Commons no compatible
- OpenSearch Plantillas de marcos de flujo de servicios

Conectores Amazon OpenSearch Service ML para Servicios de AWS

Cuando utilizas conectores de aprendizaje automático (ML) de Amazon OpenSearch Service con otro Servicio de AWS, debes configurar un rol de IAM para conectar el OpenSearch Servicio a ese servicio de forma segura. Servicios de AWS que puedes configurar un conector para incluir Amazon SageMaker AI y Amazon Bedrock. En este tutorial, explicamos cómo crear un conector desde OpenSearch Service hasta SageMaker Runtime. Para obtener más información sobre los conectores, consulte Conectores compatibles.

Temas

- Requisitos previos
- Crea un conector de OpenSearch servicio

Requisitos previos

Para crear un conector, debe tener un punto de enlace de Amazon SageMaker AI Domain y un rol de IAM que otorgue acceso al OpenSearch servicio.

Configurar un dominio de Amazon SageMaker AI

Consulte Implementación de un modelo en Amazon SageMaker AI en la Guía para desarrolladores de Amazon SageMaker AI para implementar su modelo de aprendizaje automático. Anote la URL del punto de conexión de su modelo, que necesitará para crear un conector de IA.

Creación de un rol de IAM

Configure un rol de IAM para delegar los permisos SageMaker de tiempo de ejecución al OpenSearch servicio. Para crear un rol nuevo, consulte Creación de roles de IAM (consola) en la Guía del usuario de IAM. Si lo desea, puede utilizar un rol existente siempre que tenga el mismo conjunto de privilegios. Si crea un rol nuevo en lugar de usar un rol AWS administrado, `opensearch-sagemaker-role` sustitúyalo en este tutorial por el nombre de su propio rol.

1. Adjunta la siguiente política de IAM gestionada a tu nueva función para permitir que el OpenSearch Servicio acceda a tu terminal de SageMaker IA. Para adjuntar una política a un rol, consulte Adición de permisos de identidad de IAM.

   JSON

   ```
   {
       "Version":"2012-10-17",
       "Statement": [
           {
               "Action": [
                   "sagemaker:InvokeEndpointAsync",
                   "sagemaker:InvokeEndpoint"
               ],
               "Effect": "Allow",
               "Resource": "*"
           }
       ]
   }
   ```

2. Siga las instrucciones de Modificación de una política de confianza de rol para editar la relación de confianza del rol. En la siguiente política, *`service-principal`* sustitúyala por uno de los siguientes principios de servicio para OpenSearch Service o OpenSearch Serverless:

Para servicio OpenSearch

```
opensearchservice.amazonaws.com
```

Para sistemas OpenSearch sin servidor

```
ml.opensearchservice.amazonaws.com
```

JSON

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

Le recomendamos que utilice las claves de condición `aws:SourceAccount` y `aws:SourceArn` para limitar el acceso a un dominio específico. `SourceAccount`Es el Cuenta de AWS ID que pertenece al propietario del dominio y `SourceArn` es el ARN del dominio. Por ejemplo, puede agregar el siguiente bloque de condición a la política de confianza:

```
"Condition": {
    "StringEquals": {
        "aws:SourceAccount": "account-id"
    },
    "ArnLike": {
        "aws:SourceArn": "arn:aws:es:region:account-id:domain/domain-name"
    }
}
```

Configuración de permisos

Para crear el conector, necesita permiso para transferir la función de IAM a OpenSearch Service. También necesita tener acceso a la acción `es:ESHttpPost`. Para conceder estos dos permisos, asocie la siguiente política al rol de IAM cuyas credenciales se utilicen para firmar la solicitud:

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/*"
        }
    ]
}
```

Si su usuario o rol no tiene permisos de `iam:PassRole` para transferir su rol, puede que se produzca un error de autorización cuando intente registrar un repositorio en el siguiente paso.

Asigne la función de aprendizaje automático en los OpenSearch paneles de control (si utiliza un control de acceso detallado)

El control de acceso detallado presenta un paso adicional al configurar un conector. Incluso si utiliza autenticación HTTP básica para todos los demás fines, debe asignar el rol `ml_full_access` al rol de IAM que tenga permisos `iam:PassRole` para transferir `opensearch-sagemaker-role`.

1. Navegue hasta el complemento OpenSearch Dashboards de su dominio de servicio. OpenSearch Puedes encontrar el punto de conexión de Dashboards en el panel de control de tu dominio, en la consola de OpenSearch servicio.
2. En el menú principal, seleccione Seguridad, Roles y seleccione el rol ml_full_access.

3. Seleccione Usuarios asignados, Administrar mapeo.
4. En Roles de backend, agregue el ARN del rol que tenga permisos para transferir `opensearch-sagemaker-role`.

```
arn:aws:iam::account-id:role/role-name
```

5. Seleccione Asignar y confirme que el usuario o el rol aparecen en Usuarios asignados.

Crea un conector de OpenSearch servicio

Para crear un conector, envíe una `POST` solicitud al punto final del dominio de OpenSearch servicio. Puede usar curl, el cliente Python de muestra, Postman u otro método para enviar una solicitud firmada. Tenga en cuenta que no puede usar una solicitud `POST` en la consola de Kibana. La solicitud tiene el siguiente 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}\" } }"
      }
   ]
```

```
}
```

Si el dominio reside en una nube privada virtual (VPC), la computadora debe estar conectada a la VPC para que la solicitud cree correctamente el conector de IA. El acceso a una VPC depende de la configuración de red, pero generalmente implica conectarse a una VPN o una red corporativa. Para comprobar si puedes acceder a tu dominio de OpenSearch servicio, navega `https://your-vpc-domain.region.es.amazonaws.com` en un navegador web y comprueba que recibes la respuesta JSON predeterminada.

Cliente Python de muestra

El cliente de Python es más simple de automatizar que una solicitud HTTP y tiene una mejor reutilización. Para crear el conector de IA con el cliente Python, guarde el siguiente código de ejemplo en un archivo Python. El cliente necesita los paquetes AWS SDK para Python (Boto3), `requests` y `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)
```

Conectores Amazon OpenSearch Service ML para plataformas de terceros

En este tutorial, explicamos cómo crear un conector de OpenSearch Service a Cohere. Para obtener más información sobre los conectores, consulte Conectores compatibles.

Cuando utilizas un conector de aprendizaje automático (ML) de Amazon OpenSearch Service con un modelo remoto externo, debes almacenar tus credenciales de autorización específicas en él AWS Secrets Manager. Puede ser una clave de API o una combinación de nombre de usuario y contraseña. Esto significa que también debe crear un rol de IAM que permita al OpenSearch Servicio acceder a Secrets Manager para leer.

Temas

- Requisitos previos
- Crea un conector de OpenSearch servicio

Requisitos previos

Para crear un conector para Cohere o cualquier proveedor externo con el OpenSearch Servicio, debe tener un rol de IAM que le dé acceso al OpenSearch Servicio AWS Secrets Manager, donde almacene sus credenciales. También debe almacenar sus credenciales en Secrets Manager.

Creación de un rol de IAM

Configure un rol de IAM para delegar los permisos de Secrets Manager a OpenSearch Service. También puede usar el rol de `SecretManagerReadWrite` existente. Para crear un rol nuevo, consulte Creación de roles de IAM (consola) en la Guía del usuario de IAM. Si crea un nuevo rol en lugar de usar un rol AWS administrado, `opensearch-secretmanager-role` sustitúyalo en este tutorial por el nombre de su propio rol.

1. Adjunta la siguiente política de IAM gestionada a tu nueva función para permitir que OpenSearch Service acceda a tus valores de Secrets Manager. Para adjuntar una política a un rol, consulte Adición de permisos de identidad de IAM.

   JSON

   ```
   {
       "Version":"2012-10-17",
       "Statement": [
           {
               "Action": [
                   "secretsmanager:GetSecretValue"
               ],
               "Effect": "Allow",
               "Resource": "*"
           }
       ]
   }
   ```

2. Siga las instrucciones de Modificación de una política de confianza de rol para editar la relación de confianza del rol. En la siguiente política, *`service-principal`* sustitúyala por uno de los siguientes principios de servicio para OpenSearch Service o OpenSearch Serverless:

   Para servicio OpenSearch

   `opensearchservice.amazonaws.com`

Para sistemas OpenSearch sin servidor

`ml.opensearchservice.amazonaws.com`

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
                "sts:AssumeRole"
            ],
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "service-principle"
                ]
            }
        }
    ]
}
```

Le recomendamos que utilice las claves de condición `aws:SourceAccount` y `aws:SourceArn` para limitar el acceso a un dominio específico. `SourceAccount`Es el Cuenta de AWS ID que pertenece al propietario del dominio y `SourceArn` es el ARN del dominio. Por ejemplo, puede agregar el siguiente bloque de condición a la política de confianza:

```
"Condition": {
    "StringEquals": {
        "aws:SourceAccount": "account-id"
    },
    "ArnLike": {
        "aws:SourceArn": "arn:aws:es:region:account-id:domain/domain-name"
    }
}
```

Configuración de permisos

Para crear el conector, necesita permiso para transferir la función de IAM a OpenSearch Service. También necesita tener acceso a la acción `es:ESHttpPost`. Para conceder estos dos permisos, asocie la siguiente política al rol de IAM cuyas credenciales se utilicen para firmar la solicitud:

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/*"
    }
  ]
}
```

Si su usuario o rol no tiene permisos de `iam:PassRole` para transferir su rol, puede que se produzca un error de autorización cuando intente registrar un repositorio en el siguiente paso.

Configurar AWS Secrets Manager

Para almacenar sus credenciales de autorización en Secrets Manager, consulte Create an AWS Secrets Manager secret en la Guía del usuario de AWS Secrets Manager .

Después de que Secrets Manager acepta su par clave-valor como secreto, recibirá un ARN con el formato: `arn:aws:secretsmanager:us-west-2:123456789012:secret:MySecret-a1b2c3`. Mantenga un registro de este ARN, a medida que lo usa y su clave cuando cree un conector en el siguiente paso.

Asigne la función del aprendizaje automático en los OpenSearch paneles de control (si utiliza un control de acceso detallado)

El control de acceso detallado presenta un paso adicional al configurar un conector. Incluso si utiliza autenticación HTTP básica para todos los demás fines, debe asignar el rol `ml_full_access` al rol de IAM que tenga permisos `iam:PassRole` para transferir `opensearch-sagemaker-role`.

1. Navegue hasta el complemento OpenSearch Dashboards de su dominio de servicio. OpenSearch Puedes encontrar el punto de conexión de Dashboards en el panel de control de tu dominio, en la consola de OpenSearch servicio.
2. En el menú principal, seleccione Seguridad, Roles y seleccione el rol ml_full_access.
3. Seleccione Usuarios asignados, Administrar mapeo.
4. En Roles de backend, agregue el ARN del rol que tenga permisos para transferir `opensearch-sagemaker-role`.

   ```
   arn:aws:iam::account-id:role/role-name
   ```

5. Seleccione Asignar y confirme que el usuario o el rol aparecen en Usuarios asignados.

Crea un conector de OpenSearch servicio

Para crear un conector, envíe una `POST` solicitud al punto final del dominio de OpenSearch servicio. Puede usar curl, el cliente Python de muestra, Postman u otro método para enviar una solicitud firmada. Tenga en cuenta que no puede usar una solicitud `POST` en la consola de Kibana. La solicitud tiene el siguiente 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\" }"
        }
    ]
}
```

El cuerpo de esta solicitud es diferente de una solicitud de conector de código abierto en dos aspectos. Dentro del `credential` campo, se pasa el ARN del rol de IAM que permite al OpenSearch Servicio leer Secrets Manager, junto con el ARN del secreto qué. En el campo `headers`, se hace referencia al secreto mediante la clave secreta y al hecho de que proviene de un ARN.

Si el dominio reside en una nube privada virtual (VPC), la computadora debe estar conectada a la VPC para que la solicitud cree correctamente el conector de IA. El acceso a una VPC depende de la configuración de red, pero generalmente implica conectarse a una VPN o una red corporativa. Para comprobar que puedes acceder a tu dominio de OpenSearch servicio, navega `https://`*`your-vpc-domain`*`.`*`region`*`.es.amazonaws.com` en un navegador web y comprueba que recibes la respuesta JSON predeterminada.

Cliente Python de muestra

El cliente de Python es más simple de automatizar que una solicitud HTTP y tiene una mejor reutilización. Para crear el conector de IA con el cliente Python, guarde el siguiente código de ejemplo en un archivo Python. El cliente necesita los paquetes AWS SDK para Python (Boto3), `requests` y `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)
```

Se utiliza CloudFormation para configurar la inferencia remota para la búsqueda semántica

A partir de OpenSearch la versión 2.9, puede utilizar la inferencia remota con la búsqueda semántica para alojar sus propios modelos de aprendizaje automático (ML). La inferencia remota utiliza el complemento ML Commons.

Con la inferencia remota, puede alojar sus inferencias de modelos de forma remota en servicios de aprendizaje automático, como Amazon SageMaker AI y Amazon Bedrock, y conectarlos a Amazon OpenSearch Service con conectores de aprendizaje automático.

Para facilitar la configuración de la inferencia remota, Amazon OpenSearch Service proporciona una AWS CloudFormationplantilla en la consola. CloudFormation es un Servicio de AWS lugar en el que puede aprovisionar AWS y administrar recursos de terceros tratando la infraestructura como código.

La OpenSearch CloudFormation plantilla automatiza el proceso de aprovisionamiento de modelos, de modo que puede crear fácilmente un modelo en su dominio de OpenSearch servicio y, a continuación, utilizar el ID del modelo para ingerir datos y ejecutar consultas de búsqueda neuronal.

Si utiliza codificadores neuronales dispersos con la versión 2.12 y posteriores del OpenSearch Servicio, le recomendamos que utilice el modelo de tokenizador de forma local en lugar de implementarlo de forma remota. Para obtener más información, consulte los modelos de codificación dispersa en la documentación. OpenSearch

Temas

- Plantillas disponibles CloudFormation
- Requisitos previos
- Plantillas de Amazon Bedrock
- Configuración de la búsqueda de agentes con Bedrock Claude
- Plantillas de integración de servidores MCP
- SageMaker Plantillas de Amazon
- Inferencia remota para plantillas de resaltado semántico

Plantillas disponibles CloudFormation

Las siguientes plantillas AWS CloudFormation de aprendizaje automático (ML) están disponibles para su uso:

Plantillas de Amazon Bedrock

Integración de Amazon Titan Text Embeddings

> Se conecta a los modelos de ML alojados de Amazon Bedrock, elimina la necesidad de implementar modelos independientes y utiliza puntos de conexión de Amazon Bedrock

predeterminados. Para obtener más información, consulte Amazon Titan Text Embeddings en la Guía del usuario de Amazon Bedrock.

Integración Cohere Embed

Proporciona acceso a los modelos Cohere Embed y está optimizado para flujos de trabajo específicos de procesamiento de texto. Para obtener más información, consulte Embed en el sitio web de documentos de Cohere.

Amazon Titan Multimodal Embeddings

Admite la incrustación de texto e imágenes y permite funciones de búsqueda multimodal. Para obtener más información, consulte Amazon Titan Multimodal Embeddings en la Guía del usuario de Amazon Bedrock.

Plantillas de integración de servidores MCP

Integración de servidores MCP

Implementa un Amazon Bedrock AgentCore Runtime, proporciona un punto de enlace de agente, gestiona la autenticación entrante y saliente y admite OAuth la autenticación empresarial.

SageMaker Plantillas de Amazon

Integración con modelos de incrustación de texto a través de Amazon SageMaker

Implementa modelos de incrustación de texto en Amazon SageMaker Runtime, crea funciones de IAM para el acceso a los artefactos del modelo y establece conectores de aprendizaje automático para la búsqueda semántica.

Integración con codificadores dispersos mediante SageMaker

Configura modelos de codificación dispersos para la búsqueda neuronal, crea AWS Lambda funciones para la gestión de conectores y devuelve el modelo IDs para su uso inmediato.

Requisitos previos

Para usar una CloudFormation plantilla con OpenSearch Service, complete los siguientes requisitos previos.

Configure un dominio de OpenSearch servicio

Antes de poder usar una CloudFormation plantilla, debes configurar un dominio de Amazon OpenSearch Service con la versión 2.9 o posterior y un control de acceso detallado activado. Crea un rol OpenSearch de backend del servicio para dar permiso al plugin ML Commons para que cree tu conector por ti.

La CloudFormation plantilla crea una función de IAM de Lambda para usted con el nombre predeterminado`LambdaInvokeOpenSearchMLCommonsRole`, que puede anular si desea elegir un nombre diferente. Una vez que la plantilla cree este rol de IAM, debe conceder permiso a la función Lambda para llamar a OpenSearch su dominio de servicio. Para ello, asigne la función nombrada `ml_full_access` a su función de backend del OpenSearch servicio siguiendo los siguientes pasos:

1. Navega hasta el complemento OpenSearch Dashboards de tu dominio de OpenSearch servicio. Puedes encontrar el punto de conexión de Dashboards en el panel de control de tu dominio, en la consola de OpenSearch servicio.
2. En el menú principal, seleccione Seguridad, Roles y seleccione el rol ml_full_access.
3. Seleccione Usuarios asignados, Administrar mapeo.
4. En Roles de backend, agregue el ARN del rol de Lambda que necesita permiso para llamar a su dominio.

   ```
   arn:aws:iam::account-id:role/role-name
   ```

5. Seleccione Asignar y confirme que el usuario o el rol aparecen en Usuarios asignados.

Una vez que haya asignado la función, vaya a la configuración de seguridad de su dominio y añada la función Lambda IAM a OpenSearch su política de acceso al servicio.

Habilite los permisos en su Cuenta de AWS

Cuenta de AWS Debe tener permiso para acceder CloudFormation a Lambda, junto con lo que Servicio de AWS elija para su plantilla, ya sea SageMaker Runtime o Amazon Bedrock.

Si utiliza Amazon Bedrock, también debe registrar su modelo. Consulte Acceso al modelo en la Guía del usuario de Amazon Bedrock para registrar su modelo.

Si utiliza su propio bucket de Amazon S3 para proporcionar artefactos modelo, debe añadir la función de CloudFormation IAM a su política de acceso a S3. Para obtener más información, consulta Adición y eliminación de permisos de identidad de IAM en la Guía del usuario de IAM.

Plantillas de Amazon Bedrock

Las CloudFormation plantillas de Amazon Bedrock proporcionan los AWS recursos necesarios para crear conectores entre OpenSearch Service y Amazon Bedrock.

En primer lugar, la plantilla crea una función de IAM que permite a la futura función de Lambda acceder a OpenSearch su dominio de servicio. A continuación, la plantilla crea la función de Lambda, que hace que el dominio cree un conector mediante el complemento ML Commons. Una vez que OpenSearch Service crea el conector, finaliza la configuración de la inferencia remota y puede ejecutar búsquedas semánticas mediante las operaciones de la API de Amazon Bedrock.

> **Note**
>
> Dado que Amazon Bedrock aloja sus propios modelos de aprendizaje automático, no es necesario implementar un modelo en SageMaker Runtime. En su lugar, la plantilla utiliza un punto de conexión predeterminado para Amazon Bedrock y omite los pasos de aprovisionamiento del punto de conexión.

Para usar la plantilla Amazon Bedrock CloudFormation

1. Abre la consola OpenSearch de Amazon Service.
2. En el panel de navegación de la izquierda, seleccione Integraciones.
3. En Integrar con el modelo de Amazon Titan Text Embeddings a través de Amazon Bedrock, elija Configurar dominio, Configurar dominio público.
4. Siga la petición para configurar el modelo.

> **Note**
>
> OpenSearch El servicio también proporciona una plantilla independiente para configurar un dominio de Amazon VPC. Si usa esta plantilla, debe proporcionar el ID de VPC de Amazon para la función de Lambda.

Además, OpenSearch Service proporciona las siguientes plantillas de Amazon Bedrock para conectarse al modelo Cohere y al modelo Amazon Titan Multimodal Embeddings:

- `Integration with Cohere Embed through Amazon Bedrock`

- `Integrate with Amazon Bedrock Titan Multi-modal`

Configuración de la búsqueda de agentes con Bedrock Claude

La búsqueda de agentes autónomos utiliza agentes autónomos para ejecutar búsquedas complejas en su nombre, ya que comprende las intenciones de los usuarios, organiza las herramientas adecuadas, genera consultas optimizadas y proporciona resúmenes transparentes de sus decisiones a través de una interfaz de lenguaje natural. Estos agentes se basan en modelos de razonamiento, como Bedrock Claude.

Siga los pasos que se indican a continuación para abrir y ejecutar una CloudFormation plantilla que configure automáticamente los modelos de Bedrock Claude para la búsqueda de agentes y aprenda a configurar y crear sus agentes en el complemento AI Search Flows de Dashboards. OpenSearch

Habilitar el acceso a Bedrock Claude

1. Requisito previo: si su dominio utiliza un control de acceso detallado, `arn:aws:iam::`*`your-account-id`*`:role/`*`LambdaInvokeOpenSearchMLCommonsRole`* asigne el rol de backend al rol antes de ejecutar la plantilla`ml_full_access`. Este rol de IAM se creará automáticamente CloudFormation si aún no existe. Para obtener más información sobre cómo configurar el mapeo, consulteAsigne la función del aprendizaje automático en los OpenSearch paneles de control (si utiliza un control de acceso detallado).
2. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
3. En el panel de navegación izquierdo, seleccione Integraciones.
4. En Integración con Bedrock Claude para Agentic Search, selecciona Configurar dominio. Asegúrese de que su dominio tenga la versión 3.3 o superior.
5. En la CloudFormation plantilla, introduce el punto de conexión del dominio de OpenSearch servicio y selecciona un modelo. Los campos restantes son opcionales o se rellenan previamente. Haga clic en Crear pila y espere a que se complete el aprovisionamiento.
6. En la consola de Amazon OpenSearch Service, selecciona Dominios y selecciona tu dominio. Haga clic en la URL de OpenSearch los paneles para acceder a los OpenSearch paneles.

Creación de agentes y ejecución de Agentic Search

1. Desde los OpenSearch paneles de control, abra el menú de la izquierda. Selecciona OpenSearch Plugins > AI Search Flows para acceder al plugin.

2. En la página Flujos de trabajo, seleccione la pestaña Nuevo flujo de trabajo y, en la tarjeta de búsqueda de agentes, haga clic en Crear.
3. Proporcione un nombre exclusivo para la configuración de búsqueda y haga clic en Crear.
4. En Configurar agente, haga clic en Crear nuevo agente. Seleccione el modelo Bedrock Claude recién creado y, a continuación, haga clic en Crear agente. Si el botón está desactivado, seleccione Configuración avanzada > Interfaz LLM y asegúrese de que haya seleccionado una interfaz válida. Todos los modelos CloudFormation serán modelos Bedrock Claude, por lo que puede seleccionar Bedrock Claude, si aún no lo está, y luego hacer clic en Crear agente.
5. En Flujo de pruebas, intenta realizar búsquedas con agentes. Proporcione una consulta de búsqueda en lenguaje natural y haga clic en Buscar.

Para obtener la documentación completa del complemento AI Search Flows, consulte Configuración de la búsqueda de agentes en la OpenSearch documentación.

Para obtener más información sobre cómo funciona Agentic Search, consulte Agentic Search en la documentación. OpenSearch

Plantillas de integración de servidores MCP

Con las plantillas de servidor del Model Context Protocol (MCP), puede implementar un servidor MCP OpenSearch alojado en Amazon Bedrock AgentCore, lo que reduce la complejidad de la integración entre los agentes y las herramientas de IA. OpenSearch Para obtener más información, consulte ¿Qué es Amazon Bedrock? AgentCore .

Características de la plantilla

Esta plantilla incluye las siguientes características clave para implementar y administrar su servidor MCP.

Implementación de un servidor MCP administrado

Se implementa opensearch-mcp-server-pymediante Amazon Bedrock AgentCore Runtime y proporciona un punto de enlace de agente que envía las solicitudes por proxy al servidor MCP subyacente. Para obtener más información, consulte en. opensearch-mcp-server-pyGitHub

Autenticación y seguridad

Gestiona tanto la autenticación entrante (de los usuarios al servidor MCP) como la autenticación saliente (del servidor MCP al OpenSearch), y es compatible con OAuth la autenticación empresarial.

> **Note**
>
> La plantilla del servidor MCP solo está disponible en las siguientes Regiones de AWS:
>
> - Este de EE. UU. (Norte de Virginia)
> - Oeste de EE. UU. (Oregón)
> - Europa (Fráncfort)
> - Asia-Pacífico (Sídney)

Para utilizar la plantilla de servidor MCP

Siga estos pasos para implementar la plantilla de servidor MCP y conectarla a su dominio. OpenSearch

1. Abre la consola OpenSearch de Amazon Service.
2. En el panel de navegación de la izquierda, seleccione Integraciones.
3. Localice la plantilla de integración del servidor MCP.
4. Elija Configurar dominio. A continuación, introduce el punto de enlace OpenSearch de tu dominio.

La plantilla crea un AgentCore Runtime y los siguientes componentes, si no se especifican los parámetros opcionales correspondientes:

- Un repositorio de Amazon ECR.
- Un grupo de usuarios de Amazon Cognito como autorizador OAuth
- Una función de ejecución utilizada por Runtime AgentCore

Después de completar este procedimiento, debe seguir estos pasos posteriores a la creación:

1. Para Amazon OpenSearch Service: asigne el ARN de su rol de ejecución a un rol de OpenSearch backend para controlar el acceso a su dominio.

   Para Amazon OpenSearch Serverless: cree una política de acceso a los datos que permita a su rol de ejecución acceder a su colección.
2. Obtenga un token de OAuth acceso de su autorizador. A continuación, utilice este token para acceder al servidor MCP en la URL que aparece en la salida de la CloudFormation pila.

Para obtener más información, consulte Acciones políticas para Serverless OpenSearch.

Integración con agentes de IA

Tras la implementación, puede integrar el servidor MCP con cualquier agente compatible con MCP. Para obtener más información, consulte Invocar a su servidor MCP desplegado en la Guía para desarrolladores de Amazon Bedrock.

Integración de desarrolladores

Puede agregar el punto de conexión del servidor MCP a la configuración de su agente. También puede usarlo con la CLI para Amazon Q Developer, agentes personalizados u otros agentes compatibles con MCP.

Implementación empresarial

Los agentes alojados en Central pueden conectarse a varios servicios con OpenSearch un solo componente. Este agente es compatible con los sistemas de autenticación empresariales OAuth y se escala para admitir varios usuarios y casos de uso.

Ejemplo de uso del marco 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")
```

Para obtener más información, consulte Alojamiento de un servidor OpenSearch MCP con Amazon Bedrock AgentCore en el OpenSearch sitio web.

SageMaker Plantillas de Amazon

SageMaker CloudFormation Las plantillas de Amazon definen varios AWS recursos para configurar el complemento neuronal y la búsqueda semántica por usted.

Comience por utilizar la SageMaker plantilla Integración con modelos de incrustación de texto a través de Amazon para implementar un modelo de incrustación de texto en SageMaker Runtime como servidor. Si no proporciona un punto de enlace de modelo, CloudFormation crea un rol de IAM que permite a SageMaker Runtime descargar artefactos de modelos de Amazon S3 e

implementarlos en el servidor. Si proporciona un punto final, CloudFormation crea un rol de IAM que permite a la función Lambda acceder al dominio OpenSearch del servicio o, si el rol ya existe, actualiza y reutiliza el rol. El punto de conexión sirve al modelo remoto que se utiliza para el conector de ML con el complemento ML Commons.

A continuación, utilice la SageMaker plantilla Integration with Sparse Encoders through Amazon para crear una función Lambda que haga que su dominio configure conectores de inferencia remota. Una vez creado el conector en OpenSearch Service, la inferencia remota puede ejecutar una búsqueda semántica utilizando el modelo remoto en tiempo de ejecución. SageMaker La plantilla le devuelve el ID del modelo de su dominio para que pueda empezar la búsqueda.

Para usar las CloudFormation plantillas de Amazon SageMaker AI

1. Abre la consola OpenSearch de Amazon Service.
2. En el panel de navegación de la izquierda, seleccione Integraciones.
3. En cada una de las plantillas de Amazon SageMaker AI, selecciona Configurar dominio y Configurar dominio público.
4. Siga las instrucciones de la CloudFormation consola para aprovisionar la pila y configurar un modelo.

> **Note**
>
> OpenSearch El servicio también proporciona una plantilla independiente para configurar el dominio de VPC. Si usa esta plantilla, debe proporcionar el ID de VPC para la función de Lambda.

Inferencia remota para plantillas de resaltado semántico

El resaltado semántico es una característica de búsqueda avanzada que mejora la relevancia de los resultados al analizar el significado y el contexto de las consultas en lugar de basarse únicamente en las coincidencias exactas de las palabras clave. Esta función utiliza modelos de machine learning para evaluar la similitud semántica entre las consultas de búsqueda y el contenido del documento, identificando y resaltando las frases o pasajes más relevantes desde el punto de vista contextual de los documentos. A diferencia de los métodos de resaltado tradicionales, que se centran en las coincidencias exactas de los términos, el resaltado semántico aprovecha los modelos de IA para evaluar cada oración utilizando información contextual tanto de la consulta como del texto

circundante, lo que le permite mostrar información pertinente incluso cuando los términos de búsqueda exactos no estén presentes en los pasajes resaltados. Este enfoque es particularmente valioso para las implementaciones de búsqueda impulsadas por la IA, en las que los usuarios dan prioridad al significado semántico por encima de la coincidencia literal de palabras, lo que permite a los administradores de búsquedas ofrecer experiencias de búsqueda más inteligentes y contextualmente conscientes que destacan las extensiones de contenido significativas en lugar de solo las apariciones de palabras clave. Para obtener más información, consulte Uso del resaltado semántico.

Utilice el siguiente procedimiento para abrir y ejecutar una CloudFormation plantilla que configure automáticamente los SageMaker modelos de Amazon para el resaltado semántico.

Para usar la plantilla de resaltado semántico CloudFormation

1. Abre la consola OpenSearch de Amazon Service en https://console.aws.amazon.com/aos/casa.
2. En el panel de navegación izquierdo, seleccione Integraciones.
3. En Habilitar el resaltado semántico mediante SageMaker la integración de Amazon, selecciona Configurar dominio y Configurar dominio público.
4. Siga la petición para configurar el modelo.

> **Note**
>
> OpenSearch El servicio también proporciona una plantilla independiente para configurar el dominio de VPC. Si usa esta plantilla, debe proporcionar el ID de VPC para la función de Lambda.

Configuración de ML Commons no compatible

Amazon OpenSearch Service no admite el uso de las siguientes configuraciones de ML Commons:

- `plugins.ml_commons.allow_registering_model_via_url`
- `plugins.ml_commons.allow_registering_model_via_local_file`

> **Important**
>
> En los clústeres de producción, no desactive la configuración del clúster `plugins.ml_commons.only_run_on_ml_node` (no la configure en `false`). La opción de deshabilitar esta protección es para facilitar el desarrollo, pero los clústeres de producción deberían utilizar los conectores. Para obtener más información, consulte the section called "Conectores para Servicios de AWS".

Para obtener más información sobre la configuración de ML Commons, consulte ML Commons cluster settings.

OpenSearch Plantillas de marcos de flujo de servicios

Las plantillas OpenSearch de Amazon Service Flow Framework te permiten automatizar tareas complejas de configuración y preprocesamiento de OpenSearch servicios al proporcionar plantillas para casos de uso comunes. Por ejemplo, puede utilizar plantillas de Flow Framework para automatizar las tareas de configuración de machine learning. Las plantillas OpenSearch de Amazon Service Flow Framework proporcionan una descripción compacta del proceso de configuración en un documento JSON o YAML. Estas plantillas describen las configuraciones de flujo de trabajo automatizadas para la generación de consultas o chats conversacionales, los conectores de IA, las herramientas, los agentes y otros componentes que preparan a OpenSearch Service para su uso interno en modelos generativos.

Las plantillas OpenSearch de Amazon Service Flow Framework se pueden personalizar para adaptarlas a tus necesidades específicas. Para ver un ejemplo de una plantilla de Flow Framework personalizada, consulte flow-framework. Para ver las plantillas proporcionadas por el OpenSearch servicio, consulte plantillas de flujo de trabajo. Para obtener una documentación completa, que incluye los pasos detallados, una referencia de la API y una referencia de todos los ajustes disponibles, consulte la automatización de la configuración en la documentación de código abierto. OpenSearch

> **Note**
>
> Flow-framework no admite el filtrado de roles de back-end para el Servicio 2.17. OpenSearch

Creación de conectores ML en Service OpenSearch

Las plantillas OpenSearch de Amazon Service Flow Framework permiten configurar e instalar conectores ML mediante la API de creación de conectores que se ofrece en ml-commons. Puede utilizar los conectores ML para conectar el OpenSearch Servicio a otros AWS servicios o plataformas de terceros. Para obtener más información al respecto, consulte Creating connectors for third-party ML platforms. La API del marco de flujo de Amazon OpenSearch OpenSearch Service te permite automatizar las tareas de configuración y preprocesamiento del servicio y se puede utilizar para crear conectores ML.

Antes de poder crear un conector en OpenSearch Service, debe hacer lo siguiente:

- Crea un dominio de Amazon SageMaker AI.
- Crear un rol de IAM.
- Configure el permiso para transferir roles.
- Mapee las funciones de flow-framework y ml-commons en los paneles. OpenSearch

Para obtener más información sobre cómo configurar conectores ML para AWS servicios, consulta Conectores ML de Amazon OpenSearch Service para AWS servicios. Para obtener más información sobre el uso de conectores de OpenSearch Service ML con plataformas de terceros, consulte Conectores de Amazon OpenSearch Service ML para plataformas de terceros.

Creación de un conector mediante un servicio flow-framework

Para crear una plantilla de marco de flujo con un conector, tendrás que enviar una `POST` solicitud al punto de enlace de tu dominio de OpenSearch servicio. Puede usar cURL, un cliente Python de muestra, Postman u otro método para enviar una solicitud firmada. La solicitud `POST` tiene el siguiente 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"
            }
```

```
          }
        }
      ]
    }
  }
}
```

Si el dominio reside en una nube privada virtual (Amazon VPC), debe haberse conectado a la Amazon VPC para que la solicitud cree correctamente el conector de IA. El acceso a una Amazon VPC depende de la configuración de red, pero generalmente implica conectarse a una VPN o una red corporativa. Para comprobar que puedes acceder a tu dominio de OpenSearch servicio, navega `https://`*`your-vpc-domain`*`.`*`region`*`.es.amazonaws.com` en un navegador web y comprueba que recibes la respuesta JSON predeterminada. (*`placeholder text`*Sustitúyalos por sus propios valores.

Cliente Python de muestra

El cliente de Python es más simple de automatizar que una solicitud HTTP y tiene una mejor reutilización. Para crear el conector de IA con el cliente Python, guarde el siguiente código de ejemplo en un archivo Python. El cliente requiere los paquetes AWS SDK para Python (Boto3), Requests:HTTP for Humans y 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)
```

Plantillas de flujo de trabajo predefinidas

Amazon OpenSearch Service proporciona varias plantillas de flujo de trabajo para algunos casos de uso habituales del aprendizaje automático (ML). El uso de una plantilla simplifica las configuraciones complejas y proporciona muchos valores predeterminados para casos de uso como la búsqueda semántica o conversacional. Puede especificar una plantilla de flujo de trabajo al llamar a la API Create Workflow.

- Para usar una plantilla de flujo de trabajo proporcionada por el OpenSearch servicio, especifique el caso de uso de la plantilla como parámetro de `use_case` consulta.
- Para usar una plantilla de flujo de trabajo personalizada, proporcione la plantilla completa en el cuerpo de la solicitud. Para ver un ejemplo de plantilla personalizada, consulte un ejemplo de plantilla JSON o un ejemplo de plantilla YAML.

Casos de uso de plantillas

Esta tabla proporciona información general de las diferentes plantillas disponibles, una descripción de las plantillas y los parámetros necesarios.

Caso de uso de plantillas	Description (Descripción)	Parámetros requeridos
`bedrock_titan_embe`	Crea e implementa un modelo de incrustación de Amazon Bedrock (de	`create_connector.credential.roleArn`

Caso de uso de plantillas	Description (Descripción)	Parámetros requeridos
dding_model_deploy	forma predeterminada, titan-embed-text-v1)	
bedrock_titan_embedding_model_deploy	Crea e implementa un modelo de incrustación multimodal de Amazon Bedrock (de forma predeterminada, titan-embed-text-v1)	create_connector.credential.roleArn
cohere_embedding_model_deploy	Crea e implementa un modelo de incrustación de Cohere (de forma predeterminada, embed-english-v3.0).	create_connector.credential.roleArn , create_connector.credential.secretArn
cohere_chat_model_deploy	Crea e implementa un modelo de chat de Cohere (de forma predeterminada, Cohere Command).	create_connector.credential.roleArn , create_connector.credential.secretArn
open_ai_embedding_model_deploy	Crea e implementa un modelo de incrustación de OpenAI (de forma predeterminada text-embedding-ada, -002).	create_connector.credential.roleArn , create_connector.credential.secretArn
openai_chat_model_deploy	Crea e implementa un modelo de chat de OpenAI (por defecto, gpt-3.5-turbo).	create_connector.credential.roleArn , create_connector.credential.secretArn
semantic_search_with_cohere_embedding	Configura la búsqueda semántica e implementa un modelo de incrustación de Cohere. Debe proporcionar la clave de API para el modelo de Cohere.	create_connector.credential.roleArn , create_connector.credential.secretArn

Caso de uso de plantillas	Description (Descripción)	Parámetros requeridos
`semantic_search_with_cohere_embedding_query_enricher`	Configura la búsqueda semántica e implementa un modelo de incrustación de Cohere. Agrega un procesador de búsqueda query_enricher que establece un ID de modelo predeterminado para las consultas neuronales. Debe proporcionar la clave de API para el modelo de Cohere.	`create_connector.credential.roleArn` , `create_connector.credential.secretArn`
`multimodal_search_with_bedrock_titan`	Implementa un modelo multimodal de Amazon Bedrock y configura una canalización de ingesta con un procesador text_image_embedding y un índice k-NN para la búsqueda multimodal. Debe proporcionar sus credenciales de AWS .	`create_connector.credential.roleArn`

> **Note**
>
> Para todas las plantillas que requieren un ARN secreto, lo predeterminado es almacenar el secreto con el nombre de clave «clave» en AWS Secrets Manager.

Plantillas predeterminadas con modelos entrenados previamente

Amazon OpenSearch Service ofrece dos plantillas de flujo de trabajo predeterminadas adicionales que no están disponibles en el servicio de código abierto OpenSearch .

Caso de uso de plantillas	Description (Descripción)
`semantic_search_with_local_model`	Configura la búsqueda semántica e implementa un modelo entrenado previamente (`msmarco-distilbert-base-tas-b`). Añade un procesador `neural_query_enricher` de

Caso de uso de plantillas	Description (Descripción)
	búsqueda que establece un ID de modelo predeterminado para las consultas neuronales y crea un índice k-NN vinculado llamado ". my-nlp-index
`hybrid_search_with_local_model`	Configura la búsqueda híbrida e implementa un modelo entrenado previamente (`msmarco-distilbert-base-tas-b`). Añade un procesador `neural_query_enricher` de búsqueda que establece un ID de modelo predeterminado para las consultas neuronales y crea un índice k-NN vinculado denominado ". my-nlp-index

Configuración de permisos

Si crea un dominio nuevo con la versión 2.13 o una posterior, los permisos ya están en vigor. Si habilita el marco de flujo en un dominio de OpenSearch servicio preexistente con la versión 2.11 o anterior y, a continuación, lo actualiza a la versión 2.13 o posterior, debe definir la función. `flow_framework_manager` Los usuarios que no sean administradores deben estar asignados a este rol para poder administrar índices templados en los dominios mediante un control de acceso detallado. Para crear el rol `flow_framework_manager` de forma manual, siga estos pasos:

1. En los OpenSearch paneles, vaya a Seguridad y elija Permisos.
2. Seleccione Crear grupo de acciones y configure los siguientes grupos:

Nombre del grupo	Permisos
`flow_framework_full_access`	• `cluster:admin/opensearch/flow_framework/*` • `cluster_monitor`
`flow_framework_read_accesss`	• `cluster:admin/opensearch/flow_framework/workflow/get`

Nombre del grupo	Permisos
	• `cluster:admin/opensearch/flow_framework/workflow/search` • `cluster:admin/opensearch/flow_framework/workflow_state/get` • `cluster:admin/opensearch/flow_framework/workflow_state/search`

3. Seleccione Roles y, a continuación, Crear rol.
4. Asigne el nombre flow_framework_manager al rol.
5. Para Permisos de clúster, seleccione `flow_framework_full_access` y `flow_framework_read_access`.
6. Para Índice, escriba `*`.
7. Para Permisos de índice, seleccione `indices:admin/aliases/get`, `indices:admin/mappings/get`, y `indices_monitor`.
8. Elija Crear.
9. Después de crear el rol, debe asignarlo a cualquier rol de usuario o backend que administre los índices de Flow Framework.

Análisis de seguridad para Amazon OpenSearch Service

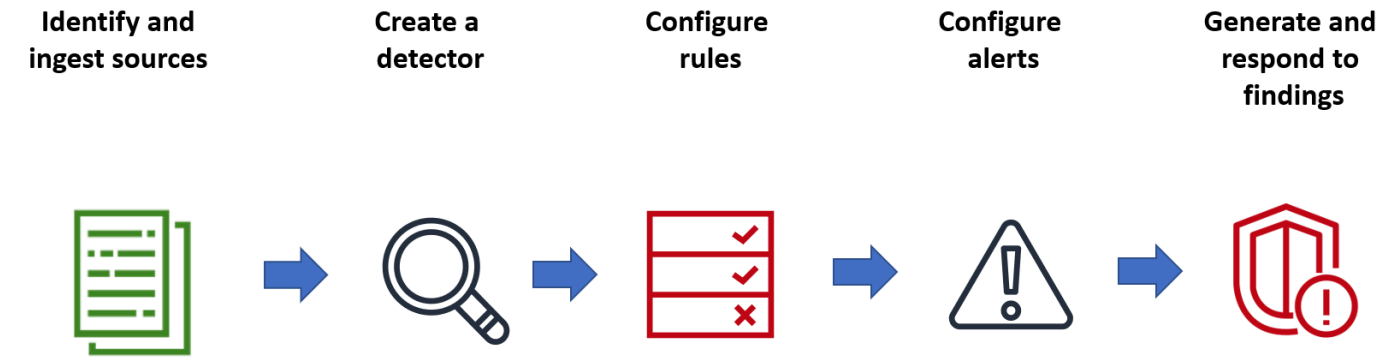
Security Analytics es una OpenSearch solución que proporciona visibilidad de la infraestructura de su organización, monitorea la actividad anómala, detecta posibles amenazas de seguridad en tiempo real y activa alertas a destinos preconfigurados. Puede monitorear la actividad maliciosa de sus registros de eventos de seguridad evaluando continuamente las reglas de seguridad y revisando los resultados de seguridad generados automáticamente. Además, Security Analytics puede generar alertas automatizadas y enviarlas a un canal de notificación específico, como Slack o el correo electrónico.

Puede usar el complemento Security Analytics para detectar amenazas comunes out-of-the-box y generar información de seguridad crítica a partir de los registros de eventos de seguridad existentes, como los registros de firewall, los registros de Windows y los registros de auditoría de autenticación. Para usar Security Analytics, su dominio debe ejecutar la OpenSearch versión 2.5 o posterior.

> **Note**
>
> En esta documentación se proporciona una breve descripción de Security Analytics for Amazon OpenSearch Service. Define los conceptos clave y proporciona los pasos para configurar los permisos. Para obtener una documentación completa, que incluye una guía de configuración, una referencia de la API y una referencia de todos los ajustes disponibles, consulte Security Analytics en la OpenSearch documentación.

Componentes y conceptos de Security Analytics

Una serie de herramientas y características constituyen la base del funcionamiento de Security Analytics. Los principales componentes que componen el complemento incluyen los detectores, los tipos de registro, las reglas, los resultados y las alertas.



Tipos de registro

OpenSearch admite varios tipos de registros y proporciona out-of-the-box mapeos para cada tipo. Al crear un detector, debe especificar el tipo de registro y configurar un intervalo de tiempo y, a partir de ahí, Security Analytics activa automáticamente un conjunto de reglas relevante que se ejecutan en ese intervalo.

Detectores

Los detectores identifican una variedad de amenazas de ciberseguridad para un tipo de registro en todos sus índices de datos. El detector se configura para que utilice tanto reglas personalizadas como reglas Sigma preconfiguradas que evalúan los eventos que ocurren en el sistema. A continuación, el detector genera los resultados de seguridad a partir de estos eventos. Para obtener más información sobre los detectores, consulte Creación de detectores en la OpenSearch documentación.

Reglas

Las reglas de detección de amenazas definen las condiciones que los detectores aplican a los datos de registro incorporados para identificar un evento de seguridad. Security Analytics permite importar, crear y personalizar reglas para cumplir con sus requisitos, y también proporciona reglas Sigma empaquetadas previamente y de código abierto para detectar las amenazas más comunes en sus registros. Security Analytics asigna muchas reglas a una base de conocimientos cada vez mayor sobre tácticas y técnicas de los adversarios, mantenida por la organización MITRE ATT&CK. Puede usar ambos OpenSearch paneles o el APIs para crear y usar reglas. Para obtener más información sobre las reglas, consulte Trabajar con reglas en la OpenSearch documentación.

Hallazgos

Cuando un detector hace coincidir una regla con un evento de registro, genera un resultado. Cada resultado incluye una combinación única de reglas seleccionadas, un tipo de registro y la gravedad de la regla. Los resultados no necesariamente apuntan a amenazas inminentes dentro del sistema, pero siempre aíslan un evento de interés. Para obtener más información sobre los hallazgos, consulte Trabajar con los hallazgos en la OpenSearch documentación.

Alertas

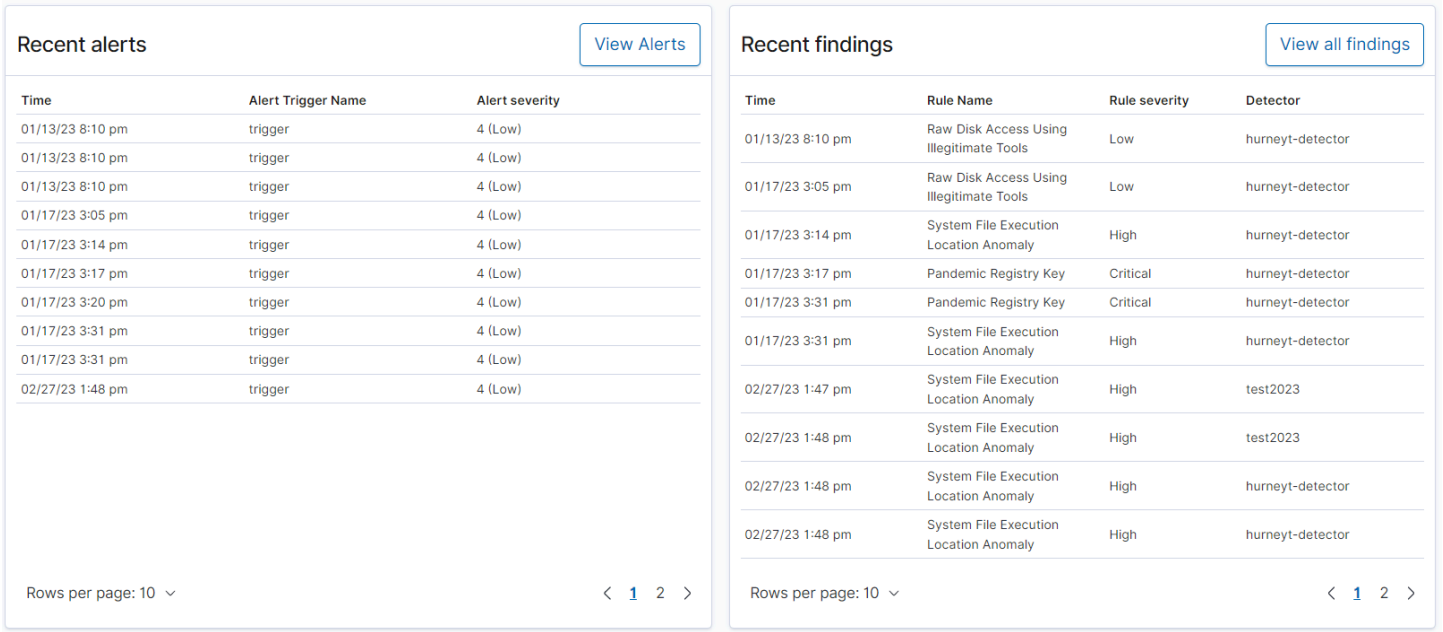
Al crear un detector, puede especificar una o varias condiciones que activan una alerta. Una alerta es una notificación que se envía a un canal preferido, como Slack o el correo electrónico. Puede configurar la alerta para que se active cuando el detector coincida con una o varias reglas y puede personalizar el mensaje de notificación. Para obtener más información sobre las alertas, consulte Trabajar con alertas en la OpenSearch documentación.

Exploración de Security Analytics

Puede usar los OpenSearch paneles de control para visualizar y obtener información sobre su complemento de análisis de seguridad. La vista Información general proporciona información como resultados y recuentos de alertas, resultados y alertas recientes, reglas de detección frecuentes y una lista de detectores. Puede ver una vista resumida compuesta por varias visualizaciones. El siguiente gráfico, por ejemplo, muestra la tendencia de los resultados y las alertas de varios tipos de registros durante un período de tiempo determinado.



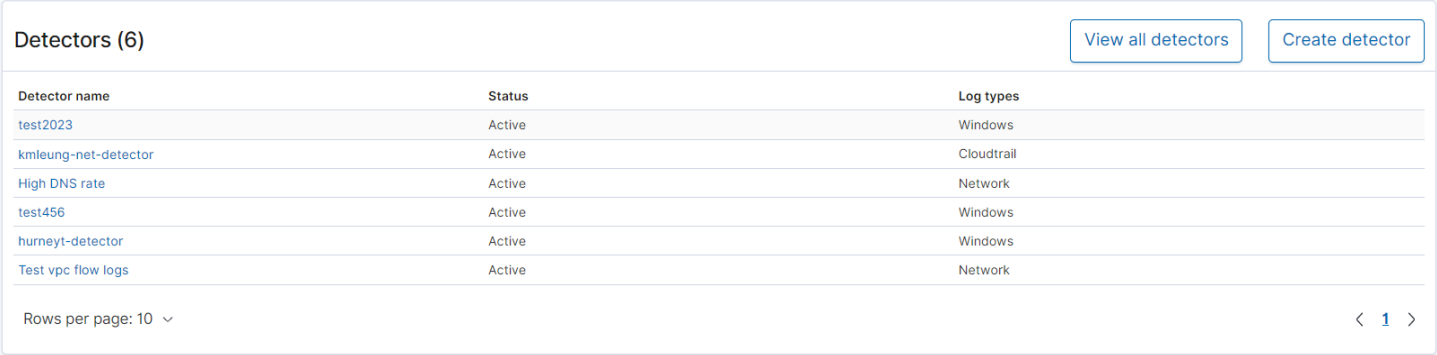
Más abajo en la página, puede revisar los resultados y alertas más recientes.



Además, puede ver una distribución de las reglas que se activan con más frecuencia en todos los detectores activos. Esto puede ayudarle a detectar e investigar diferentes tipos de actividades maliciosas en todos los tipos de registros.



Por último, puede ver el estado de los detectores configurados. Desde este panel, también puede acceder al flujo de trabajo de creación de detectores.



Para configurar Security Analytics, cree reglas en la página de Reglas y utilícelas para escribir los detectores en la página de Detectores. Para tener una visión más precisa de los resultados de Security Analytics, puede utilizar las páginas de Resultados y Alertas.

Configuración de permisos

Si habilita Security Analytics en un dominio de OpenSearch servicio preexistente, es posible que la `security_analytics_manager` función no esté definida en el dominio. Los usuarios que no sean administradores deben estar asignados a este rol para poder administrar índices templados en los dominios mediante un control de acceso detallado. Para crear el rol `security_analytics_manager` de forma manual, siga estos pasos:

1. En los OpenSearch paneles, vaya a Seguridad y elija Permisos.
2. Seleccione Crear grupo de acciones y configure los siguientes grupos:

Nombre del grupo	Permisos
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. Seleccione Roles y, a continuación, Crear rol.
4. Asigne el nombre security_analytics_manager al rol.

5. Para Permisos de clúster, seleccione `security_analytics_full_access` y `security_analytics_read_access`.
6. Para Índice, escriba `*`.
7. Para Permisos de índice, seleccione `indices:admin/mapping/put` y `indices:admin/mappings/get`.
8. Seleccione Crear.
9. Después de crear el rol, debe mapearlo a cualquier rol de usuario o backend que administre los índices de Security Analytics.

Resolución de problemas

No existe tal error de índice

Si no tiene detectores y abre el panel de Security Analytics, es posible que vea una notificación en la parte inferior derecha que diga `[index_not_found_exception] no such index [.opensearch-sap-detectors-config]`. Puede ignorar esta notificación, que desaparece en unos segundos y no volverá a aparecer una vez que crea un detector.

Observabilidad en Amazon Service OpenSearch

La observabilidad es la práctica de obtener información sobre el estado interno y el rendimiento de sus sistemas complejos mediante el examen de sus resultados. El monitoreo tradicional puede indicarle que su sistema está inactivo, mientras que la observabilidad le ayuda a entender por qué está inactivo, ya que le permite hacer nuevas preguntas sobre sus datos.

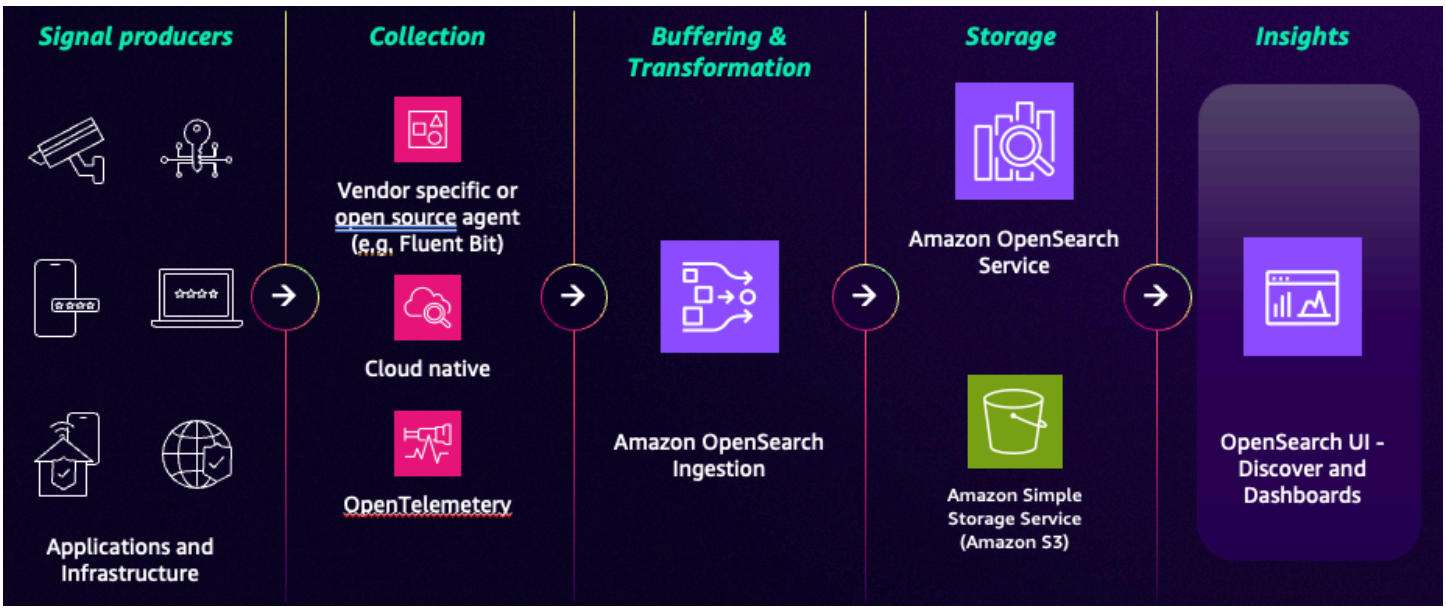
Amazon OpenSearch Service proporciona una solución unificada mediante la recopilación y correlación de tipos clave de datos de telemetría.

- Los registros proporcionan registros cronológicos de eventos, como errores de aplicaciones, solicitudes de los usuarios o mensajes de estado del sistema.
- Los rastros representan el end-to-end recorrido de una solicitud a medida que recorre los diferentes servicios de un sistema distribuido.

Al reunir estos tipos de datos, Amazon OpenSearch Service ayuda a los equipos de operaciones, a los ingenieros de fiabilidad del sitio y a los desarrolladores a detectar, diagnosticar y resolver los problemas operativos con mayor rapidez.

El flujo de trabajo de observabilidad en Service OpenSearch

Para transferir los datos de sus aplicaciones a OpenSearch Service para su análisis, se utiliza un proceso de varias etapas:



1. Recopilación

   El proceso comienza en el origen, con agentes o colectores ligeros que recopilan datos de telemetría de los productores de señales, como las aplicaciones y la infraestructura. Los dos agentes de código abierto más comunes son:

   - OpenTelemetry(OTel) colector: el método estándar y preferido del sector para recopilar registros y trazas.
   - Fluent Bit: un procesador y reenviador de registros liviano y de alto rendimiento que admite el OTel esquema de registros y trazas.

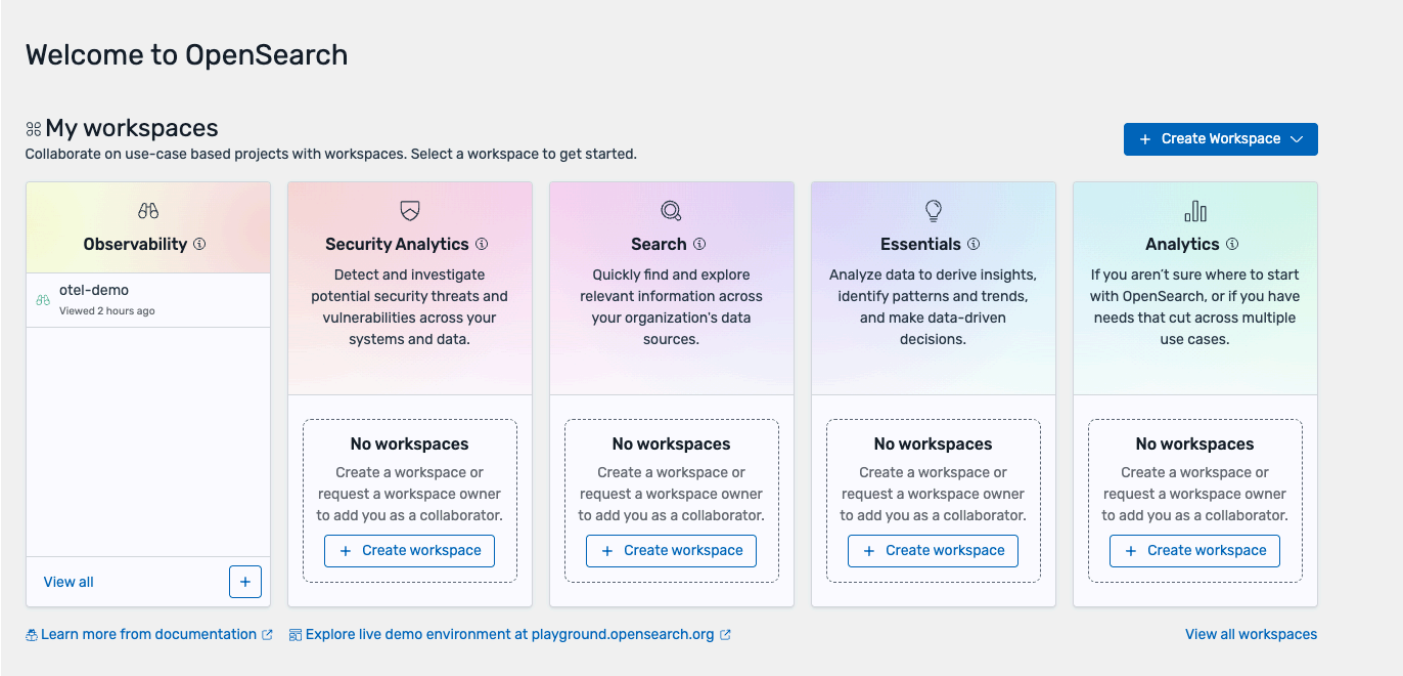
2. Ingestión (Amazon OpenSearch Ingestion)

   Una vez recopilada la telemetría, los datos se envían a OpenSearch Ingestion, una canalización de datos totalmente gestionada y sin servidor. Puede crear canalizaciones personalizadas para:

   - Filtrar: elimine los datos innecesarios para reducir los costos de almacenamiento.
   - Enriquecer: añada metadatos valiosos, como información geográfica de una dirección IP.
   - Transforme y normalice: estructure los registros no estructurados en un formato coherente.
   - Enrutar: envíe distintos tipos de datos a distintos índices OpenSearch de servicio o a Amazon S3.

3. Análisis y visualización

   Tras el procesamiento, los datos se cargan en un dominio o colección de OpenSearch servicios. Puede almacenar, indexar y analizar grandes cantidades de datos prácticamente en tiempo real. Interactúa con estos datos a través de una interfaz de visualización, como el espacio de trabajo de observabilidad de la interfaz de OpenSearch usuario, para ejecutar consultas, crear paneles y configurar alertas.



OpenSearch Interfaz de usuario y paneles OpenSearch

OpenSearch El servicio proporciona dos interfaces de usuario distintas para facilitar la observación. Le recomendamos que utilice la OpenSearch interfaz de usuario y configure un espacio de trabajo de observabilidad para las nuevas instalaciones y que migre desde los paneles de control existentes OpenSearch . A continuación, se muestra una tabla que describe las ventajas de la OpenSearch interfaz de usuario en comparación con los paneles de control tradicionales. OpenSearch

Característica	OpenSearch INTERFAZ DE USUARIO	OpenSearch Cuadros de mando
Orígenes de datos	Varias fuentes: se pueden conectar a varios dominios OpenSearch de servicio, colecciones OpenSearch sin servidor y otras fuentes de datos.	Fuente única: ubicada junto a un único dominio de servicio. OpenSearch
Actualizaciones	Las nuevas funciones llegan primero aquí porque no están	Las nuevas funciones están vinculadas a la OpenSearc h versión. Es posible que

Característica	OpenSearch INTERFAZ DE USUARIO	OpenSearch Cuadros de mando
	vinculadas a una versión específica OpenSearch .	las actualizaciones queden obsoletas en el futuro.
Disponibilidad.	Se alojan para Nube de AWS garantizar que no haya tiempo de inactividad durante las actualizaciones de los clústeres.	Puede dejar de estar disponibl e temporalmente durante el mantenimiento y las actualiza ciones del dominio
Colaboración	Cuenta con espacios de trabajo para una colaboración en equipo selecta en flujos de trabajo específicos.	La colaboración se basa en compartir los objetos guardados en un único dominio.

Nota: Para que empezar sea más fácil, hemos creado un nuevo flujo de trabajo de introducción para los registros en la consola de Amazon OpenSearch Service que configurará una nueva canalización de ingesta OTel personalizada, te permitirá seleccionar un OpenSearch clúster existente y crear una nueva aplicación de OpenSearch interfaz de usuario con un espacio de trabajo de observabilidad creado. Lo único que tiene que hacer es dirigir a sus OTel agentes al nuevo punto final de ingestión y estará listo para obtener información sobre sus OTel datos formateados.

Dado que la ingesta y el análisis se gestionan de forma diferente en los registros y en los seguimientos, hemos creado secciones independientes para profundizar en ellas.

Registros

OpenSearch La ingestión puede transformar los datos de registro no estructurados en un formato estructurado durante la ingestión. OpenSearch La ingestión proporciona procesadores que normalizan y enriquecen los datos antes de indexarlos. Algunos ejemplos de procesadores útiles son:

- `grok`— Analiza y estructura los datos de texto no estructurados, como los registros de acceso al servidor web, en campos distintos.

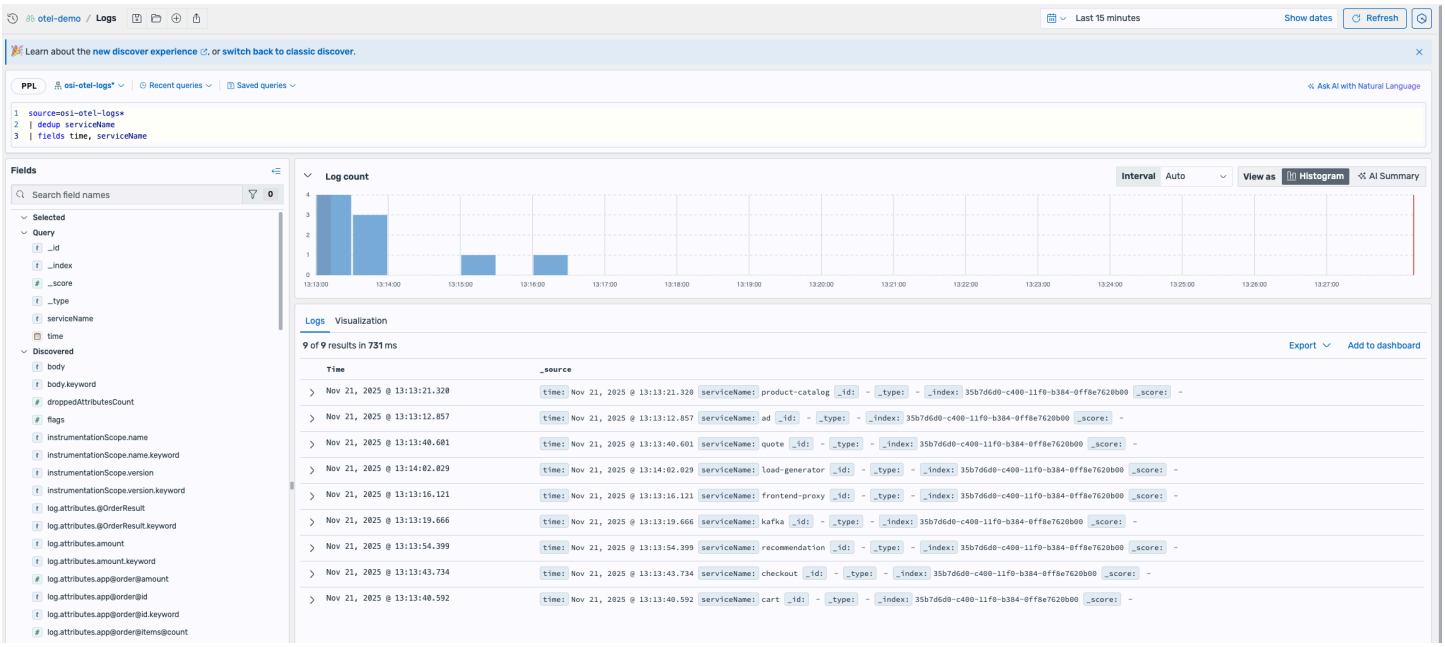
- `date`— Analiza una fecha de un campo de registro y la establece como marca temporal del evento.
- `parse_json`— Analiza un campo de cadena que contiene un objeto JSON.

Nota: Para facilitar el inicio, hemos creado un nuevo flujo de trabajo de introducción para los registros en la consola de Amazon OpenSearch Service que configurará un nuevo canal de ingestión personalizado de Otel, lo dirigirá a un OpenSearch clúster existente y creará una nueva aplicación de OpenSearch interfaz de usuario con un espacio de trabajo de observabilidad creado. Todo lo que tiene que hacer es dirigir a sus agentes de Otel al nuevo punto de ingestión.

OpenSearch Espacio de trabajo de interfaz de usuario y observabilidad

Una vez que los datos de tus registros se hayan incorporado a Amazon OpenSearch Service, utilizarás las herramientas que proporciona el espacio de trabajo de observabilidad de Amazon OpenSearch Service en la OpenSearch interfaz de usuario para analizarlos. El espacio de trabajo de observabilidad proporciona herramientas especializadas diseñadas para extraer información significativa en Discover y Dashboards.

El espacio de trabajo de observabilidad incluye una nueva experiencia Discover que utiliza el lenguaje de procesamiento canalizado (PPL) complementado con un asistente de lenguaje natural desarrollado por Amazon Q Developer for Business. La asistencia lingüística facilita que cualquiera pueda empezar a utilizar los lenguajes canalizados. Tras afinar la consulta, cree visualizaciones y paneles directamente desde el nuevo Discover sin tener que pasar a otras partes de la herramienta. Para consultar sus datos mediante DQL o SQL, cambie a la antigua experiencia de Discover.



Consulta tus registros mediante PPL

Dispone de varias opciones para consultar sus registros a fin de recopilar información sobre el funcionamiento de su aplicación o servicio.

El lenguaje de procesamiento por canalización (PPL) es un lenguaje de consulta con una sintaxis (|) basada en canalizaciones para encadenar comandos. Puede usarlo para crear expresiones potentes para analizar sus registros.

Nota: Para desbloquear el PPL más reciente commands/functions en la versión OpenSearch 2.19, tendrás que cambiar el indicador de una función en las herramientas para OpenSearch desarrolladores mediante la siguiente consulta (no es necesaria en la versión OpenSearch 3.3):

```
PUT /_plugins/_query/settings { "transient" : { "plugins.calcite.enabled" : true } }
```

Busca los hosts con más errores

En este ejemplo, se analizan los registros para determinar los hosts de servicio con el mayor número total de errores.

```
source = my-index |
    where level = "ERROR" |
    stats count() as error_count by host |
    sort -error_count |
```

```
    head 5
```

Calcula el tiempo medio de solicitud

En este ejemplo, se analizan los registros para calcular el tiempo medio de solicitud de cada código de estado del registro.

```
source = my-index |
    stats avg(request_time) by status_code
```

Para obtener más información sobre la PPL, consulta el manual de referencia de la PPL en opensearch.org.

Consulta tus registros mediante IA

En este ejemplo, se analizan los registros para mostrar los errores registrados en los últimos 5 minutos.

```
Show me all of the error logs from the last 5 minutes
```



Consulta tus registros mediante SQL

SQL proporciona una forma familiar de consultar los datos de registro.

En este ejemplo, se analizan los registros para mostrar los errores por marca de tiempo.

```
SELECT timestamp, severity_text, body, service_name
FROM opentelemetry_logs
WHERE severity_text = 'ERROR' AND service_name = 'my-service'
ORDER BY timestamp DESC;
```

Para obtener más información sobre SQL, consulte el manual de referencia de SQL en. GitHub

Consulta de los registros mediante DQL

El DQL es bueno para buscar y filtrar rápidamente.

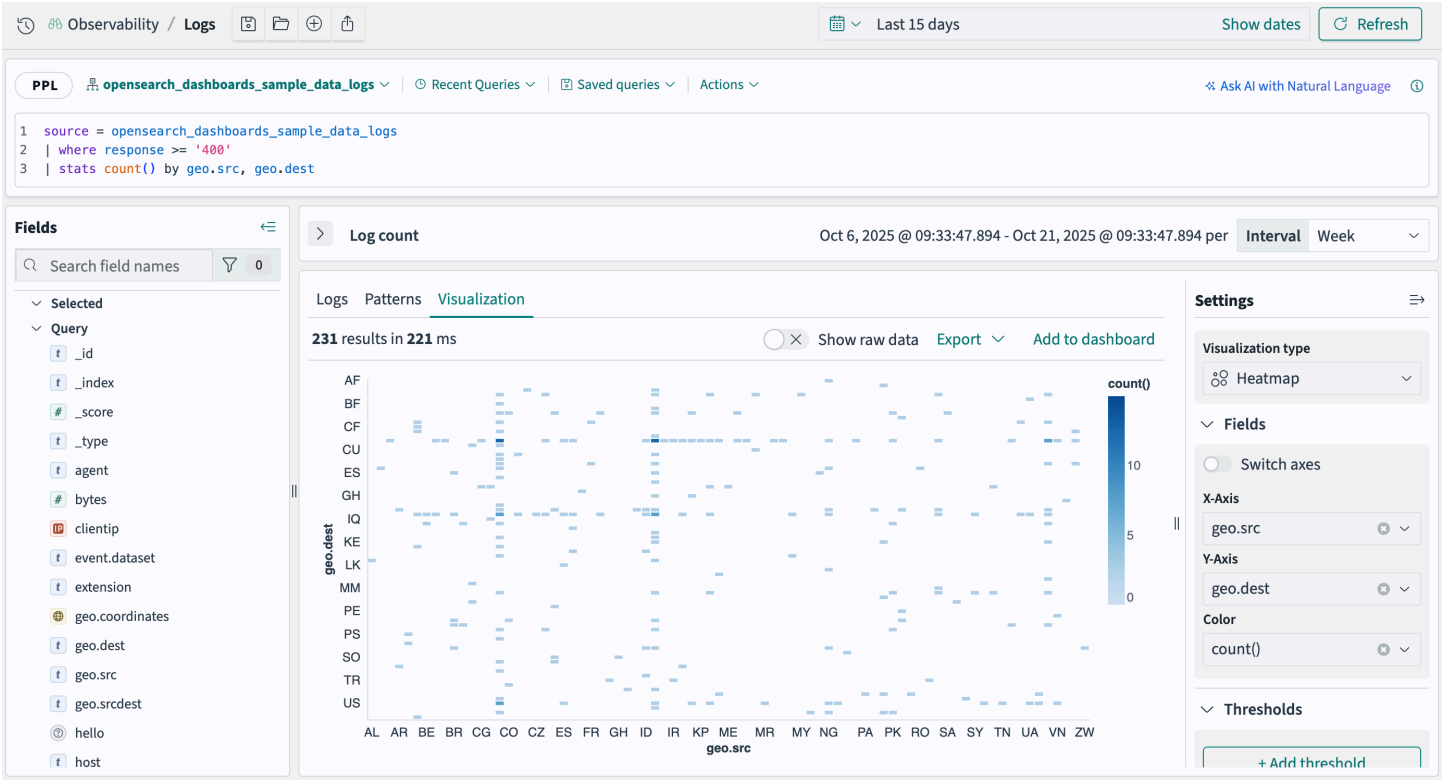
En este ejemplo se analizan los registros y se muestran errores y excepciones.

```
error OR exception
```

Para obtener más información sobre DQL, consulte el manual de referencia de DQL en opensearch.org.

Cuadros de mando y alertas para los registros

En la nueva experiencia de Discover con PPL, puede crear visualizaciones desde la pestaña de visualizaciones de Discover. Elija entre 12 tipos de visualización y edítelos sobre la marcha antes de añadirlos a un panel de control. En la antigua experiencia de Discover, buscaría Visualize en el panel de navegación de la izquierda para crear una nueva visualización y Dashboards para añadir las visualizaciones a sus paneles.



Puede definir los monitores de alertas mediante PPL o el DSL de consultas de OpenSearch servicio para ejecutar consultas programadas. Una condición de activación, como un número específico

de registros de errores, activa una alerta. Puedes enviar notificaciones a través de canales como Amazon Simple Notification Service o webhooks.

Para obtener más información sobre las alertas, consulta la documentación de alertas en opensearch.org.

Rastros

Amazon OpenSearch Service ofrece funciones integrales de rastreo distribuido que le ayudan a comprender el rendimiento de las aplicaciones y a diagnosticar problemas en toda su arquitectura de microservicios. Al ingerir OpenTelemetry (OTel) los datos de rastreo con OpenSearch Ingestion, OpenSearch Service procesa y estructura automáticamente la información de telemetría, lo que le brinda end-to-end visibilidad de los flujos de solicitudes a través de sus sistemas distribuidos.

Procesamiento e ingestión de datos de rastreo

OpenSearch Ingestion proporciona procesadores especializados que normalizan y enriquecen los datos de rastreo durante la ingestión, lo que garantiza que la telemetría siga patrones consistentes y esté lista para el análisis. Los procesadores clave para los datos de rastreo incluyen:

- `service_map`— Crea automáticamente gráficos de dependencia de los servicios a partir de las relaciones entre tramos, que muestran cómo fluyen las solicitudes entre los servicios.
- `trace_group`— Agrega los intervalos relacionados en grupos de rastreo lógicos en función de los atributos del intervalo de entrada, como el método HTTP y la ruta.
- `otel_trace_raw`— Procesa los datos de OpenTelemetry rastreo sin procesar y extrae los atributos del tramo, los atributos de los recursos y la información del alcance de la instrumentación en campos en los que se pueden realizar búsquedas.

OpenSearch Espacio de trabajo de interfaz de usuario y observabilidad

Una vez que tus datos de rastreo se hayan incorporado a Amazon OpenSearch Service, utilizarás las herramientas que proporciona el espacio de trabajo de observabilidad de Amazon OpenSearch Service en la OpenSearch interfaz de usuario para analizarlos. El espacio de trabajo de observabilidad proporciona visualizaciones y herramientas de análisis especializadas diseñadas para ayudarte a entender el rendimiento del servicio, identificar los cuellos de botella y solucionar problemas en tu arquitectura distribuida.

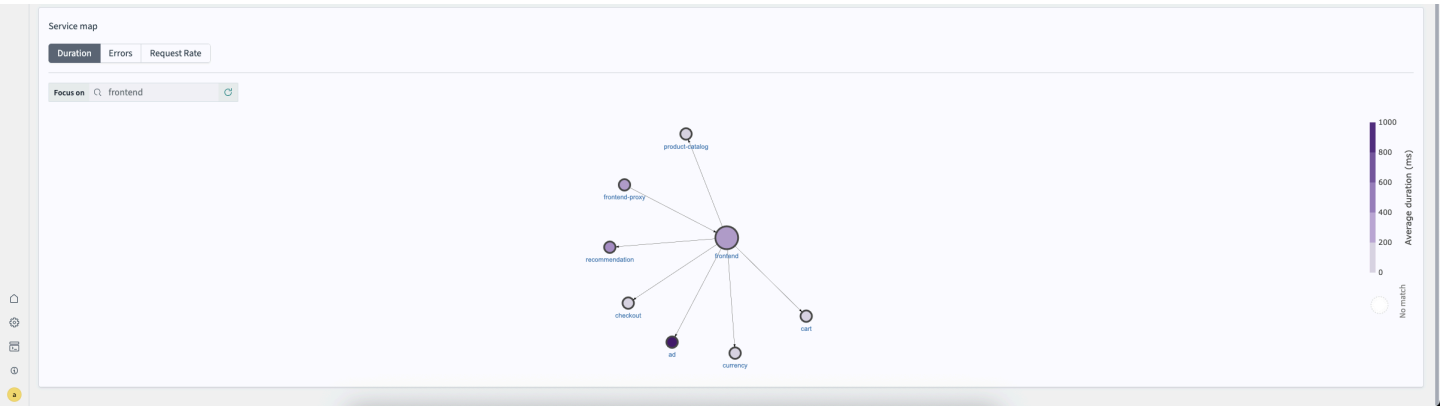
El espacio de trabajo de observabilidad incluye una vista de servicios que muestra las métricas RED (tasa, tasa de error y duración) de todos los servicios instrumentados, junto con un mapa de servicios interactivo que muestra las dependencias y los patrones de comunicación. La vista de trazas le permite buscar trazas específicas mediante trazas IDs o tramos y, a continuación IDs, analizar en detalle los gráficos de cascada y los intervalos para comprender el recorrido completo de las solicitudes a través del sistema.

Características principales de

Vista de servicios

La vista de servicios proporciona una visión general completa del estado y el rendimiento de su aplicación mediante:

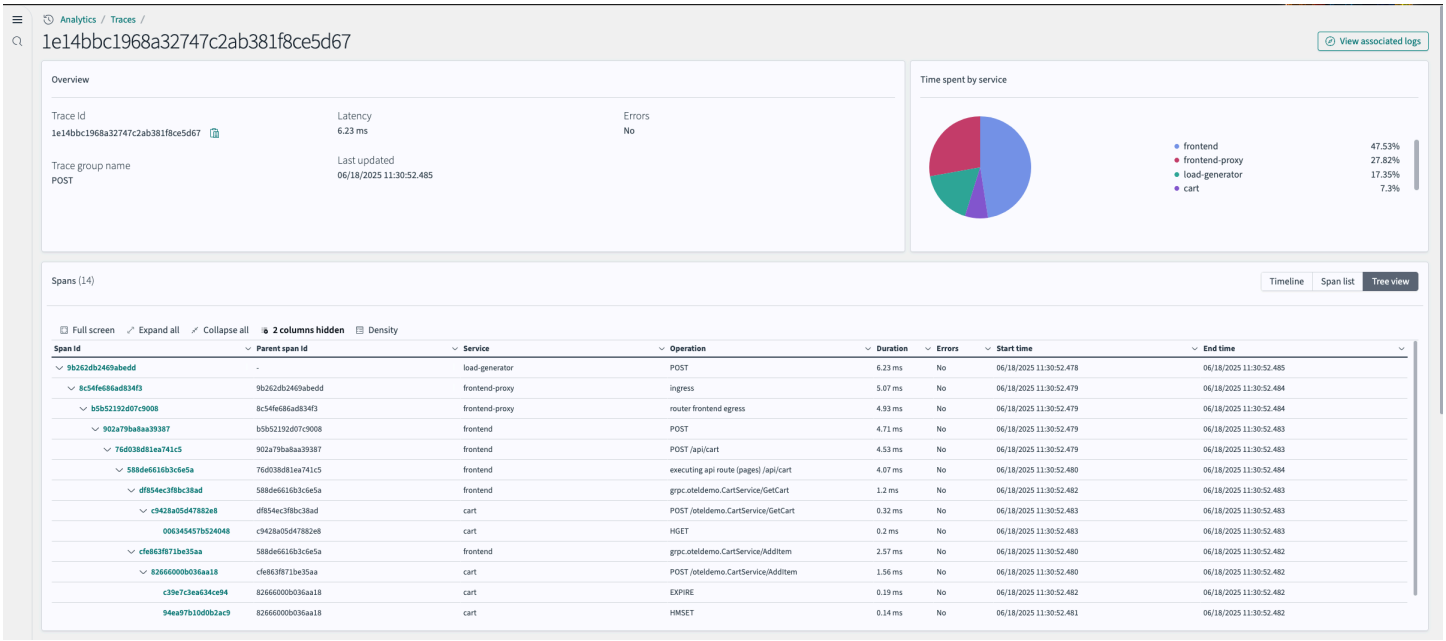
- Panel de métricas de RED: controle la tasa (solicitudes por segundo), la tasa de errores (porcentaje de solicitudes fallidas) y la duración (percentiles de latencia) de cada servicio de su sistema distribuido. Estas métricas le proporcionan información inmediata sobre el estado del servicio y le ayudan a identificar rápidamente la degradación del rendimiento.
- Mapa de servicios interactivo: visualice cómo se comunican sus servicios entre sí mediante un gráfico de dependencias generado automáticamente. El mapa de servicios muestra los flujos de solicitudes entre los servicios, lo que le ayuda a comprender la arquitectura del sistema e identificar los cuellos de botella o los fallos en cascada.
- Indicadores de estado del servicio: identifique rápidamente los servicios problemáticos en función de las tasas de error y los umbrales de latencia. Los servicios están codificados por colores para destacar aquellos que requieren atención inmediata, lo que facilita la priorización de las iniciativas de solución de problemas.
- Cuadro de diálogo de correlación de servicios: consulte en profundidad cualquier servicio para analizar los registros y trazas relacionados. Esta vista integrada conecta las métricas de nivel de servicio con los datos de rastreo detallados y las entradas de registro asociadas, lo que permite un análisis más rápido de la causa raíz sin tener que cambiar de una herramienta a otra.



Vista de trazas

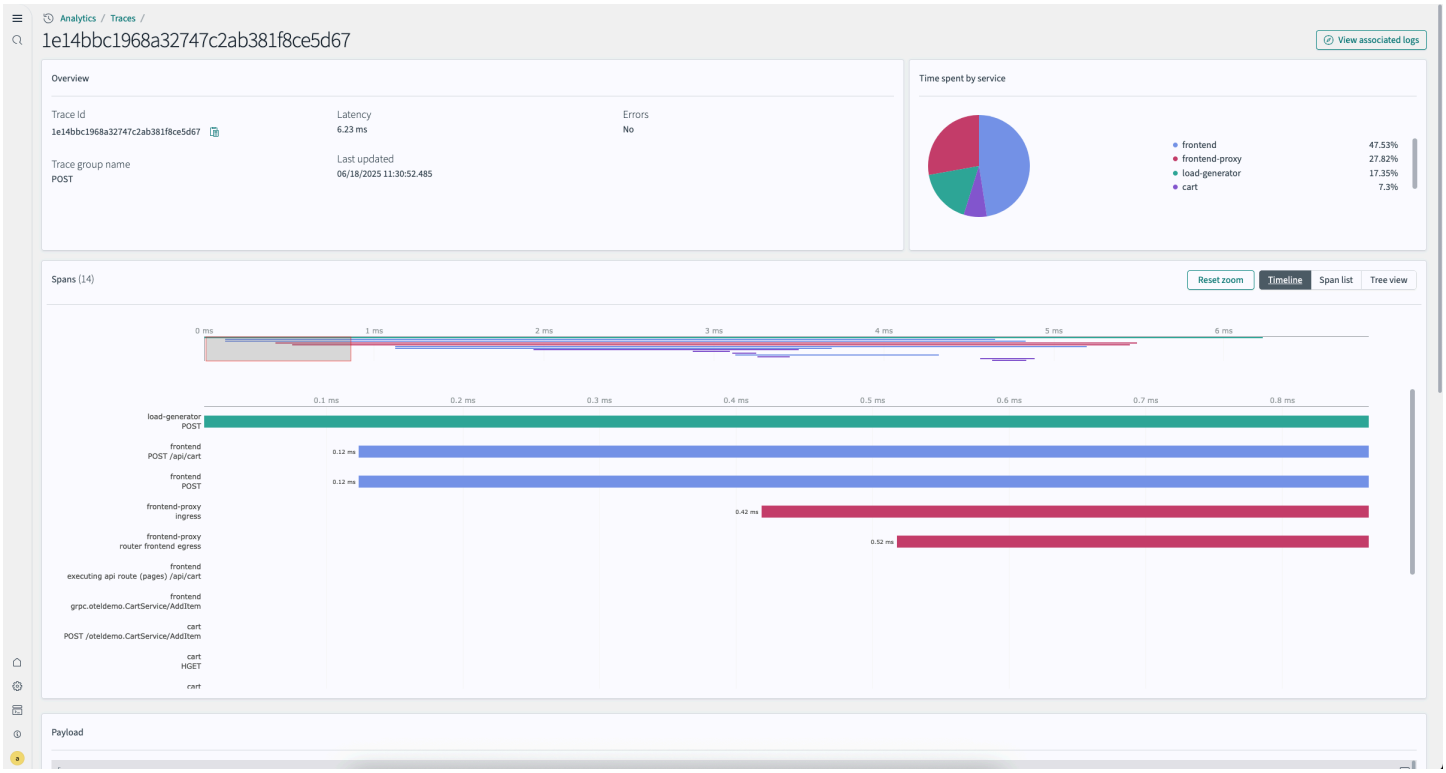
La vista Traces permite investigar en profundidad las solicitudes individuales a través de su sistema distribuido:

- Agrupación de trazas por método y ruta HTTP: organiza automáticamente las trazas en grupos lógicos en función de los puntos finales de la API y muestra métricas agregadas, como la latencia media, la tasa de errores y las tendencias de rendimiento a lo largo del tiempo. Esto le ayuda a identificar qué puntos finales tienen problemas y a realizar un seguimiento de las mejoras en el rendimiento.
- Búsqueda de ID de rastreo e ID de tramo: localice rápidamente trazas específicas mediante identificadores de rastreo o identificadores de tramo. Esto resulta especialmente útil cuando se investigan los problemas notificados por los usuarios o se correlacionan con los registros de errores que contienen el contexto del rastreo.
- Gráficos de cascada: visualice el cronograma completo de una solicitud a medida que avanza en sus servicios. La vista en cascada muestra el tiempo y la duración de los intervalos, lo que facilita la identificación de las operaciones lentas, el procesamiento secuencial frente al paralelo y dónde se dedica el tiempo en el sistema distribuido.
- Vista de árbol con desglose jerárquico de intervalos: explore las relaciones padre-hijo entre los intervalos para comprender la jerarquía de llamadas dentro de un rastreo. Esta vista le ayuda a ver cómo se ramifica una solicitud entre los servicios e identificar qué llamadas de servicio contribuyen a la latencia general.
- Panel de registros asociados: vea los registros que se produjeron durante el mismo período de tiempo que su rastreo, filtrados por el servicio relevante y el contexto del rastreo. Esta correlación entre los rastreos y los registros mejora considerablemente la solución de problemas, ya que proporciona tanto el flujo de solicitudes como los registros detallados de la aplicación en una única interfaz.



Capacidades avanzadas

- Análisis de correlación: vincule sin problemas las trazas, los intervalos y los servicios con los registros correspondientes. El espacio de trabajo de observabilidad correlaciona automáticamente los datos de telemetría mediante el contexto de rastreo, lo que le permite cambiar entre distintas vistas de la misma solicitud sin perder el contexto.
- Nombres de índice personalizados y compatibilidad entre clústeres: configure el OpenSearch servicio para leer los datos de rastreo de patrones de índice personalizados o de varios clústeres. OpenSearch Esta flexibilidad admite escenarios de implementación complejos y le permite organizar los datos de telemetría de acuerdo con sus necesidades operativas.
- Límites del mapa de servicios configurables: ajuste la cantidad de servicios y conexiones que se muestran en el mapa de servicios para gestionar topologías a gran escala. En el caso de sistemas con cientos de servicios, puede filtrar el mapa para centrarse en subconjuntos de servicios específicos o ajustar los límites de representación para mantener el rendimiento.
- Navegación con minimapas para diagramas de Gantt: navegue por gráficos de cascadas de gran tamaño de forma eficiente utilizando la vista general del minimapa. Esta función resulta especialmente útil a la hora de analizar trazos con muchos intervalos, ya que te permite pasar rápidamente a diferentes secciones de la línea temporal.



Traces proporciona at-a-glance visibilidad del rendimiento de las aplicaciones en función de los datos del protocolo OpenTelemetry (OTel). Le ayuda a comprender cómo fluyen las solicitudes a través de su sistema distribuido mediante el seguimiento de su end-to-end recorrido entre los servicios.

Mejores prácticas operativas para Amazon OpenSearch Service

Este capítulo proporciona las mejores prácticas para operar los dominios de Amazon OpenSearch Service e incluye pautas generales que se aplican a muchos casos de uso. Cada carga de trabajo es única, con características únicas, por lo que ninguna recomendación genérica es exactamente adecuada para cada caso de uso. La práctica general más importante consiste en implementar, probar y ajustar sus dominios en un ciclo continuo para encontrar la configuración, la estabilidad y el costo óptimos para su carga de trabajo.

Temas

- Monitorización y alertas
- Estrategia de particiones
- Stability
- Desempeño
- Seguridad
- Optimización de costos
- CloudWatch Alarmas recomendadas para Amazon OpenSearch Service
- Dimensionamiento de los dominios de Amazon OpenSearch Service
- Escala de petabytes en Amazon Service OpenSearch
- Nodos coordinadores dedicados en Amazon OpenSearch Service
- Nodos maestros dedicados en Amazon OpenSearch Service

Monitorización y alertas

Las siguientes prácticas recomendadas se aplican a la supervisión de sus dominios OpenSearch de servicio.

Configure CloudWatch las alarmas

OpenSearch El servicio emite métricas de rendimiento a Amazon CloudWatch. Revisa periódicamente las métricas de tu clúster e instancia y configura las CloudWatch alarmas recomendadas en función del rendimiento de tu carga de trabajo.

Habilitación de la publicación de registros

OpenSearch El servicio expone los registros de OpenSearch errores, los registros lentos de búsqueda, los registros lentos de indexación y los registros de auditoría en Amazon CloudWatch Logs. Los registros lentos de búsqueda, los registros lentos de índice y los registros de errores son útiles para la resolución de problemas de rendimiento y estabilidad. Los registros de auditoría, que solo están disponibles si habilita control de acceso preciso, realizan un seguimiento de la actividad del usuario. Para obtener más información, consulte los registros en la OpenSearch documentación.

Los registros lentos de búsqueda e indexación de registros lentos son una herramienta importante para comprender y solucionar problemas del rendimiento de sus operaciones de búsqueda e indexación. Habilite la entrega lenta de registros de búsqueda e índice para todos los dominios de producción. También debe configurar los umbrales de registro; de lo contrario, CloudWatch no capturará los registros.

Estrategia de particiones

Los fragmentos distribuyen la carga de trabajo entre los nodos de datos de su OpenSearch dominio de servicio. Los índices configurados correctamente pueden ayudar a mejorar el rendimiento general del dominio.

Cuando envías datos al OpenSearch Servicio, los envías a un índice. Un índice es análogo a una tabla de base de datos, con documentos como las filas y campos como las columnas. Al crear el índice, indica OpenSearch cuántos fragmentos principales desea crear. Los fragmentos principales son particiones independientes del conjunto de datos completo. OpenSearch El servicio distribuye automáticamente los datos entre los fragmentos principales de un índice. También puede configurar réplicas del índice. Cada partición de réplica incluye un conjunto completo de copias de las particiones principales de ese índice.

OpenSearch El servicio mapea los fragmentos de cada índice en los nodos de datos del clúster. Garantiza que las particiones principales y de réplica del índice residan en nodos de datos diferentes. La primera réplica garantiza que tenga dos copias de los datos en el índice. Siempre debe usar al menos una réplica. Las réplicas adicionales proporcionan redundancia y capacidad de lectura adicionales.

OpenSearch envía solicitudes de indexación a todos los nodos de datos que contienen fragmentos que pertenecen al índice. Envía solicitudes de indexación primero a los nodos de datos que contienen particiones principales y, después, a los nodos de datos que contienen particiones de

réplica. El nodo coordinador dirige las solicitudes de búsqueda, ya sea a una partición principal o a una réplica, para todas las particiones que pertenecen al índice.

Por ejemplo, para un índice con cinco particiones principales y una réplica, cada solicitud de indexación toca 10 particiones. En contraste, las solicitudes de búsqueda se envían a las particiones n, donde n es el número de particiones principales. Para un índice con cinco particiones principales y una réplica, cada consulta de búsqueda toca cinco particiones (principales o réplicas) de ese índice.

Determinar los recuentos de particiones y nodos de datos

Utilice las siguientes prácticas recomendadas para determinar los recuentos de nodos de datos y particiones para su dominio.

Tamaño de las particiones: el tamaño de los datos en el disco es un resultado directo del tamaño de los datos de origen y cambia a medida que se indexan más datos. La source-to-index proporción puede variar considerablemente, de 1:10 a 10:1 o más, pero por lo general es de alrededor de 1:1,10. Puede usar esa relación para predecir el tamaño del índice en el disco. Puede también indexar algunos datos y recuperar los tamaños reales del índice para determinar la relación de la carga de trabajo. Después de tener un tamaño de índice previsto, establezca un recuento de particiones para que cada partición tenga entre 10 y 30 GiB (para cargas de trabajo de búsqueda) o entre 30 y 50 GiB (para cargas de trabajo de registros). 50 GiB debería ser el máximo; asegúrese de planificar el crecimiento.

Recuento de particiones: la distribución de particiones a los nodos de datos tiene un gran impacto en el rendimiento de un dominio. Cuando tenga índices con múltiples particiones, procure que la cantidad de particiones sea un múltiplo de la cantidad de nodos de datos. Esto ayuda a garantizar que las particiones se distribuyan de manera uniforme entre los nodos de datos y evita los nodos activos. Por ejemplo, si tiene 12 particiones principales, el recuento de nodos de datos debe ser 2, 3, 4, 6 o 12. Sin embargo, el recuento de particiones es secundario al tamaño de la partición; si tiene 5 GiB de datos, debe seguir utilizando una sola partición.

Particiones por nodo de datos: el número total de particiones que puede contener un nodo es proporcional a la memoria en montón de máquina virtual Java (JVM) del nodo. Trate de obtener 25 particiones o menos por GiB de memoria en montón. Por ejemplo, un nodo con 32 GiB de memoria en montón no debe contener más de 800 particiones. Si bien la distribución de los fragmentos puede variar en función de los patrones de carga de trabajo, hay un límite de 1000 fragmentos por nodo para Elasticsearch y de OpenSearch 1,1 a 2,15 y 4000 para los de 2,17 o más. OpenSearch La API

cat/allocation proporciona una vista rápida de la cantidad de particiones y el almacenamiento total de particiones en los nodos de datos.

Proporción de partición a CPU: cuando una partición participa en una solicitud de indexación o búsqueda, utiliza una vCPU para procesar la solicitud. Como práctica recomendada, utilice un punto de escala inicial de 1,5 vCPU por partición. Si su tipo de instancia tiene 8 vCPUs, configure el recuento de nodos de datos de manera que cada nodo no tenga más de seis fragmentos. Tenga en cuenta que se trata de una aproximación. Asegúrese de probar la carga de trabajo y escalar el clúster en consecuencia.

Para obtener recomendaciones sobre el volumen de almacenamiento, el tamaño de las particiones y el tipo de instancia, consulte los recursos siguientes:

- the section called “Determinación del tamaño de dominios”
- the section called “Escala de petabytes”

Evite el sesgo de almacenamiento

El sesgo de almacenamiento ocurre cuando uno o más nodos de un clúster contienen una mayor proporción de almacenamiento para uno o más índices que los demás. Los indicios de que hay sesgo de almacenamiento incluyen utilización desigual de la CPU, latencia intermitente y desigual, y colocación en cola desigual en los nodos de datos. Para determinar si tiene problemas de sesgo, consulte las siguientes secciones de solución de problemas:

- the section called “Partición de nodos y sesgo de almacenamiento”
- the section called “Partición del índice y el sesgo de almacenamiento”

Stability

Las siguientes prácticas recomendadas se aplican para mantener un dominio de OpenSearch servicio estable y en buen estado.

Manténgase al día con OpenSearch

Actualizaciones del software del servicio

OpenSearch El servicio publica periódicamente actualizaciones de software que añaden funciones o mejoran sus dominios. Las actualizaciones no cambian la versión del motor de búsqueda

OpenSearch ni la de Elasticsearch. Te recomendamos programar una hora periódica para ejecutar la operación de la DescribeDomainAPI e iniciar una actualización del software del `UpdateStatus` servicio, si es así. `ELIGIBLE` Si no actualizas tu dominio dentro de un período de tiempo determinado (normalmente dos semanas), el OpenSearch Servicio realizará la actualización automáticamente.

OpenSearch actualizaciones de versión

OpenSearch El servicio añade periódicamente soporte para las versiones mantenidas por la comunidad de OpenSearch. Actualice siempre a las OpenSearch versiones más recientes cuando estén disponibles.

OpenSearch El servicio actualiza simultáneamente ambos OpenSearch OpenSearch paneles (o Elasticsearch y Kibana si tu dominio ejecuta un motor antiguo). Si el clúster tiene nodos maestros dedicados, las actualizaciones se completan sin tiempo de inactividad. De lo contrario, es posible que el clúster deje de responder durante varios segundos después de la actualización mientras elige un nodo principal. OpenSearch Es posible que los paneles no estén disponibles durante una parte o durante toda la actualización.

Hay dos formas de actualizar un dominio:

- Actualización local: esta opción es más fácil porque mantiene el mismo clúster.
- Actualización de instantáneas y restauración: esta opción es buena para probar nuevas versiones en un clúster nuevo o migrar entre clústeres.

Independientemente del proceso de actualización que utilice, recomendamos mantener un dominio que sea únicamente para desarrollo y pruebas y actualizarlo a la nueva versión antes de actualizar el dominio de producción. Para el tipo de implementación, seleccione Desarrollo y pruebas cuando cree el dominio de prueba. Asegúrese de actualizar todos los clientes a versiones compatibles inmediatamente después de la actualización del dominio.

Cómo mejorar el rendimiento de las instantáneas

Para evitar que la instantánea se bloquee durante el procesamiento, el tipo de instancia del nodo maestro dedicado debe coincidir con el número de particiones. Para más información, consulte the section called “Elección de tipos de instancias para nodos principales dedicados”. Además, cada nodo no debe tener más de las 25 particiones recomendadas por cada GiB de memoria dinámica de Java. Para más información, consulte the section called “Selección del número de particiones”.

Habilite los nodos maestros dedicados

Nodos maestros dedicados para mejorar la estabilidad de los clústeres. Un nodo maestro dedicado realiza tareas de administración de clústeres, pero no contiene datos de índice ni responde a las solicitudes de los clientes. Al asumir de este modo las tareas de administración de clústeres, aumenta la estabilidad de su dominio y hace posible que se lleven a cabo algunos cambios de configuración sin tiempo de inactividad.

Habilite y utilice tres nodos maestros dedicados para lograr una estabilidad óptima del dominio en tres zonas de disponibilidad. La implementación con Multi-AZ con modo de espera le permite configurar tres nodos maestros dedicados. Para obtener recomendaciones de tipos de instancias, consulte the section called “Elección de tipos de instancias para nodos principales dedicados”.

Implemente en varias zonas de disponibilidad

Para evitar que se pierdan datos y minimizar el tiempo de inactividad del clúster en caso de que se produzca una interrupción del servicio, puede distribuir los nodos en dos o tres zonas de disponibilidad de la misma Región de AWS. La práctica recomendada es realizar la implementación mediante Multi-AZ con modo de espera, que configura tres zonas de disponibilidad: dos zonas activas y una en espera, y con dos particiones de réplicas por índice. Esta configuración permite a OpenSearch Service distribuir los fragmentos de réplica a fragmentos AZs distintos de los principales correspondientes. No hay cargos por transferencia de datos entre zonas de disponibilidad para las comunicaciones de clústeres entre zonas de disponibilidad.

Las zonas de disponibilidad son ubicaciones aisladas dentro de cada región. Con una configuración de dos zonas de disponibilidad, perder una zona de disponibilidad significa que se pierde la mitad de toda la capacidad del dominio. El cambio a tres zonas de disponibilidad reduce aún más el impacto de perder una sola zona de disponibilidad.

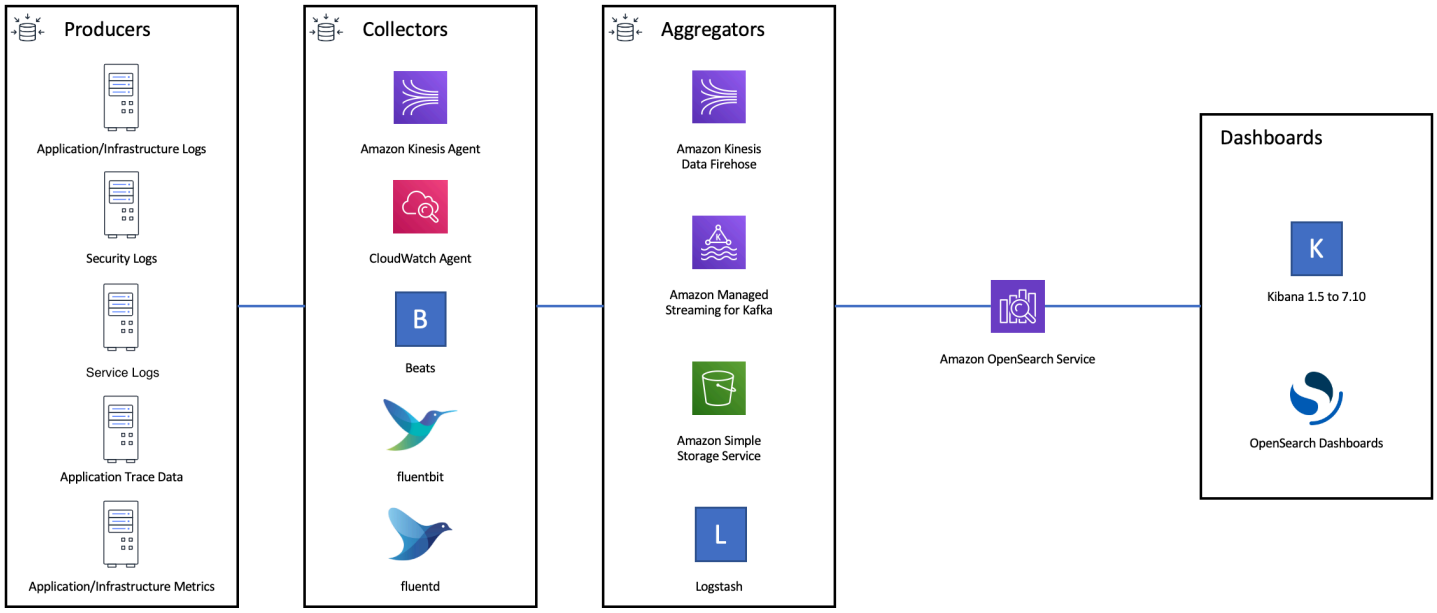
Controle el flujo de incorporación y el almacenamiento en búfer

Recomendamos que se limite el recuento total de solicitudes mediante la operación de la API _bulk. Es más eficiente enviar una solicitud `_bulk` que contiene 5000 documentos que enviar 5000 solicitudes que contienen un solo documento.

Para lograr una estabilidad operativa óptima, a veces es necesario limitar o incluso detener el flujo ascendente de las solicitudes de indexación. Limitar la tasa de solicitudes de índices es un mecanismo importante para hacer frente a picos inesperados u ocasionales en las solicitudes que,

de otro modo, podrían abrumar al clúster. Considere la posibilidad de crear un mecanismo de control de flujo en su arquitectura ascendente.

El siguiente diagrama muestra varias opciones de componentes para una arquitectura de incorporación de registros. Configure la capa de agregación para dejar espacio suficiente para almacenar en búfer los datos entrantes para picos de tráfico repentinos y un breve mantenimiento del dominio.



Cree asignaciones para cargas de trabajo de búsqueda

Para las cargas de trabajo de búsqueda, cree mapeos que definan cómo se OpenSearch almacenan e indexan los documentos y sus campos. Establezca `dynamic` en `strict` para evitar la adición accidental de nuevos campos.

```
PUT my-index
{
  "mappings": {
    "dynamic": "strict",
    "properties": {
      "title": { "type" : "text" },
      "author": { "type" : "integer" },
      "year": { "type" : "text" }
    }
  }
}
```

Utilice plantillas de índice

Puedes usar una plantilla de índice como una forma de indicar OpenSearch cómo configurar un índice cuando se crea. Configure las plantillas de índices antes de crear índices. Luego, cuando crea un índice, hereda la configuración y las asignaciones de la plantilla. Puede aplicar más de una plantilla a un único índice, de manera que pueda especificar la configuración en una plantilla y las asignaciones en otra. Esta estrategia permite una plantilla para configuraciones comunes en varios índices y plantillas separadas para configuraciones y asignaciones más específicas.

Los siguientes parámetros son útiles para configurar en plantillas:

- Número total de particiones primarias y de réplicas
- Intervalo de actualización (frecuencia con la que se actualizan y se realizan cambios recientes en el índice disponibles para la búsqueda)
- Control de mapeo dinámico
- Asignaciones de campo explícito

La siguiente plantilla de ejemplo contiene cada una de estas configuraciones:

```
{
   "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"
         }
      }
   }
}
```

Aunque cambien con poca frecuencia, tener la configuración y los mapeos definidos de forma centralizada OpenSearch es más fácil de administrar que actualizar varios clientes anteriores.

Administre los índices con la administración de estado de índice

Si administra registros o datos de serie temporal, le recomendamos que utilice Administración de estados de índice (ISM). ISM le permite automatizar las tareas regulares de administración del ciclo de vida. Con ISM, puede crear políticas que invoquen la renovación de alias de índice, tomen instantáneas de índices, muevan índices entre niveles de almacenamiento y eliminen índices antiguos. Incluso puede usar la operación sustitución de ISM como estrategia alternativa de administración del ciclo de vida de los datos para evitar el sesgo de las particiones.

En primer lugar, establezca una política de ISM. Por ejemplo, consulte the section called “Ejemplos de política”. A continuación, puede adjuntar la política a uno o más índices. Si incluye un campo de plantilla de ISM en la política, OpenSearch Service aplicará automáticamente la política a cualquier índice que coincida con el patrón especificado.

Eliminar los índices que no se utilizan

Revise periódicamente los índices de su clúster e identifique los que no estén en uso. Tome una instantánea de esos índices para que se almacenen en S3 y, a continuación, elimínelos. Cuando se eliminan los índices no utilizados, se reduce el recuento de particiones y se permite una distribución del almacenamiento de información y una utilización de recursos más equilibradas entre los nodos. Incluso cuando están inactivos, los índices consumen algunos recursos durante las actividades de mantenimiento de índices internos.

En lugar de eliminar manualmente los índices no utilizados, puede usar ISM para tomar automáticamente una instantánea y eliminar los índices después de un periodo de tiempo determinado.

Utilice varios dominios para disfrutar de una alta disponibilidad

Para lograr una alta disponibilidad más allá del 99,9 % de tiempo de actividad en varias regiones, considere la posibilidad de usar dos dominios. Para conjuntos de datos pequeños o que cambian lentamente, puede configurar la replicación entre clústerespara mantener un modelo activo-pasivo. En este modelo, solo se escribe el dominio principal, pero se puede leer cualquiera de los dos dominios. Para conjuntos de datos más grandes y datos que cambian rápidamente, configure la entrega dual en la canalización de incorporación para que todos los datos se escriban de forma independiente en ambos dominios en un modelo activo-activo.

Diseñe sus aplicaciones ascendentes y descendentes teniendo en cuenta la conmutación por error. Asegúrese de probar el proceso de conmutación por error junto con otros procesos de recuperación de desastres.

Desempeño

Las siguientes prácticas recomendadas se aplican para ajustar sus dominios para lograr un rendimiento óptimo.

Optimice el tamaño y la compresión de solicitudes

El tamaño masivo depende de los datos, el análisis y la configuración del clúster, pero un buen punto de partida es de 3 a 5 MiB por solicitud masiva.

Envíe solicitudes y reciba respuestas de sus OpenSearch dominios mediante la compresión gzip para reducir el tamaño de la carga útil de las solicitudes y respuestas. Puedes usar la compresión gzip con el cliente OpenSearch Python o incluir los siguientes encabezados desde el lado del cliente:

- `'Accept-Encoding': 'gzip'`
- `'Content-Encoding': 'gzip'`

Para optimizar el tamaño de las solicitudes, comience con un tamaño de solicitud de 3 MiB. Luego, aumente lentamente el tamaño de la solicitud hasta que el desempeño de la indexación deje de mejorar.

> **Note**
>
> Para habilitar la compresión gzip en dominios que ejecutan Elasticsearch versión 6.x, debe establecer `http_compression.enabled` a nivel de clúster. Esta configuración se cumple de forma predeterminada en las versiones 7.x de Elasticsearch y en todas las versiones de. OpenSearch

Reduzca el tamaño de las respuestas de solicitudes masivas

Para reducir el tamaño de OpenSearch las respuestas, excluya los campos innecesarios con el parámetro. `filter_path` Asegúrese de no filtrar ningún campo que sea necesario para identificar

o reintentar las solicitudes fallidas. Para obtener más información y ejemplos, consulta the section called “Reducción del tamaño de la respuesta”.

Ajuste intervalos de actualización

OpenSearch los índices tienen una coherencia de lectura eventual. Una operación de actualización hace que todas las actualizaciones que se realizan en un índice estén disponibles para la búsqueda. El intervalo de actualización predeterminado es de un segundo, lo que OpenSearch significa que se actualiza cada segundo mientras se escribe en un índice.

Cuando menos frecuente sea la actualización de un índice (mayor intervalo de actualización), mejor será el rendimiento general de la indexación. La desventaja de aumentar el intervalo de actualización es que hay un retraso mayor entre la actualización del índice y el momento en que los nuevos datos están disponibles para la búsqueda. Establezca el intervalo de actualización tan alto como pueda tolerar para mejorar el rendimiento general.

Recomendamos configurar el parámetro `refresh_interval` para todos los índices a 30 segundos o más.

Habilite el ajuste automático

Auto-Tune utiliza las métricas de rendimiento y uso del OpenSearch clúster para sugerir cambios en el tamaño de las colas, el tamaño de la caché y la configuración de la máquina virtual Java (JVM) en los nodos. Estos cambios opcionales mejoran la velocidad y la estabilidad del clúster. Puede volver a la configuración de OpenSearch servicio predeterminada en cualquier momento. La función de ajuste automático está habilitada de forma predeterminada en nuevos dominios, a menos que la desactive explícitamente.

Recomendamos que se habilite el ajuste automático en todos los dominios y se establezca un período de mantenimiento periódico o se revisen periódicamente sus recomendaciones.

Seguridad

Las siguientes prácticas recomendadas se aplican a la protección de sus dominios.

Cómo habilitar el control de acceso detallado

El control de acceso detallado le permite controlar quién puede acceder a determinados datos dentro de un dominio del Servicio. OpenSearch En comparación con el control de acceso generalizado, el control de acceso detallado proporciona a cada clúster, índice, documento y campo su propia política

de acceso especificada. Los criterios de acceso pueden basarse en una serie de factores, incluido el rol de la persona quien solicita el acceso y la acción que pretende realizar con los datos. Por ejemplo, puede dar a un usuario acceso para escribir en un índice, y a otra se le puede dar acceso solo para leer los datos del índice sin realizar cambios.

El control de acceso detallado permite que los datos con diferentes requisitos de acceso existan en el mismo espacio de almacenamiento sin tener problemas de seguridad o cumplimiento.

Se recomienda habilitar un control de acceso detallado para sus dominios.

Implemente dominios dentro de una VPC

Colocar su dominio de OpenSearch servicio en una nube privada virtual (VPC) permite una comunicación segura entre el OpenSearch Servicio y otros servicios dentro de la VPC, sin necesidad de una puerta de enlace a Internet, un dispositivo NAT o una conexión VPN. Todo el tráfico permanece seguro dentro de la nube. AWS Debido a su aislamiento lógico, los dominios que residen dentro de una VPC tienen una capa adicional de seguridad en comparación con los dominios que utilizan puntos de conexión públicos.

Le recomendamos que utilice crear sus dominios dentro de una VPC.

Aplique una política de acceso restrictivo

Incluso si su dominio se implementa dentro de una VPC, es una práctica recomendada la implementación de la seguridad en capas. Asegúrese de comprobar la configuración de sus políticas de acceso actuales.

Aplica a tus dominios una política de acceso restrictiva basada en los recursos y sigue el principio de privilegios mínimos al conceder acceso a la API de configuración y a las operaciones de la OpenSearch API. Como regla general, evite usar la entidad principal de usuario anónimo `"Principal": {"AWS": "*" }` en sus políticas de acceso.

Sin embargo, hay algunas situaciones en las que es aceptable usar una política de acceso abierto, como cuando se habilita un control de acceso detallado. Una política de acceso abierto puede permitirle acceder al dominio en los casos en que la firma de solicitudes sea difícil o imposible, por ejemplo, desde ciertos clientes y herramientas.

Habilite el cifrado en reposo

OpenSearch Los dominios de servicio ofrecen el cifrado de los datos en reposo para evitar el acceso no autorizado a los datos. El cifrado en reposo utiliza AWS Key Management Service (AWS KMS)

para almacenar y administrar las claves de cifrado, y el algoritmo estándar de cifrado avanzado con claves de 256 bits (AES-256) para realizar el cifrado.

Si el dominio almacena información confidencial, habilite el cifrado de datos en reposo.

Habilita el cifrado node-to-node

Node-to-node El cifrado proporciona un nivel de seguridad adicional además de las funciones de seguridad predeterminadas del OpenSearch Servicio. Implementa la seguridad de la capa de transporte (TLS) para todas las comunicaciones entre los nodos que están aprovisionados. OpenSearch Node-to-nodecifrado: todos los datos que se envíen a su dominio de OpenSearch servicio a través de HTTPS permanecen cifrados en tránsito mientras se distribuyen y replican entre los nodos.

Si tu dominio almacena datos confidenciales, habilita el node-to-node cifrado.

Supervisa con AWS Security Hub CSPM

Supervise su uso del OpenSearch Servicio en relación con las mejores prácticas de seguridad mediante el uso de AWS Security Hub CSPM. Security Hub CSPM utiliza controles de seguridad para evaluar las configuraciones de los recursos y los estándares de seguridad para cumplir con varios marcos de conformidad. Para obtener más información sobre el uso de Security Hub CSPM para evaluar los recursos del OpenSearch servicio, consulte los Amazon OpenSearch Service controles en la Guía del AWS Security Hub usuario.

Optimización de costos

Las siguientes prácticas recomendadas se aplican a la optimización y el ahorro de los costes del OpenSearch servicio.

Use los tipos de instancia de última generación

OpenSearch El servicio siempre adopta nuevos tipos de instancias de Amazon EC2 que ofrecen un mejor rendimiento a un coste menor. Recomendamos usar siempre las instancias de última generación.

Evite usar instancias T2 o `t3.small` para dominios de producción, ya que pueden volverse inestables bajo una carga pesada sostenida. Las instancias `r6g.large` son una opción para cargas de trabajo de producción pequeñas (tanto como nodos de datos, como nodos maestros dedicados).

Use los volúmenes gp3 de Amazon EBS más recientes

OpenSearch los nodos de datos requieren un almacenamiento de baja latencia y alto rendimiento para proporcionar una indexación y consultas rápidas. Utilizando volúmenes gp3 de Amazon EBS, se obtiene un mayor rendimiento de referencia (IOPS y rendimiento) a un costo un 9,6 % inferior que con el tipo de volumen gp2 de Amazon EBS que se ofrecía anteriormente. Puede aprovisionar IOPS y rendimiento adicionales independientemente del tamaño de los volúmenes mediante gp3. Además, estos volúmenes son más estables que los de la generación anterior, ya que no utilizan créditos de ráfaga. El tipo de volumen gp3 también duplica los límites de tamaño de per-data-node volumen del tipo de volumen gp2. Con estos volúmenes de mayor tamaño, puede reducir el costo de los datos pasivos mediante el incremento de la cantidad de almacenamiento por nodo de datos.

Utilice UltraWarm y almacene en frío los datos de registro de series temporales

Si lo utiliza OpenSearch para el análisis de registros, traslade sus datos a un UltraWarm almacenamiento en frío para reducir los costes. Utilice la Administración de estado de índices (ISM) para migrar datos entre niveles de almacenamiento y administrar la retención de datos.

UltraWarmproporciona una forma rentable de almacenar grandes cantidades de datos de solo lectura en OpenSearch Service. UltraWarm utiliza Amazon S3 para el almacenamiento, lo que significa que los datos son inmutables y solo se necesita una copia. Solo paga por un almacenamiento que es equivalente al tamaño de las particiones principales de sus índices. Las latencias de las UltraWarm consultas aumentan con la cantidad de datos de S3 que se necesitan para atender la consulta. Una vez que los datos se han almacenado en caché en los nodos, las consultas a los UltraWarm índices tienen un rendimiento similar al de las consultas a los índices activos.

El almacenamiento en frío también cuenta con el respaldo de S3. Cuando necesite consultar datos inactivos, puede adjuntarlos de forma selectiva a los nodos existentes. UltraWarm Los datos inactivos incurren en el mismo coste de almacenamiento gestionado que los objetos almacenados en frío UltraWarm, pero los objetos almacenados en frío no consumen recursos de los UltraWarm nodos. Por lo tanto, el almacenamiento en frío proporciona una cantidad significativa de capacidad de almacenamiento sin afectar al tamaño o al recuento de los UltraWarm nodos.

UltraWarm se vuelve rentable cuando tiene que migrar aproximadamente 2,5 TiB de datos desde un almacenamiento activo. Supervise su tasa de llenado y planifique trasladar los índices a ellos UltraWarm antes de alcanzar ese volumen de datos.

Revise las recomendaciones para instancias reservadas

Considere la posibilidad de adquirir instancias reservadas (RIs) una vez que tenga una buena base de referencia sobre el rendimiento y el consumo de cómputo. Los descuentos comienzan en torno al 30 % para reservas sin pago anticipado de 1 año y pueden aumentar hasta un 50 % para todos los compromisos de pagos iniciales de 3 años.

Después de observar un funcionamiento estable durante al menos 14 días, revise Acceder a las recomendaciones de reservas en la Guía del usuario de AWS Cost Management . El encabezado Amazon OpenSearch Service muestra recomendaciones de compra específicas de RI y ahorros proyectados.

CloudWatch Alarmas recomendadas para Amazon OpenSearch Service

CloudWatch las alarmas realizan una acción cuando una CloudWatch métrica supera un valor específico durante un período de tiempo determinado. Por ejemplo, es posible que AWS desee enviarle un correo electrónico si el estado del clúster es `red` superior a un minuto. En esta sección se incluyen algunas alarmas recomendadas para Amazon OpenSearch Service y cómo responder a ellas.

Puede implementar estas alarmas automáticamente mediante CloudFormation. Para ver una pila de muestras, consulta el GitHubrepositorio relacionado.

> **Note**
>
> Si despliegas la CloudFormation pila, las `KMSKeyInaccessible` alarmas `KMSKeyError` y permanecerán en un `Insufficient Data` estado dado que estas métricas solo aparecen si un dominio encuentra un problema con su clave de cifrado.

Para obtener más información sobre la configuración de alarmas, consulte Creación de CloudWatch alarmas de Amazon en la Guía del CloudWatch usuario de Amazon.

Alarma	Problema
El valor máximo de `ClusterSt`	Al menos una partición principal y sus réplicas no están asignados a un nodo. Consulte the section called “Estado rojo del clúster”.

Alarma	Problema
`atus.red` es >= 1 durante 1 minuto, 1 periodo consecutivo	
El valor máximo de `ClusterSt atus.yellow` es >= 1 durante 1 minuto, 5 periodos consecuti vos	Al menos una partición de réplica no está asignada a un nodo. Consulte the section called “Estado amarillo del clúster”.
El valor mínimo de `FreeStorageSpace` es <= 20 480 durante 1 minuto, 1 periodo consecutivo	El espacio de almacenamiento disponible de un nodo en su clúster se redujo hasta los 20 GiB. Consulte the section called “Falta de espacio de almacenamiento disponible”. Este valor está en MiB, por lo que, en lugar de 20 480, recomendamos que lo configure en un 25 % del espacio de almacenamiento de cada nodo.
El valor de `ClusterIn dexWrites Blocked` es >= 1 durante 5 minutos, 1 periodo consecutivo	El clúster bloquea las solicitudes de escritura. Consulte the section called “ClusterBlockException”.
El valor mínimo de `Nodes` es < x 1 día, 1 periodo consecutivo	x es el número de nodos del clúster. Esta alarma indica que al menos un nodo del clúster se mantuvo inaccesible durante un día. Consulte the section called “Nodos de clúster defectuosos”.

Alarma	Problema

El valor máximo de `Automated SnapshotFailure` es >= 1 durante 1 minuto, 1 periodo consecutivo	Se produjo un error en una instantánea automatizada. Este error suele ser el resultado de un estado rojo del clúster. Consulte the section called “Estado rojo del clúster”.
	
	Para ver un resumen de todas las instantáneas automatizadas e información sobre los errores, también puede probar las siguientes solicitudes:
	
	`GET domain_endpoint /_snapshot/cs-automated/_all`
	`GET domain_endpoint /_snapshot/cs-automated-enc/_all`

El valor máximo de `CPUUtilization` o `WarmCPUUtilization` es >= 80 % durante 15 minutos, 3 periodos consecutivos	A veces puede producirse un uso de CPU del 100 %, pero el uso sostenido elevado es problemático. Considere la posibilidad de utilizar tipos de instancias más grandes o de agregar instancias.
El valor máximo de `JVMMemoryPressure` es >= 95 % durante 1 minuto, 3 periodos consecutivos	El clúster podría encontrar errores de memoria insuficiente si aumenta el uso. Considere la posibilidad de escalar verticalmente. OpenSearch El servicio utiliza la mitad de la RAM de una instancia para el montón de Java, hasta un tamaño de pila de 32 GiB. Puede escalar las instancias verticalmente hasta 64 GiB de RAM y después escalarlas horizontalmente mediante el agregado de instancias.
El valor máximo de `OldGenJVMMemoryPressure` es >= 80 % durante 1 minuto, 3 periodos consecutivos	

Alarma	Problema
El valor máximo de `MasterCPUUtilization` es >= 50 % durante 15 minutos, 3 periodos consecutivos	Considere la posibilidad de utilizar tipos de instancias más grandes para los nodos maestros dedicados. Debido al rol que desempeñan en la estabilidad del clúster y las implementaciones azul/verde, los nodos maestros dedicados deberían tener un uso de CPU medio menor que los nodos de datos.
El valor máximo de `MasterJVMMemoryPressure` es >= 95 % durante 1 minuto, 3 periodos consecutivos	
El valor máximo de `MasterOldGenJVMMemoryPressure` es >= 80 % durante 1 minuto, 3 periodos consecutivos	
El valor de `KMSKeyError` es >= 1 durante 1 minuto, 1 periodo consecutivo	La clave de AWS KMS cifrado que se usa para cifrar los datos en reposo de tu dominio está deshabilitada. Vuelva a habilitarla para restablecer las operaciones normales. Para más información, consulte the section called “Cifrado en reposo”.
El valor de `KMSKeyInaccessible` es >= 1 durante 1 minuto, 1 periodo consecutivo	La clave de AWS KMS cifrado que se utiliza para cifrar los datos inactivos de tu dominio se ha eliminado o ha revocado su concesión al Servicio. OpenSearch No puede recuperar los dominios que están en este estado. Sin embargo, si tiene una instantánea manual, puede utilizarla para migrar a un nuevo dominio. Para más información, consulte the section called “Cifrado en reposo”.

Alarma	Problema
El valor de `shards.active` es >= 30 000 durante 1 minuto, 1 periodo consecutivo	El número total de particiones primarias y de réplicas activas es superior a 30 000. Es posible que esté rotando los índices con demasiada frecuencia. Considere la posibilidad de utilizar ISM para eliminar los índices una vez que alcancen una antigüedad determinada.
5*xx* alarmas >= 10 % de `OpenSearchRequests`	Es posible que uno o varios nodos de datos estén sobrecargados, o que las solicitudes no se completen dentro del periodo de tiempo de espera. Considere la posibilidad de cambiar a tipos de instancia más grandes o de agregar más nodos al clúster. Compruebe que está siguiendo las prácticas recomendadas para la arquitectura de particiones y clústeres.
El valor máximo de `MasterReachableFromNode` es < 1 durante 5 minutos, 1 periodo consecutivo	Esta alarma indica que el nodo maestro se ha detenido o es inaccesible. Estos errores suelen ser el resultado de un problema de conectividad de red o de AWS dependencia.
El valor medio de `ThreadpoolWriteQueue` es >= 100 durante 1 minuto, 1 periodo consecutivo	El clúster está experimentando una alta simultaneidad de indexación. Revise y controle las solicitudes de indexación, o aumente los recursos del clúster.
El valor medio de `ThreadpoolSearchQueue` es >= 500 durante 1 minuto, 1 periodo consecutivo	El clúster está experimentando una alta simultaneidad de búsqueda. Considere la posibilidad de escalar el clúster. También se puede aumentar el tamaño de la cola de búsqueda, pero si se incrementa excesivamente puede provocar errores de memoria.

Alarma	Problema
El valor máximo de `Threadpool SearchQueue` es >= 5000 durante 1 minuto, 1 periodo consecutivo	
El aumento de `Threadpool SearchRejected` SUM es >=1{expresión matemática DIFF ( )} durante 1 minuto, 1 periodo consecutivo	Estas alarmas le notifican los problemas del dominio que podrían afectar el rendimiento y la estabilidad.
El aumento de `Threadpool WriteRejected` SUM es >=1{expresión matemática DIFF ( )} durante 1 minuto, 1 periodo consecutivo	

> **Note**
>
> Si solo desea visualizar métricas, consulte the section called “Monitoreo de métricas del clúster”.

Otras alarmas para tener en cuenta

Considere la posibilidad de configurar las siguientes alarmas en función de las funciones del OpenSearch servicio que utilice habitualmente.

Alarma	Problema
`WarmFreeStorageSpace` es >=10 %	Has alcanzado el 10% del total de tu almacenamiento caliente gratuito. `WarmFreeStorageSpace` mide la suma del espacio de almacenamiento caliente libre en MiB. UltraWarm utiliza Amazon S3 en lugar de discos adjuntos.
El valor de `HotToWarmMigrationQueueSize` es >= 20 durante 1 minuto, 3 periodos consecutivos	Al mismo tiempo, un gran número de índices pasan del modo activo al UltraWarm almacenamiento. Considere la posibilidad de escalar el clúster.
El valor de `HotToWarmMigrationSuccessLatency` es >= 1 día, 1 periodo consecutivo	Si está intentando revertir los índices diarios, configure esta alarma para que se le notifique si el `HotToWarmMigrationSuccessCount` x latencia es superior a 24 horas.
El valor máximo de `WarmJVMMemoryPressure` es >= 95 % durante 1 minuto, 3 periodos consecutivos	El clúster podría encontrar errores de memoria insuficiente si aumenta el uso. Considere el escalado vertical. OpenSearch El servicio utiliza la mitad de la RAM de una instancia para el montón de Java, hasta un tamaño de pila de 32 GiB. Puede escalar las instancias verticalmente hasta 64 GiB de RAM y después escalarlas horizontalmente mediante el agregado de instancias.
El valor máximo de `WarmOldGenJVMMemoryPressure` es >= 80 % durante 1 minuto, 3 periodos consecutivos	

Alarma	Problema
El valor de `WarmToColdMigrationQueueSize` es >= 20 durante 1 minuto, 3 periodos consecutivos	Un gran número de índices se están trasladando simultáneamente al almacenamiento en frío. UltraWarm Considere la posibilidad de escalar el clúster.
El valor de `HotToWarmMigrationFailureCount` es >= 1 durante 1 minuto, 1 periodo consecutivo	Se pueden producir errores en las migraciones si se realizan durante instantáneas, reubicaciones de particiones o fusiones forzadas. Los errores durante las instantáneas o las reubicaciones de particiones suelen deberse a errores de nodo o problemas de conectividad de S3. La falta de espacio en el disco suele ser la causa subyacente de los errores en las fusiones forzosas.
El valor de `WarmToColdMigrationFailureCount` es >= 1 durante 1 minuto, 1 periodo consecutivo	Las migraciones suelen fallar cuando se produce un error en los intentos de migrar metadatos de índice al almacenamiento en frío. También pueden producirse errores al eliminar el estado de clúster de índice en caliente.
El valor de `WarmToColdMigrationLatency` es >= 1 día, 1 periodo consecutivo	Si está intentando revertir los índices diarios, configure esta alarma para que se le notifique si el `WarmToColdMigrationSuccessCount` x latencia es superior a 24 horas.
El valor de `AlertingDegraded` es >= 1 durante 1 minuto, 1 periodo consecutivo	El índice de alerta está en rojo, o uno o más nodos no ajustan a la programación.

Alarma	Problema
El valor de `ADPluginUnhealthy` es >= 1 durante 1 minuto, 1 periodo consecutivo	El complemento de detección de anomalías no funciona correctamente, ya sea debido a altas tasas de error o porque uno de los índices utilizados está en rojo.
El valor de `AsynchronousSearchFailureRate` es >= 1 durante 1 minuto, 1 periodo consecutivo	Al menos una búsqueda asíncrona ha fallado en el último minuto, lo que probablemente significa que el nodo coordinador ha fallado. El ciclo de vida de una solicitud de búsqueda asíncrona se administra únicamente en el nodo coordinador, por lo que si el coordinador cae, la solicitud falla.
El valor de `AsynchronousSearchStoreHealth` es >= 1 durante 1 minuto, 1 periodo consecutivo	El estado del almacén de respuestas de búsqueda asíncrona en el índice persistente está en rojo. Es posible que esté almacenando respuestas asíncronas de gran tamaño, lo que puede desestabilizar un clúster. Intente limitar las respuestas de búsqueda asíncronas a 10 MB o menos.
El valor de `SQLUnhealthy` es >= 1 durante 1 minuto, 3 periodos consecutivos	El complemento SQL devuelve 5 xx códigos de respuesta o pasa una consulta de DSL no válida a ella. OpenSearch Solucione los problemas de las solicitudes que sus clientes hacen al complemento.
El valor de `LTRStatus.red` es >= 1 durante 1 minuto, 1 periodo consecutivo	Al menos uno de los índices necesarios para ejecutar el complemento Learning to Rank carece de particiones principales y no es funcional.

Dimensionamiento de los dominios de Amazon OpenSearch Service

No existe un método perfecto para dimensionar los dominios de Amazon OpenSearch Service. Sin embargo, si comienza por comprender sus necesidades de almacenamiento, el servicio y OpenSearch en sí mismo, puede realizar una estimación inicial fundamentada de sus necesidades de hardware. Este cálculo puede servir como un buen punto de partida para la mayoría de los aspectos esenciales de la determinación del tamaño de los dominios: probarlos con cargas de trabajo representativas y monitorear su rendimiento.

Temas

- Cálculo de requisitos de almacenamiento
- Selección del número de particiones
- Selección de tipos de instancias y pruebas

Cálculo de requisitos de almacenamiento

La mayoría OpenSearch de las cargas de trabajo se clasifican en una de estas dos amplias categorías:

- Índice de larga duración: se escribe código que procesa los datos en uno o más OpenSearch índices y, a continuación, se actualizan esos índices periódicamente a medida que cambian los datos de origen. Algunos ejemplos comunes son las búsquedas en sitios web, documentos y comercio electrónico.
- Índices continuos: los datos fluyen de manera continua en un conjunto de índices temporales, con un periodo de indexación y de retención (como un conjunto de índices diarios que se conservan durante dos semanas). Algunos ejemplos comunes son el análisis de registros, el procesamiento de series temporales y el análisis de secuencias de clics.

Para las cargas de trabajo de índices de larga duración, puede examinar el origen de datos en disco y determinar fácilmente la cantidad de espacio de almacenamiento que consumen. Si los datos provienen de varias fuentes, solo tiene que agregar esas fuentes juntas.

Para los índices continuos, puede multiplicar la cantidad de datos generados durante un periodo de tiempo representativo por el periodo de retención. Por ejemplo, si genera 200 MiB de datos de

registro por hora, o sea 4,7 GiB por día, eso significa que contará con un total de 66 GiB de datos en un momento dado si hubo un periodo de retención de dos semanas.

Sin embargo, el tamaño de los datos de origen es solo uno de los aspectos de las necesidades de almacenamiento. También debe considerar lo siguiente:

- Número de réplicas: cada réplica es una copia completa de la partición principal; el tamaño de almacén del índice muestra el tamaño que ocupan la partición principal y de réplica. De forma predeterminada, cada OpenSearch índice tiene una réplica. Recomendamos tener al menos una para evitar la pérdida de datos. Las réplicas también mejoran el rendimiento, por lo que tal vez convenga tener más si tiene una carga de trabajo que realiza muchas operaciones de lectura. Utilice `PUT /my-index/_settings` para actualizar la configuración `number_of_replicas` para su índice.
- OpenSearch sobrecarga de indexación: el tamaño en disco de un índice varía. El tamaño total de los datos de origen más el índice suele ser del 110 % del origen, y el índice de hasta el 10 % de los datos de origen. Después de indexar los datos, puede utilizar la API `_cat/indices?v` y el valor de `pri.store.size` para calcular la sobrecarga exacta. `_cat/allocation?v` también proporciona un resumen útil.
- Espacio reservado por el sistema operativo: de forma predeterminada, Linux reserva el 5 % del sistema de archivos para el usuario `root` para procesos críticos, recuperación del sistema y como medida de seguridad frente a problemas de fragmentación del disco.
- OpenSearch Sobrecarga del OpenSearch servicio: el servicio reserva el 20% del espacio de almacenamiento de cada instancia (hasta 20 GiB) para fusiones de segmentos, registros y otras operaciones internas.

  Debido a este máximo de 20 GiB, la cantidad total de espacio reservado puede variar drásticamente en función del número de instancias del dominio. Por ejemplo, un dominio podría tener tres instancias `m6g.xlarge.search`, cada una con 500 GiB de espacio de almacenamiento, lo que equivaldría a un total de 1,46 TiB. En este caso, el espacio reservado total solo es de 60 GiB. Otro dominio podría tener 10 instancias `m3.medium.search`, cada una con 100 GiB de espacio de almacenamiento, lo que equivaldría a un total de 0,98 TiB. Aquí, el espacio reservado total es de 200 GiB, aunque el primer dominio sea un 50 % mayor.

  En la siguiente fórmula, aplicamos una estimación “en el peor de los casos” de los gastos generales. Este cálculo incluye espacio libre adicional para ayudar a minimizar el impacto de los errores de nodo y las interrupciones en la zona de disponibilidad.

En resumen, si tiene 66 GiB de datos en un momento dado y quiere una réplica, el requisito de almacenamiento mínimo será aproximadamente 66 * 2 * 1,1 / 0,95 / 0,8 = 191 GiB. Se podría generalizar este cálculo de la manera siguiente:

Datos de origen * (1 más cantidad de réplicas) * (1 más sobrecarga de indexación)/(1: espacio reservado de Linux)/(1: sobrecarga de OpenSearch servicio) = requisito mínimo de almacenamiento

O puede utilizar esta versión simplificada:

Datos de origen * (1 + número de réplicas) * 1,45 = requisito mínimo de almacenamiento

La falta de espacio de almacenamiento suficiente es una de las causas más comunes de la inestabilidad del clúster. Por lo tanto, debe verificar los números cuando elige tipos de instancia, recuentos de instancias y volúmenes de almacenamiento.

Existen otras consideraciones respecto al almacenamiento:

- Si el requisito mínimo de almacenamiento es superior a 1 PB, consulte the section called “Escala de petabytes”.
- Si tiene índices continuos y desea usar una arquitectura caliente/templada, consulte the section called “UltraWarm almacenamiento”.

Selección del número de particiones

Una vez que conozca los requisitos de almacenamiento, puede investigar la estrategia de indexación. De forma predeterminada, en OpenSearch Service, cada índice se divide en cinco fragmentos principales y una réplica (10 fragmentos en total). Este comportamiento es diferente al del código abierto OpenSearch, que de forma predeterminada es un fragmento principal y uno de réplica. Como no se puede cambiar fácilmente el número de particiones principales de un índice existente, debe decidir el número de particiones antes de indexar el primer documento.

El objetivo general de la selección de un número de particiones es distribuir un índice de manera uniforme entre todos los nodos de datos del clúster. Sin embargo, estas particiones no deberían ser demasiado grandes ni demasiado numerosas. Una pauta general es tratar de mantener el tamaño de la partición entre 10 y 30 GiB para las cargas de trabajo en las que la latencia de búsqueda es un objetivo de rendimiento clave y entre 30 y 50 GiB para las cargas de trabajo con mucha escritura, como el análisis de registros.

Los fragmentos grandes pueden dificultar la recuperación en caso de un error, pero dado que cada fragmento utiliza cierta cantidad de CPU y memoria, tener demasiados fragmentos pequeños

puede provocar problemas de rendimiento y errores de memoria insuficiente. OpenSearch En otras palabras, los fragmentos deben ser lo suficientemente pequeños como para que la instancia de OpenSearch servicio subyacente pueda gestionarlos, pero no tan pequeños como para que supongan una carga innecesaria para el hardware.

Por ejemplo, supongamos que tiene 66 GiB de datos. No espera que el número aumente a lo largo del tiempo y desea mantener las particiones en torno a los 30 GiB cada una. El número de particiones debería ser, por tanto, aproximadamente 66 * 1,1/30 = 3. Se podría generalizar este cálculo de la manera siguiente:

(Datos de origen + espacio para crecer) * (1 + sobrecarga de indexación) / tamaño de partición deseado = número aproximado de particiones principales

Esta ecuación ayuda a compensar el aumento de los datos a lo largo del tiempo. Si prevé que estos mismos 66 GiB de datos se cuadrupliquen a lo largo del próximo año, el número aproximado de particiones es (66 + 198) * 1,1/30 = 10. Recuerde, sin embargo, que aún no tiene esos 198 GiB de datos adicionales. Asegúrese de que estos preparativos para el futuro no creen particiones innecesariamente diminutas que consuman grandes cantidades de CPU y memoria en la actualidad. En este caso, tendrá 66 * 1,1/10 particiones = 7,26 GiB por partición, lo que consumirá recursos adicionales y queda debajo del intervalo de tamaño recomendado. Podrías considerar el middle-of-the-road enfoque más amplio de seis fragmentos, lo que te deja con fragmentos de 12 GiB en la actualidad y fragmentos de 48 GiB en el futuro. De nuevo, es posible que prefiera empezar con tres particiones y reindexar sus datos cuando las particiones superen los 50 GiB.

Un problema mucho menos común implica limitar el número de particiones por nodo. Si el tamaño de las particiones es adecuado, normalmente se queda sin espacio en disco mucho antes de alcanzar este límite. Por ejemplo, una instancia de `m6g.large.search` tiene un tamaño máximo de disco de 512 GiB. Si permanece por debajo del 80 % de uso del disco y el tamaño de sus particiones en 20 GiB, puede acomodar aproximadamente 20 particiones. Elasticsearch 7. x y versiones posteriores, y todas las versiones OpenSearch hasta la 2.15, tienen un límite de 1000 fragmentos por nodo. Para ajustar el máximo de particiones por nodo, establezca la configuración `cluster.max_shards_per_node`. Para la OpenSearch versión 2.17 y versiones posteriores, OpenSearch Service admite 1000 particiones por cada 16 GB de memoria dinámica de JVM, hasta un máximo de 4000 particiones por nodo. Para ver un ejemplo, consulte Configuración del clúster. Para obtener más información sobre la cantidad de particiones, consulte the section called “Cuotas de recuento de la partición”.

El dimensionamiento apropiado de las particiones casi siempre lo mantiene por debajo de este límite, pero también puede considerar el número de particiones por cada GiB del montón de Java. En un

nodo dado, no tenga más de 25 particiones por cada GiB del montón de Java. Por ejemplo, una instancia de `m5.large.search` tiene un montón de 4 GiB, por lo que cada nodo no debe tener más de 100 particiones. En ese recuento de particiones, cada uno tiene aproximadamente 5 GiB de tamaño, lo que está muy por debajo de nuestra recomendación.

Selección de tipos de instancias y pruebas

Después de calcular sus requisitos de almacenamiento y elegir el número de particiones que necesita, puede comenzar a tomar decisiones sobre el equipo. Los requisitos del equipo varían enormemente en función de la carga de trabajo, pero aquí también ofrecemos algunas recomendaciones básicas.

En general, los límites de almacenamiento para cada tipo de instancia se corresponden con la cantidad de CPU y memoria que podría necesitar para las cargas de trabajo. Por ejemplo, una instancia `m6g.large.search` tiene un tamaño de volumen de EBS máximo de 512 GiB, 2 núcleos de vCPU y 8 GiB de memoria. Si el clúster tiene muchas particiones, realiza un gran número de altas, actualiza documentos frecuentemente o procesa un gran número de consultas, esos recursos podrían ser insuficientes para sus necesidades. Si su clúster entra dentro de una de estas categorías, empiece con una configuración más cercana a 2 núcleos de vCPU y 8 GiB de memoria para cada 100 GiB de su requisito de almacenamiento.

> **Tip**
>
> Para ver un resumen de los recursos de hardware que se asignan a cada tipo de instancia, consulta los precios OpenSearch de Amazon Service.

Aun así, estos recursos podrían ser insuficientes. Algunos OpenSearch usuarios afirman que necesitan muchas veces esos recursos para cumplir sus requisitos. Para encontrar el equipo correcto para la carga de trabajo, debe realizar una estimación inicial fundamentada, probarla con cargas de trabajo representativas, ajustarla y probarla de nuevo.

Paso 1: realizar una estimación inicial

Para empezar, recomendamos un mínimo de tres nodos para evitar posibles OpenSearch problemas, como un estado cerebral dividido (cuando un fallo en la comunicación provoca que un clúster tenga dos nodos maestros). Aunque tenga tres nodos maestros dedicados, seguiremos recomendando que tenga un mínimo de dos nodos de datos para la replicación.

Paso 2: calcular los requisitos de almacenamiento por nodo

Si tuviera un requisito de almacenamiento de 184 GiB y el número mínimo recomendado de tres nodos, usaría la ecuación 184 / 3 = 61 GiB para encontrar la cantidad de almacenamiento que necesita cada nodo. En este ejemplo, podría elegir tres instancias `m6g.large.search`, donde cada una con un volumen de almacenamiento de EBS de 90 GiB para no quedarse corto y tener un margen de crecimiento para el futuro. Esta configuración proporciona 6 núcleos de vCPU y 24 GiB de memoria, por lo que resulta adecuada para las cargas de trabajo más ligeras.

Para ver un ejemplo más sustancial, considere un requisito de almacenamiento de 14 TiB (14 336 GiB) y una carga de trabajo pesada. En este caso, podría decidir empezar con 2 * 144 = 288 núcleos de vCPU y 8 * 144 = 1 152 GiB de memoria. Estos números se corresponden con aproximadamente 18 instancias `i3.4xlarge.search`. Si no necesita el almacenamiento local rápido, también puede probar con 18 instancias `r6g.4xlarge.search`, cada una con un volumen de almacenamiento de EBS de 1 TiB.

Si el clúster incluye cientos de terabytes de datos, consulte the section called “Escala de petabytes”.

Paso 3: realizar pruebas representativas

Tras configurar el clúster, puede añadir los índices utilizando el número de fragmentos que calculó anteriormente, realizar algunas pruebas representativas con los clientes utilizando un conjunto de datos realista y supervisar CloudWatch las métricas para ver cómo gestiona el clúster la carga de trabajo.

Paso 4: suceder o iterar

Si el rendimiento satisface sus necesidades, las pruebas se realizan correctamente y CloudWatch las métricas son normales, el clúster está listo para usarse. Recuerde configurar CloudWatch alarmas para detectar un uso deficiente de los recursos.

Si el desempeño no es aceptable, no se superan las pruebas o los valores de `CPUUtilization` o `JVMMemoryPressure` son altos, es posible que tenga que elegir un tipo de instancia diferente (o agregar instancias) y continuar con las pruebas. A medida que agrega instancias, reequilibra OpenSearch automáticamente la distribución de los fragmentos en todo el clúster.

Debido a que es más fácil medir el exceso de capacidad en un clúster sobrealimentado que el déficit en uno infraalimentado, recomendamos comenzar con un clúster más grande de lo que crea necesario. A continuación, debe realizar pruebas y reducir verticalmente el tamaño hasta tener un

clúster eficiente con los recursos adicionales precisos para garantizar operaciones estables durante los períodos de mayor actividad.

Los clústeres en producción o los clústeres con estados complejos se benefician de los nodos maestros dedicados, que mejoran el desempeño y la fiabilidad del clúster.

Escala de petabytes en Amazon Service OpenSearch

Los dominios OpenSearch de Amazon Service ofrecen almacenamiento adjunto de hasta 10 PB. Puede configurar un dominio con 1000 tipos de `OR1.16xlarge.search` instancias, cada uno con 36 TB de almacenamiento. Debido a la gran diferencia de escala, las recomendaciones para los dominios de este tamaño difieren de nuestras recomendaciones generales. En esta sección, se explican consideraciones sobre la creación de dominios, los costos, el almacenamiento y el tamaño de las particiones.

Si bien en esta sección se hace referencia con frecuencia a los tipos de `i3.16xlarge.search` instancias, puedes usar varios otros tipos de instancias para alcanzar los 10 PB de almacenamiento total del dominio.

Creación de dominios

Los dominios de este tamaño superan el límite predeterminado de 80 instancias por dominio. Para solicitar un aumento del límite de servicio de hasta 1000 instancias por dominio, abra un caso en el AWS Support Center.

Precios

Antes de crear un dominio de este tamaño, consulta la página de precios de Amazon OpenSearch Service para asegurarte de que los costes asociados se ajustan a tus expectativas. Examine the section called “UltraWarm almacenamiento” para ver si una arquitectura caliente/templada se ajusta a su caso de uso.

Almacenamiento

Los tipos de `i3` instancias están diseñados para proporcionar un almacenamiento express (NVMe) rápido, local y no volátil. Dado que este almacenamiento local suele ofrecer ventajas de rendimiento en comparación con Amazon Elastic Block Store, los volúmenes de EBS no son una opción cuando se seleccionan estos tipos de instancias en OpenSearch Service. Si prefiere el almacenamiento de EBS, utilice otro tipo de instancia, como `r6.12xlarge.search`.

Tamaño y número de fragmentos

Una OpenSearch pauta habitual es no superar los 50 GB por partición. Habida cuenta del número de fragmentos necesarios para acomodar dominios grandes y los recursos disponibles para instancias `i3.16xlarge.search`, le recomendamos un tamaño de fragmento de 100 GB.

Por ejemplo, si tiene 450 TB de datos de origen y solo quiere una réplica, su requisito de almacenamiento mínimo será aproximadamente 450 TB * 2 * 1,1/0.95 = 1.04 PB. Para obtener una explicación de este cálculo, consulte the section called “Cálculo de requisitos de almacenamiento”. Aunque 1,04 PB / 15 TB = 70 instancias, podría seleccionar 90 o más instancias `i3.16xlarge.search` para disponer de una red de seguridad de almacenamiento, hacer frente a los errores de los nodos y tener en cuenta cierta variación en la cantidad de datos a lo largo del tiempo. Cada instancia agrega otros 20 GiB a su requisito de almacenamiento mínimo, pero para discos de este tamaño, esos 20 GiB son casi insignificantes.

Controlar el número de fragmentos es complicado. OpenSearch los usuarios suelen rotar los índices a diario y conservar los datos durante una o dos semanas. En esta situación, puede que le resulte útil distinguir entre fragmentos "activos" e "inactivos". Los fragmentos activos son fragmentos que se están escribiendo o leyendo activamente. Las particiones inactivas pueden dar servicio a algunas solicitudes de lectura, pero están en gran medida inactivos. En general, debe mantener el número de fragmentos activos por debajo de varios miles. Cuando el número de particiones activas se acerca a los 10 000, surgen riesgos importantes de rendimiento y estabilidad.

Para calcular el número de fragmentos principales, utilice esta fórmula: 450 000 GB * 1,1 / 100 GB por fragmento = 4 950 fragmentos. Si se duplica ese número para dar cuenta de las réplicas se obtienen 9 900 particiones, lo que supone un grave problema si todas las particiones están activas. Pero si rota los índices y solo una séptima parte o una catorceava parte de las particiones están activas en un día determinado (1,414 o 707 particiones, respectivamente), el clúster puede funcionar correctamente. Como siempre, el paso más importante para determinar el tamaño y la configuración del dominio es realizar pruebas representativas en el cliente utilizando un conjunto de datos realista.

Nodos coordinadores dedicados en Amazon OpenSearch Service

Los nodos coordinadores dedicados de Amazon OpenSearch Service son nodos especializados que descargan las tareas de coordinación de los nodos de datos. Estas tareas incluyen la gestión de las solicitudes de búsqueda y el alojamiento OpenSearch de los paneles de control. Al separar estas

funciones, los nodos coordinadores dedicados reducen la carga de los nodos de datos, lo que les permite centrarse en las operaciones de almacenamiento, indexación y búsqueda de datos. Esto mejora el rendimiento general del clúster y la utilización de los recursos.

Además, los nodos coordinadores dedicados ayudan a reducir la cantidad de direcciones IP privadas necesarias para las configuraciones de VPC, lo que conduce a una administración de red más eficiente. Esta configuración puede suponer una mejora de hasta un 15 % en el rendimiento de indexación y un 20 % en el rendimiento de las consultas, según las características de la carga de trabajo.

Cuándo utilizar nodos coordinadores dedicados

Los nodos coordinadores dedicados son los más beneficiosos en los siguientes escenarios.

- Clústeres grandes: en entornos con un gran volumen de datos o consultas complejas, delegar las tareas de coordinación a nodos dedicados puede mejorar el rendimiento del clúster.
- Consultas frecuentes: las cargas de trabajo que implican consultas de búsqueda o agregaciones frecuentes, especialmente aquellas con histogramas de fechas complejos o múltiples agregaciones, se benefician de un procesamiento de consultas más rápido.
- Uso intensivo de los cuadros de mando: los OpenSearch cuadros de mando pueden consumir muchos recursos. Al delegar esta responsabilidad en nodos coordinadores dedicados se reduce la carga sobre los nodos de datos.

Arquitectura y comportamiento

En un OpenSearch clúster, los nodos coordinadores dedicados se ocupan de dos responsabilidades clave.

- Administración de solicitudes: estos nodos reciben las solicitudes de búsqueda entrantes y las reenvían a los nodos de datos correspondientes, que almacenan los datos relevantes. A continuación, consolidan los resultados de varios nodos de datos en un único conjunto de resultados globales, que se devuelve al cliente.
- Alojamiento de cuadros de mando: los nodos coordinadores gestionan los OpenSearch cuadros de mando, lo que alivia a los nodos de datos de la carga adicional que supone alojar los OpenSearch cuadros de mando y gestionar el tráfico relacionado.

En los dominios de VPC, a los nodos coordinadores dedicados se les asignan interfaces de red elásticas (ENIs) en lugar de nodos de datos. Esta disposición ayuda a reducir la cantidad de direcciones IP privadas necesarias VPCs, lo que mejora la eficiencia de la red. Por lo general, los nodos coordinadores dedicados representan aproximadamente el 10 % del total de nodos de datos.

Requisitos y limitaciones

Los nodos coordinadores dedicados tienen los siguientes requisitos y limitaciones.

- Los nodos coordinadores dedicados se admiten en todas OpenSearch las versiones y en las versiones 6.8 a 7.10 de Elasticsearch.
- Para habilitar los nodos coordinadores dedicados, el dominio debe tener habilitados los nodos maestros dedicados. Para obtener más información, consulte the section called “Nodos maestros dedicados”.
- El aprovisionamiento de nodos coordinadores dedicados puede dar lugar costos adicionales. Sin embargo, la mayor eficiencia de los recursos y el mejor rendimiento justifican la inversión, especialmente en el caso de clústeres grandes o complejos.

Aprovisionamiento de nodos coordinadores dedicados

Siga estos pasos para aprovisionar nodos coordinadores dedicados en un dominio existente. Asegúrese de que su dominio tenga habilitados los nodos maestros dedicados antes de aprovisionar los nodos coordinadores.

Consola

Para aprovisionar nodos coordinadores dedicados en el Consola de administración de AWS

1. Inicia sesión en la consola OpenSearch de Amazon Service desde https://console.aws.amazon.com/aos/casa.
2. Elija Dominios y, a continuación, seleccione el dominio que desea modificar.
3. En la sección Configuración de clúster, elija Editar.
4. Seleccione Activar nodos coordinadores dedicados.
5. Seleccione el tipo de instancia y el número de nodos coordinadores que aprovisionar.
6. Seleccione Save changes (Guardar cambios). Puede que el dominio tarde varios minutos en actualizarse.

AWS CLI

Para aprovisionar nodos coordinadores dedicados mediante el AWS CLI, utilice el update-domain-configcomando. El siguiente ejemplo aprovisiona tres nodos coordinadores `r6g.large.search` en un dominio.

```
aws opensearch update-domain-config \
  --domain-name my-opensearch-domain \
  --cluster-config
 InstanceCount=3,InstanceType=r6g.large.search,DedicatedCoordinatorCount=3,ZoneAwarenessEnabled
```

Este comando habilita los nodos coordinadores dedicados, establece el tipo de instancia y el recuento de los nodos coordinadores y habilita el reconocimiento de zona para una mayor disponibilidad.

Prácticas recomendadas

Tenga en cuenta las siguientes prácticas recomendadas al utilizar nodos coordinadores dedicados.

- Utilice instancias de uso general para la mayoría de los casos de uso. Proporcionan un enfoque equilibrado entre costo y rendimiento. Las instancias optimizadas para memoria son ideales para cargas de trabajo que requieren recursos de memoria sustanciales, como las que implican agregaciones complejas o búsquedas a gran escala.
- Un buen punto de partida es aprovisionar del 5 % al 10 % como nodos coordinadores dedicados. Por ejemplo, si su dominio tiene 90 nodos de datos de un tipo de instancia concreto, considere la posibilidad de aprovisionar de 5 a 9 nodos coordinadores del mismo tipo de instancia.

  > **Note**
  >
  > La disponibilidad de los tipos de instancias puede variar según la región. Al seleccionar los tipos de instancia para los nodos coordinadores, compruebe que el tipo de instancia elegido esté disponible en la región de destino. Puede comprobar la disponibilidad del tipo de instancia en la consola de OpenSearch servicio al crear o modificar su dominio.

- Para minimizar el riesgo de que se produzca un único punto de fallo, aprovisione al menos dos nodos coordinadores dedicados. Esto garantiza que el clúster permanezca operativo incluso si falla un nodo.
- Si usted utiliza búsqueda entre Regiones, aprovisione nodos coordinadores dedicados en los dominios de destino. Los dominios de origen por lo general no manejan tareas de coordinación.

- Para entornos con un uso intensivo de indexación, considere la posibilidad de utilizar instancias optimizadas para la CPU que coincidan con el tamaño de las instancias de sus nodos de datos para obtener un rendimiento óptimo.
- Para cargas de trabajo con uso intensivo de memoria, utilice un tipo de instancia ligeramente más grande para sus nodos coordinadores dedicados, a fin de ayudar a administrar el aumento en las demandas de memoria.
- Realiza un seguimiento de la CloudWatch métrica de `CoordinatorCPUUtilization` Amazon. Si supera constantemente el 80 %, podría indicar que necesita nodos coordinadores más grandes o adicionales para administrar la carga.
- Ajuste el tamaño de los nodos coordinadores dedicados para que coincidan con los nodos de datos. Por ejemplo, comience con 4 nodos coordinadores grandes de uso general cuando utilice 4 nodos de datos grandes.
- Utilice varias instancias más pequeñas en lugar de menos instancias más grandes para los nodos coordinadores, a menos que sus solicitudes o respuestas individuales necesiten una memoria (pulgadas GBs) extremadamente alta. Por ejemplo, elija 12 instancias 4xl en lugar de 6 instancias generales 8xlarge.

Recomendaciones de nodos por tamaño de clúster

Utilice las siguientes pautas como punto de partida para aprovisionar nodos coordinadores dedicados en función del tamaño del clúster. Ajusta la cantidad y el tipo de nodos en función de las características de la carga de trabajo y las métricas de rendimiento.

Tamaño del clúster	Nodos coordinadores recomendados	Tipo de instancia
Pequeño (hasta 50 nodos)	de 3 a 5 nodos	Fin general
Medio (de 50 a 100 nodos)	de 5 a 9 nodos	Optimizada para memoria
Grande (más de 100 nodos)	10 a 15 nodos	Optimizada para memoria

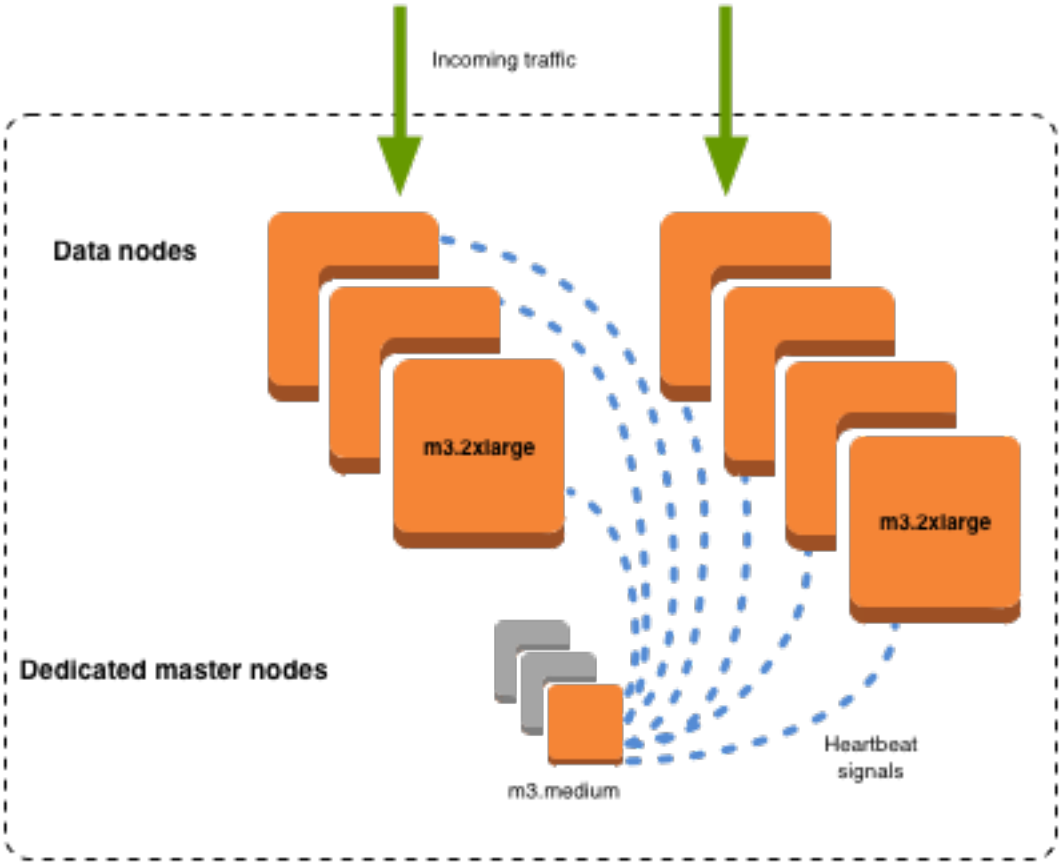
Nodos maestros dedicados en Amazon OpenSearch Service

Amazon OpenSearch Service utiliza nodos maestros dedicados para aumentar la estabilidad del clúster. Un nodo maestro dedicado realiza tareas de administración de clústeres, pero que no contiene datos ni responde a las solicitudes de carga de datos. Al asumir de este modo las tareas de administración del clúster, aumenta la estabilidad del dominio. Al igual que todos los demás tipos de nodos, se paga una tarifa por hora por cada nodo maestro dedicado.

Los nodos maestros dedicados se encargan de realizar las siguientes tareas de administración del clúster:

- Seguimiento de todos los nodos del clúster.
- Seguimiento del número de índices del clúster.
- Seguimiento del número de particiones que pertenecen a cada índice.
- Mantenimiento de la información de enrutamiento para los nodos del clúster.
- Actualización del estado del clúster después de los cambios, como la creación de un índice o la adición o eliminación de nodos en el clúster.
- Reproducción de los cambios de estado del clúster en todos los nodos que lo componen.
- Monitorización del estado de todos los nodos del clúster, mediante el envío de señales de latido, señales periódicas que monitorizan la disponibilidad de los nodos de datos del clúster.

La siguiente ilustración muestra un dominio OpenSearch de servicio con 10 instancias. Siete de ellas son nodos de datos y tres son nodos maestros dedicados. Solo uno de los nodos maestros dedicados está activo. Los dos nodos maestros dedicados grises están en reserva a la espera por si el nodo maestro dedicado activo sufre algún error. Los siete nodos de datos se encargan de atender todas las solicitudes de carga de datos, mientras que el nodo maestro dedicado activo asume las tareas de administración del clúster.



Elección del número de nodos maestros dedicados

Se recomienda utilizar Multi-AZ con modo de espera, que añade tres nodos maestros dedicados a cada dominio de OpenSearch servicio de producción. Si realiza la implementación con Multi-AZ sin modo de espera o con Single-AZ, igual le recomendamos tres nodos maestros dedicados. No elija nunca un número par de nodos maestros dedicados. Tenga en cuenta lo siguiente al elegir el número de nodos maestros dedicados:

- El OpenSearch Servicio prohíbe explícitamente tener un nodo principal dedicado, ya que no se dispone de una copia de seguridad en caso de que se produzca un error. Recibirá una excepción de validación si intenta crear un dominio con un solo nodo maestro dedicado.
- Si tiene dos nodos maestros dedicados, el clúster no tiene el quórum de nodos necesario para elegir un nuevo nodo maestro si se produce algún error.

  El quórum se calcula como el número de nodos maestros dedicados / 2 + 1 (redondeado al número entero más próximo). En este caso, 2 / 2 + 1 = 2. Como un nodo maestro dedicado produjo

un error y solo existe una copia de seguridad, el clúster no tiene quórum y no puede elegir un nuevo nodo principal.

- El uso de tres nodos maestros dedicados, el número recomendado, ofrece dos nodos de backup en caso de que se produzca un error en un nodo maestro y el quórum necesario (2) para elegir un nuevo nodo maestro.
- No es mejor utilizar cuatro nodos maestros dedicados en lugar de tres y se pueden ocasionar problemas si se utilizan múltiples zonas de disponibilidad.
  - Si un nodo maestro produce un error, tiene quórum (3) para elegir un nuevo nodo. Si dos nodos producen un error, pierde ese quórum, de la misma forma que con tres nodos maestros dedicados.
  - En una configuración de tres zonas de disponibilidad, dos AZs tienen un nodo principal dedicado y una zona de disponibilidad tiene dos. Si esa AZ sufre una interrupción, las dos restantes AZs no tienen el quórum necesario (3) para elegir un nuevo maestro.
- Contar con cinco nodos maestros dedicados funciona igual de bien que tres y permite mantener el quórum aunque pierda dos nodos. Sin embargo, dado que solo hay un nodo maestro dedicado activo en cada momento dado, esta configuración supone que usted paga por cuatro nodos inactivos. Muchos clientes consideran excesivo este nivel de protección de conmutación por error.

Si un clúster tiene un número par de nodos aptos para el modo maestro OpenSearch y las versiones 7 de Elasticsearch. x y posteriores ignoran un nodo para que la configuración de votación sea siempre un número impar. En este caso, cuatro nodos maestros dedicados equivalen en esencia a tres (y dos, a uno).

> **Note**
>
> Si el clúster no tiene el quórum necesario para elegir un nuevo nodo maestro, se produce un error en las solicitudes de lectura y de escritura en el clúster. Este comportamiento es diferente del OpenSearch predeterminado.

Elección de tipos de instancias para nodos principales dedicados

OpenSearch Cuotas de instancias y dominios de servicio

Si bien los nodos maestros dedicados no procesan solicitudes de búsqueda ni de consulta, su tamaño es proporcional al tamaño y el número de instancias, índices y particiones que son capaces

de administrar. En el caso de los clústeres de producción, recomendamos como mínimo asignar los siguientes tipos de instancias para los nodos maestros dedicados.

Estas recomendaciones se basan en las cargas de trabajo habituales, pero pueden variar en función de sus necesidades. En los clústeres con muchas particiones o asignaciones de campo, puede resultar conveniente utilizar tipos de instancias más grandes. Para obtener más información, consulta CloudWatch Alarmas recomendadas para Amazon OpenSearch Service para determinar si necesitas usar un tipo de instancia más grande.

RAM	Max Node Support para Elasticsearch y OpenSearch Service de 1.x a 2.15	Max Shard Support para Elasticsearch y OpenSearch Service de la versión 1.x a la 2.15	Max Node Support para el OpenSearch servicio 2.17 y versiones posteriores	Max Shard Support para el OpenSearch servicio 2.17 y versiones posteriores
2 GB	No aplicable	No aplicable	10	1000
4 GB	No aplicable	No aplicable	10	5000
8 GB	10	10 000	30	15 000
16 GB	30	30 000	60	30 000
32 GB	75	40 000	120	60 000
64 GB	125	75 000	240	120 000
128 GB	200	75 000	480	240 000
256 GB	No aplicable	No aplicable	1002	500 000

Referencia general para Amazon OpenSearch Service

Amazon OpenSearch Service admite una variedad de instancias, operaciones, complementos y otros recursos.

Temas

- Tipos de instancias compatibles en Amazon OpenSearch Service
- Características por versión de motor en Amazon OpenSearch Service
- Plugins por versión de motor en Amazon OpenSearch Service
- Operaciones compatibles en Amazon OpenSearch Service
- Cuotas OpenSearch de Amazon Service
- Instancias reservadas en Amazon OpenSearch Service
- Técnicas de optimización de costes para Amazon OpenSearch Service
- Otros recursos compatibles en Amazon OpenSearch Service

Tipos de instancias compatibles en Amazon OpenSearch Service

Amazon OpenSearch Service admite los siguientes tipos de instancias. No todas las regiones admiten todos los tipos de instancias. Para obtener información sobre la disponibilidad, consulta los precios OpenSearch de Amazon Service.

Para obtener información sobre qué tipo de instancia es adecuada para un caso de uso, consulte the section called “Determinación del tamaño de dominios”, the section called “Cuotas de tamaño del volumen de EBS” y the section called “Cuotas de red”.

Ten en cuenta que las siguientes familias de instancias no admiten los controles de cifrado de VPC: T2, T3, R3, I3, C4, M4, R4, C5, M5, R5, C6g, M6g, R6g y R6gd. Los controles de cifrado de VPC son una función de seguridad y cumplimiento que le ofrece un control autoritario centralizado para supervisar el estado de cifrado de sus flujos de tráfico, lo ayuda a identificar los recursos que permiten la comunicación en texto sin cifrar y, finalmente, le proporciona mecanismos para aplicar el cifrado en tránsito dentro de una región y entre ella. VPCs Para obtener más información, consulte la documentación de controles de cifrado de VPC.

Tipos de instancias de generación actual

Para obtener el mejor rendimiento, le recomendamos que utilice los siguientes tipos de instancias al crear nuevos dominios de OpenSearch servicio.

Tipo de instancia	instancias	Restricciones
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	Los tipos de instancias i4i requieren Elasticsearch 5.1 o posterior o cualquier versión de OpenSearch, y no admiten el almacenamiento por volumen de EBS.
i7i	i7i.large.search i7i.xlarge.search	Los tipos de instancia i7i requieren Elasticsearch 5.1 o posterior o cualquier versión de, y no admiten el almacenamiento por volumen de EBS. OpenSearch

Tipo de instancia	instancias	Restricciones
	i7i.2xlarge.search 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	Los tipos de instancias i8g requieren Elasticsearch 7.9 o una versión posterior o cualquier versión de, y no son compatibles con los volúmenes de almacenamiento de OpenSearch EBS.

Tipo de instancia	instancias	Restricciones
Graviton3	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 solo admite el almacenamiento de gp3. Las instancias de Graviton3 (c7g, m7g, r7g) son compatibles con todas las versiones y con las versiones 7.9 y 7.10 de Elasticsearch. OpenSearch

Tipo de instancia	instancias	Restricciones
	m7g.8xlarge.search m7g.12xlarge.search m7g.16xlarge.search r7g.medium.search r7g.large.search r7g.xlarge.search r7g.2xlarge.search r7g.4xlarge.search r7g.8xlarge.search r7g.12xlarge.search r7g.16xlarge.search r7gd.large.search r7gd.xlarge.search	

Tipo de instancia	instancias	Restricciones
	r7gd.2xlarge.search r7gd.4xlarge.search r7gd.8xlarge.search r7gd.12xlarge.search r7gd.16xlarge.search	

Tipo de instancia	instancias	Restricciones

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`
		• Los tipos de OR1 instancia requieren la versión 2.11 o una versión posterior. OpenSearch
		• OR1 las instancias solo son compatibles con los nodos maestros de otros tipos de instancias de Graviton (C6g, M6g, R6g).

Tipo de instancia	instancias	Restricciones

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 y los tipos de OM2 instancia requieren la versión OpenSearch 2.11 o una versión posterior y solo son compatibles con los tipos de instancias de nodos maestros basados en Graviton (C6g, M6g, R6g).

Tipo de instancia	instancias	Restricciones

OM2	`om2.large.search`
	`om2.xlarge.search`
	`om2.2xlarge.search`
	`om2.4xlarge.search`
	`om2.8xlarge.search`
	`om2.12xlarge.search`
	`om2.16xlarge.search`
		• OR2 y los tipos de OM2 instancia requieren la versión OpenSearch 2.11 o una versión posterior y solo son compatibles con los tipos de instancias de nodos maestros basados en Graviton (C6g, M6g, R6g).

Tipo de instancia	instancias	Restricciones

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`
		• Los tipos de OI2 instancia requieren la versión OpenSearch 2.11 o una versión posterior.
		• Estas instancias se aprovisionan con NVMe almacenamiento y no requieren un almacenamiento de EBS independiente.
		• Estas instancias (hasta 8 veces más grandes) también se pueden usar como nodos calientes en la versión 3.3 o versiones posteriores. OpenSearch

Tipo de instancia	instancias	Restricciones

Im4gn	`im4gn.large.search`
	`im4gn.xlarge.search`
	`im4gn.2xlarge.search`
	`im4gn.4xlarge.search`
	`im4gn.8xlarge.search`
	`im4gn.16xlarge.search`
		• Los tipos de instancias iM4gn requieren Elasticsearch 7.9 o una versión posterior o cualquier versión de OpenSearch, y no son compatibles con los volúmenes de almacenamiento de EBS.
		• Las instancias Im4gn solo son compatibles con otros tipos de instancias Graviton (C6g, M6g, R6g, R6gd). No se pueden combinar instancias Graviton y no Graviton en el mismo clúster.

T3	`t3.small.search` `t3.medium.search`	• Los tipos de instancias T3 requieren Elasticsearch 5.6 o posterior o cualquier versión de. OpenSearch • Puede usar los tipos de instancia T3 solo si el dominio está aprovisionado sin modo de espera. Para obtener más informaci ón, consulte the section called “Multi-AZ sin modo de espera”. • Solo puede usar los tipos de instancias T3 cuando el recuento de instancias del dominio sea 10 o menos. • Los tipos de instancias T3 no admiten el UltraWarm almacenam iento, el almacenamiento en frío ni el ajuste automático.

Tipo de instancia	instancias	Restricciones

c7i	`c7i.large.search`
	`c7i.xlarge.search`
	`c7i.2xlarge.search`
	`c7i.4xlarge.search`
	`c7i.8xlarge.search`
	`c7i.12xlarge.search`
	`c7i.16xlarge.search`
		• La instancia c7i requiere Elasticsearch 5.1 o posterior o cualquier versión de OpenSearch, y solo admite volúmenes de almacenamiento. GP3

Tipo de instancia	instancias	Restricciones

m7i	`m7i.large.search`
	`m7i.xlarge.search`
	`m7i.2xlarge.search`
	`m7i.4xlarge.search`
	`m7i.8xlarge.search`
	`m7i.12xlarge.search`
	`m7i.16xlarge.search`
		• La instancia m7i requiere Elasticsearch 5.1 o posterior o cualquier versión de, y solo admite volúmenes de almacenamiento. OpenSearch GP3

Tipo de instancia	instancias	Restricciones

r7i	`r7i.large.search`
	`r7i.xlarge.search`
	`r7i.2xlarge.search`
	`r7i.4xlarge.search`
	`r7i.8xlarge.search`
	`r7i.12xlarge.search`
	`r7i.16xlarge.search`
		• La instancia r7i requiere Elasticsearch 5.1 o posterior o cualquier versión de, y solo admite volúmenes de almacenamiento. OpenSearch GP3

Tipo de instancia	instancias	Restricciones

c8g	`c8g.large.search`
	`c8g.xlarge.search`
	`c8g.2xlarge.search`
	`c8g.4xlarge.search`
	`c8g.8xlarge.search`
	`c8g.12xlarge.search`
	`c8g.16xlarge.search`
		• Los tipos de instancia c8g requieren Elasticsearch 7.9 o posterior o cualquier versión de, y solo admiten volúmenes de almacenamiento. OpenSearch GP3

Tipo de instancia	instancias	Restricciones

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`
		• Los tipos de instancias m8g requieren Elasticsearch 7.9 o posterior o cualquier versión de, y solo admiten volúmenes de almacenamiento. OpenSearch GP3

Tipo de instancia	instancias	Restricciones

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`
		• Los tipos de instancia r8g requieren Elasticsearch 7.9 o posterior o cualquier versión de, y solo admiten volúmenes de almacenamiento. OpenSearch GP3

Tipo de instancia	instancias	Restricciones

r8gd	`r8gd.medium.search`
	`r8gd.large.search`
	`r8gd.2xlarge.search`
	`r8gd.4xlarge.search`
	`r8gd.8xlarge.search`
	`r8gd.12xlarge.search`
	`r8gd.16xlarge.search`
		• Los tipos de instancia r8gd requieren Elasticsearch 7.9 o una versión posterior o cualquier versión de los volúmenes de almacenamiento de EBS, y no son compatibles. OpenSearch

Tipos de instancias de generación anterior

OpenSearch El servicio ofrece tipos de instancias de generaciones anteriores para los usuarios que han optimizado sus aplicaciones en función de ellos y que aún no han realizado la actualización. Le recomendamos que utilice los tipos de instancia de la generación actual para obtener un rendimiento óptimo, pero seguimos admitiendo los siguientes tipos de instancia de la generación anterior.

Tipo de instancia	instancias	Restricciones

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`
		• Los tipos de instancia M3 no admiten el cifrado de datos en reposo, el control de acceso detallado ni la búsqueda en clústeres.
		• Los tipos de instancias M3 tienen restricciones adicionales según OpenSearch la versión. Para obtener más información, consulte the section called “Tipo de instancia M3 no válido”.

M4	`m4.large.search`	

Tipo de instancia	instancias	Restricciones

	`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`
		Los tipos de instancia R3 no admiten el cifrado de datos en reposo ni el control de acceso detallado.

Tipo de instancia	instancias	Restricciones

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.search i4g.8xlarge.search i4g.16xlarge.search	Los tipos de instancias i4g requieren Elasticsearch 7.9 o posterior o cualquier versión de OpenSearch, y no admiten volúmenes de almacenamiento de EBS.

Tipo de instancia	instancias	Restricciones

C5	`c5.large.search`
	`c5.xlarge.search`
	`c5.2xlarge.search`
	`c5.4xlarge.search`
	`c5.9xlarge.search`
	`c5.18xlarge.search`
		Los tipos de instancias C5 requieren Elasticsearch 5.1 o posterior o cualquier versión de. OpenSearch

Tipo de instancia	instancias	Restricciones

C6g	`c6g.large.search`
	`c6g.xlarge.search`
	`c6g.2xlarge.search`
	`c6g.4xlarge.search`
	`c6g.8xlarge.search`
	`c6g.12xlarge.search`
		• Los tipos de instancia C6g requieren Elasticsearch 7.9 o posterior o cualquier versión de. OpenSearch
		• Las instancias C6g solo son compatibles con otros tipos de instancias Graviton (Im4gn, M6g, R6g, R6gd). No se pueden combinar instancias Graviton y no Graviton en el mismo clúster.

Tipo de instancia	instancias	Restricciones

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`
		Los tipos de instancias M5 requieren Elasticsearch 5.1 o posterior o cualquier versión de. OpenSearch

Tipo de instancia	instancias	Restricciones

M6g	`m6g.large.search`
	`m6g.xlarge.search`
	`m6g.2xlarge.search`
	`m6g.4xlarge.search`
	`m6g.8xlarge.search`
	`m6g.12xlarge.search`
		• Los tipos de instancias M6g requieren Elasticsearch 7.9 o posterior o cualquier versión de. OpenSearch
		• Las instancias M6g solo son compatibles con otros tipos de instancias Graviton (Im4gn, C6g, R6g, R6gd). No se pueden combinar instancias Graviton y no Graviton en el mismo clúster.

Tipo de instancia	instancias	Restricciones

R5	`r5.large.search`
	`r5.xlarge.search`
	`r5.2xlarge.search`
	`r5.4xlarge.search`
	`r5.12xlarge.search`
		Los tipos de instancia R5 requieren Elasticsearch 5.1 o posterior o cualquier versión de. OpenSearch

Tipo de instancia	instancias	Restricciones

R6g	`r6g.large.search`
	`r6g.xlarge.search`
	`r6g.2xlarge.search`
	`r6g.4xlarge.search`
	`r6g.8xlarge.search`
	`r6g.12xlarge.search`
		• Los tipos de instancias R6g requieren Elasticsearch 7.9 o posterior o cualquier versión de. OpenSearch
		• Las instancias R6g solo son compatibles con otros tipos de instancias Graviton (Im4gn, C6g, M6g, R6gd). No se pueden combinar instancias Graviton y no Graviton en el mismo clúster.

Tipo de instancia	instancias	Restricciones

R6gd	`r6gd.large.search`
	`r6gd.xlarge.search`
	`r6gd.2xlarge.search`
	`r6gd.4xlarge.search`
	`r6gd.8xlarge.search`
	`r6gd.12xlarge.search`
	`r6gd.16xlarge.search`
		• Los tipos de instancias R6gd requieren Elasticsearch 7.9 o posterior o cualquier versión de, y no son compatibles con los volúmenes de almacenamiento de OpenSearch EBS.
		• Las instancias R6gd solo son compatibles con otros tipos de instancias Graviton (Im4gn, C6g, M6g, R6g). No se pueden combinar instancias Graviton y no Graviton en el mismo clúster.

T2	`t2.micro.search`
	`t2.small.search`
	`t2.medium.search`
		• Solo puede usar los tipos de instancias T2 cuando el recuento de instancias del dominio sea 10 o menos.
		• El tipo de instancia `t2.micro.search` solo admite Elasticsearch 1.5 y 2.3.
		• Los tipos de instancias T2 no admiten el cifrado de datos en reposo, el control de acceso detallado, el almacenamiento, el UltraWarm almacenamiento en frío, la búsqueda entre clústeres ni el ajuste automático.

> **Tip**
>
> Recomendamos utilizar tipos de instancias diferentes para nodos maestros dedicados y nodos de datos.

Características por versión de motor en Amazon OpenSearch Service

Muchas funciones OpenSearch del servicio requieren una OpenSearch versión mínima o una versión antigua de Elasticsearch OSS. Si cumple la versión mínima de una característica, pero la característica no está disponible en el dominio, actualice el software de servicio.

Característica	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
Arquitectura de varios niveles	3.3	No incluido
Origen derivado	3.1	No incluido
Índice Star-tree	2.17	No incluido
Compatibilidad con Amazon Q	2.17	No incluido
Plugins personalizados	2.15	No incluido
Búsqueda de	2.13	No incluido

Característica	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
segmentos simultánea		
Generación de consultas en lenguaje natural	2.13	No incluido
Búsqueda semántica multimodal	2.11	No incluido
Orígenes de datos de consulta directa	2.11	No incluido
Nuevos conectores de machine learning	2.9	No incluido
Búsqueda semántica	2.9	No incluido
Búsqueda puntual	2,5	No incluido
Notificaciones	2.3	No incluido
ML Commons	1.3	No incluido

Característica	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
Replicación entre clústeres	1.1	7,10
Transformaciones de índices	1.0	No incluido
Nodo coordinador dedicado	1.0	6.8
Compatibilidad con VPC	1.0	1.0
Requerir HTTPS para todo el tráfico al dominio		
Compatibilidad con Multi-AZ		
Nodos maestros dedicados		
Paquetes personalizados		

Característica	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
Puntos de conexión personalizados		
Publicación de registros lentos		
Registro de errores de publicación	1.0	5.1
Cifrado de datos en reposo		
Autenticación de Cognito para paneles OpenSearch		
Actualizaciones locales		
Instantáneas automatizadas por hora	1.0	5.3

Característica	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
Node-to-node cifrado	1.0	6.0
Compatibilidad con cliente REST de alto nivel de Java		
Compresión de solicitud y respuesta HTTP		
Alertas	1.0	6.2
SQL	1.0	6.5
Búsqueda en clústeres	1.0	6.7
Control de acceso detallado		
Autenticación SAML para paneles OpenSearch		
Ajuste automático		

Característica	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
Reindexación remota		
UltraWarm	1.0	6.8
Administración de estados de índice		
k-NN por distancia euclidiana	1.0	7.1
Detección de anomalías	1.0	7.4
k-NN por similitud coseno	1.0	7.7
Aprender a clasificar		
Lenguaje de procesamiento de canalizaciones	1.0	7.9
OpenSearch Informes de cuadros de mando		

Característica	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
OpenSearch Paneles de mando: Trace Analytics		
Instancias de Graviton basadas en ARM		
Almacenamiento en frío		
Distancia de Hamming, distancia de norma L1 y scripting Painless para k-NN	1.0	7.10
Búsqueda asíncrona		
Compatibilidad con Curator	No incluido	5.1

Para obtener información acerca de los complementos, que habilitan algunas de estas características y funcionalidades adicionales, consulte the section called “Complementos por versión de motor”. Para obtener información sobre la API de OpenSearch para cada versión, consulte the section called “Operaciones admitidas”.

Plugins por versión de motor en Amazon OpenSearch Service

Los dominios OpenSearch de Amazon Service vienen preempaquetados con complementos de la OpenSearch comunidad. El servicio implementa y administra los complementos automáticamente, pero implementa complementos diferentes según la versión del sistema operativo Elasticsearch OpenSearch heredado que elija para su dominio.

En la siguiente tabla se enumeran los complementos por OpenSearch versión, así como las versiones compatibles del sistema operativo Elasticsearch heredado. Solo incluye los complementos con los que puede interactuar; no es exhaustiva. OpenSearch El servicio utiliza complementos adicionales para habilitar las funciones principales del servicio, como el complemento S3 Repository para las instantáneas y el complemento OpenSearchPerformance Analyzer para la optimización y la supervisión. Para obtener una lista completa de todos los complementos que se ejecutan en un dominio, realice la siguiente solicitud:

```
GET _cat/plugins?v
```

Complemento	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
HanLP	2.11	No compatible
Análisis de hebreo	2.11	No compatible
Search Ranking de Amazon Personalize	2.9	No compatible
Búsqueda neuronal	2.9	No compatible
Security Analytics	2,5	No compatible

Complemento	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
OpenSearch notifications	2.3	No compatible
ML Commons	1.3	No compatible
Análisis de Sudachi (recomendado para japonés)	1.3	No compatible
STConvert	1.3	No compatible
Análisis de Pinyin	1.3	No compatible
Análisis de Nori	1.3	No compatible
OpenSearch observabilidad	1.2	No compatible
OpenSearch replicación entre clústeres	1.1	7.10
OpenSearch búsqueda asíncrona	1.0	7.10
IK (Chinese) Analysis	1.0	7.7

Complemento	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
Vietnamese Analysis		
Thai Analysis		
Learning to Rank		
OpenSearch detección de anomalías	1.0	7.4
OpenSearch k-NN	1.0	7.1
OpenSearch Index State Management	1.0	6.8
OpenSearch seguridad	1.0	6.7
OpenSearch SQL	1.0	6.5
OpenSearch alertando	1.0	6.2
Ukrainian Analysis	1.0	5.3

Complemento	Versión mínima requerida OpenSearch	Versión mínima requerida de Elasticsearch
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
Smart Chinese Analysis	1.0	5.1
Seunjeon Korean Analysis	1.0	5.1
Phonetic Analysis	1.0	2.3
Japanese (kuromoji) Analysis	1.0	Incluido en todos los dominios
ICU Analysis	1.0	Incluido en todos los dominios

Complementos opcionales

Además de los complementos predeterminados que vienen preinstalados, Amazon OpenSearch Service admite varios complementos opcionales de análisis de idioma. Puedes usar Consola de administración de AWS y AWS CLI para asociar un complemento a un dominio, desasociar un complemento de un dominio y enumerar todos los complementos. Un paquete de complementos opcional es compatible con una OpenSearch versión específica y solo se puede asociar a los dominios con esa versión.

Tenga en cuenta que para el complemento Sudachi, al volver a asociar un archivo de diccionario, no se refleja inmediatamente en el dominio. El diccionario se actualiza cuando se ejecuta la siguiente blue/green implementación en el dominio como parte de un cambio de configuración u otra actualización. Como alternativa, puede crear un nuevo paquete con los datos actualizados, crear un nuevo índice con un nuevo paquete, volver a indexar el índice existente en el nuevo índice y, a continuación, eliminar el índice anterior. Si prefiere utilizar el enfoque de reindexación, utilice un alias de índice para que no se interrumpa el tráfico.

Los complementos opcionales utilizan el tipo de paquete `ZIP-PLUGIN`. Para obtener más información sobre los complementos opcionales, consulte the section called “Paquetes”.

Operaciones compatibles en Amazon OpenSearch Service

OpenSearch El servicio es compatible con muchas versiones OpenSearch y versiones antiguas del sistema operativo Elasticsearch. En las siguientes secciones se muestran las operaciones que admite OpenSearch Service para cada versión.

Temas

- Diferencias de API destacadas

Diferencias de API destacadas

Lista nueva APIs

Para admitir clústeres grandes con una gran cantidad de índices y fragmentos, hemos introducido una nueva lista APIs con soporte de paginación, como _. list/indices and _list/shards La API List recupera las estadísticas sobre los índices y los particiones en un formato paginado. Esto agiliza la tarea de procesar las respuestas que incluyen muchos índices.

- `_list/indices`: _list/indices
- `_list/shards`: _list/shards

Cambios en los existentes APIs

Para admitir clústeres de gran tamaño, hemos incorporado a la API `_cluster/stats` la posibilidad de agregar filtros de métricas adicionales que permitan recuperar solo las respuestas estadísticas relevantes, como `_cluster/stats/<metric>/nodes/<node-filters>` y `_cluster/stats/<metric>/<index_metric>/nodes/<node-filters>`. Para obtener más información, consulte _cluster/stats.

Hemos agregado soporte en la API `_cat/shards` para la cancelación de tareas con la especificación de un parámetro de solicitud `cancel_after_time_interval`. Para obtener más información, consulte _cat/shards.

Limitación del tamaño de la respuesta para la API _cat

Para admitir clústeres grandes con un número total de instancias superior a 200 en nodos de datos y calientes, tenemos un límite de 10 000 en la cantidad de índices devueltos por la `_cat/segments API`. Si el número de índices de la respuesta supera este límite, la API devuelve un error 429. Para evitarlo, puede especificar un filtro de patrón de índice en la consulta, por ejemplo `_cat/segments/<index-pattern>`.

Configuración y estadísticas

OpenSearch El servicio solo acepta solicitudes PUT a la `_cluster/settings` API que utilizan el formulario de configuración «plano». Rechaza las solicitudes que utilizan el formato de configuración expandido.

```
// Accepted
PUT _cluster/settings
{
  "persistent" : {
    "action.auto_create_index" : false
  }
}

// Rejected
PUT _cluster/settings
```

```
{
  "persistent": {
    "action": {
      "auto_create_index": false
    }
  }
}
```

El cliente REST Java de alto nivel utiliza el formato expandido, por lo que, si necesita enviar solicitudes de configuración, use el cliente de bajo nivel.

Antes de Elasticsearch 5.3, la `_cluster/settings` API de los dominios de OpenSearch servicio solo admitía el PUT método HTTP, no el GET método. OpenSearch y las versiones posteriores de Elasticsearch admiten GET este método, como se muestra en el siguiente ejemplo:

```
GET https://domain-name.region.es.amazonaws.com/_cluster/settings?pretty
```

A continuación se muestra un ejemplo de rentabilidad:

```
{
  "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"
      }
    }
```

```
  }
}
```

Si comparas las respuestas de un OpenSearch clúster de código abierto y de un OpenSearch servicio para determinadas configuraciones y estadísticas APIs, es posible que veas que faltan campos. OpenSearch El servicio redacta cierta información que expone aspectos internos del servicio, como la ruta de datos del sistema de archivos `_nodes/stats` o el nombre y la versión del sistema operativo. `_nodes`

Reducir

La API `_shrink` puede hacer que las actualizaciones, los cambios de configuración y las eliminaciones de dominio no se lleven a cabo correctamente. No recomendamos utilizarla en dominios que ejecuten las versiones 5.3 o 5.1 de Elasticsearch. Estas versiones tienen un error que puede causar un error en la restauración de instantáneas de índices encogidos.

Si utilizas la `_shrink` API en otras OpenSearch versiones o en otras versiones de Elasticsearch, realiza la siguiente solicitud antes de iniciar la operación de reducción:

```
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
  }
}
```

A continuación, realice la siguiente solicitud después de completar la operación de reducción:

```
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
  }
}
```

Lista nueva APIs

Para admitir clústeres grandes con una gran cantidad de índices y fragmentos, hemos introducido una nueva lista APIs con soporte de paginación, es decir, y. `_list/indices` `_list/shards` La API List recupera las estadísticas sobre los índices y los particiones en un formato paginado. Esto agiliza la tarea de procesar las respuestas que incluyen muchos índices. Para obtener más información acerca de `_list/indices`, consulte List indices. Para obtener más información acerca de `_list/shards`, consulte List shards.

Cambios en los existentes APIs

Para admitir clústeres grandes, hemos agregado soporte en `_cluster/stats/<metric>/nodes/<node-filters>` y`_cluster/stats/<metric>/<index_metric>/nodes/<node-filters>`. Para obtener más información sobre `_cluster/stats`, consulte Estadísticas de clústeres.

Limitar el tamaño de la respuesta para _cat APIs

Para admitir clústeres grandes con un número total de instancias superior a 200 en nodos de datos y calientes, tenemos un límite de 10 000 en la cantidad de índices que devuelve la API _cat/segments. Si el número de índices de la respuesta supera este límite, la API devuelve un error 429. Para evitarlo, puede especificar un filtro de patrón de índice en la consulta (por ejemplo, `_cat/segments/<index-pattern>`).

Además, la API `_cat/shards` ahora admite la cancelación de tareas especificando el parámetro de solicitud `cancel_after_time_interval`. Para obtener más información acerca de las particiones, consulte CAT shards.

Elección de los tipos de instancias para nodos maestro dedicados

En la siguiente tabla, se proporcionan recomendaciones de elección según los tipos de instancia adecuados para los nodos maestros dedicados:

RAM	Número máximo de nodos compatibles	Cantidad máxima de particiones admitidas	
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 versión 2.19

Para obtener información sobre las operaciones de la versión OpenSearch 2.19, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico. Para obtener más información acerca de los cambios de esta versión, consulte las notas de la versión 2.19.

OpenSearch versión 2.17

Para la OpenSearch versión 2.17, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

> **Note**
>
> A partir de la OpenSearch versión 2.17, la `cluster.max_shards_per_node` configuración no se puede modificar. Para la OpenSearch versión 2.17 y versiones posteriores,

OpenSearch Service admite 1000 fragmentos por cada 16 GB de memoria dinámica de JVM, hasta un máximo de 4000 fragmentos por nodo.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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`
- `/_list`
- `/_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`
- `/_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`

  - `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/_sm`
- `/_plugins/_sql`
- `/_percolate`
- `/_rank_eval`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, utilice el cuerpo de la solicitud, no la cadena de consulta, para pasar los valores a Service. `scroll_id` OpenSearch
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch

operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM y otras.

5. Consulte the section called “Reducir”.

> **Note**
>
> Actualmente, el cambio de la funcionalidad del ajuste `cluster.max_shards_per_node` no está habilitado para los clientes con Multi-AZ (zona de disponibilidad) en espera.

OpenSearch versión 2.15

Para la OpenSearch versión 2.15, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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/pipeline`
- `/_search/point_in_time`
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`

  - 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
  - cluster.search.request.slowlog.threshold.trace
  - search.phase_took_enabled
- /_cluster/state
- /_cluster/stats
- /_count
- /_dashboards
- /_plugins/_ism
- /_plugins/_ml
- /_plugins/_notifications
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_security_analytics
- /_plugins/_sm
- /_plugins/_sql
- /_percolate
- /_rank_eval
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 2.13

Para la OpenSearch versión 2.13, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search/pipeline`
- `/_search/point_in_time`
- `/_search profile`

- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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
  - cluster.search.request.slowlog.threshold.trace
- /_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
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

  - `search.phase_took_enabled`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_dashboards`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 2.11

Para la OpenSearch versión 2.11, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`

- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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
- /_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
- /_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. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 2.9

Para la OpenSearch versión 2.9, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search/pipeline`
- `/_search/point_in_time`
- `/_search profile`

- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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
- /_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
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los scroll_id valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar scroll_id los valores a OpenSearch Service.

3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método PUT. Para obtener información acerca del método GET, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 2.7

En la OpenSearch versión 2.7, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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`
- `/_plugins/_ism`
- `/_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`

- 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/_ml
- /_plugins/_notifications
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_security_analytics
- /_plugins/_sm
- /_plugins/_sql
- /_percolate
- /_rank_eval
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 2.5

Para la OpenSearch versión 2.5, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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/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. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 2.3

Para la OpenSearch versión 2.3, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-nam*
- `/_delete_by_query` [1]
- `/_explain`
- `/_refresh`
- `/_reindex` [1]

- *e* /_forcemerge , /*index-name* /update/*id* y /*index-name* /_close)
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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_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
- /_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

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 1.3

Para la OpenSearch versión 1.3, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`

- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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
- /_nodes
- /_plugins/_asynchronous_search
- /_plugins/_alerting
- /_plugins/_anomaly_detection
- /_plugins/_ism
- /_plugins/_ml
- /_plugins/_ppl
- /_plugins/_security
- /_plugins/_sql
- /_percolate
- /_rank_eval
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los scroll_id valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar scroll_id los valores a OpenSearch Service.

3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método PUT. Para obtener información acerca del método GET, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 1.2

Para la OpenSearch versión 1.2, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* /_forcemerge , */index-name* /update/*id* y */index-name* /_close)
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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
- /_plugins/_ism
- /_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

    - 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/_ppl
  - /_plugins/_security
  - /_plugins/_sql
  - /_percolate
  - /_rank_eval

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 1.1

Para la OpenSearch versión 1.1, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

OpenSearch versión 1.0

Para la OpenSearch versión 1.0, el OpenSearch servicio admite las siguientes operaciones. Para obtener información sobre la mayoría de las operaciones, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.

- Todas las operaciones de la ruta de índice (como */index-nam*
- `/_delete_by_query` [1]
- `/_explain`
- `/_refresh`
- `/_reindex` [1]

e /_forcemerge , /*index-name* /update/*id* y /*index-name* /_close)
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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_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
- /_rank_eval
- /_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

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres de los `scroll_id` valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 7.10

Para Elasticsearch 7.10, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_index_template` [6]
- `/_ingest/pipeline`
- `/_index_template`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_resolve/index`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`

- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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
- /_nodes
- /_opendistro/_alerting
- /_opendistro/_asynchronous_search
- /_opendistro/_anomaly_detection
- /_opendistro/_ism
- /_opendistro/_ppl
- /_opendistro/_security
- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_plugins/_replication
- /_rank_eval
- /_status
- /_tasks
- /_template [6]
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en scroll_id los valores, utilice el cuerpo de la solicitud, no la cadena de consulta, para pasar scroll_id los valores a Service. OpenSearch
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método PUT. Para obtener información acerca del método GET, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones

genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.

5. Consulte the section called “Reducir”.
6. Las plantillas de índice heredadas (`_template`) fueron reemplazadas por plantillas que se pueden combinar (`_index_template`), a partir de Elasticsearch 7.8. Las plantillas que se pueden combinar tienen prioridad sobre las plantillas heredadas. Si ninguna plantilla que se puede combinar coincide con un índice determinado, una plantilla heredada puede coincidir y ser aplicada. La `_template` operación sigue funcionando en las versiones posteriores de Elasticsearch OSS OpenSearch y en las versiones posteriores, pero las llamadas GET a los dos tipos de plantillas arrojan resultados diferentes.

Elasticsearch versión 7.9

Para Elasticsearch 7.9, OpenSearch Service admite las siguientes operaciones.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [4]:
- `/_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`
- `/_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]

- 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/_ppl
- /_opendistro/_security
- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_rank_eval
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las OpenSearch operaciones genéricas que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.
6. Las plantillas de índice heredadas (`_template`) fueron reemplazadas por plantillas que se pueden combinar (`_index_template`), a partir de Elasticsearch 7.8. Las plantillas que se

pueden combinar tienen prioridad sobre las plantillas heredadas. Si ninguna plantilla que se puede combinar coincide con un índice determinado, una plantilla heredada puede coincidir y ser aplicada. La `_template` operación sigue funcionando en las versiones posteriores de Elasticsearch OSS OpenSearch y en las versiones posteriores, pero las llamadas GET a los dos tipos de plantillas arrojan resultados diferentes.

Elasticsearch versión 7.8

Para Elasticsearch 7.8, OpenSearch Service admite las siguientes operaciones.

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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`
- `/_index_template` [6]
- `/_ingest/pipeline`
- `/_ltr`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_opendistro/_anomaly_detection`
- `/_opendistro/_ism`
- `/_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`

- `indices.breaker.request.limit`
- `indices.breaker.total.limit`
- `cluster.max_shards_per_node`
- `/_opendistro/_security`
- `/_opendistro/_sql`
- `/_percolate`
- `/_plugin/kibana`
- `/_rank_eval`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.
6. Las plantillas de índice heredadas (`_template`) fueron reemplazadas por plantillas que se pueden combinar (`_index_template`), a partir de Elasticsearch 7.8. Las plantillas que se pueden combinar tienen prioridad sobre las plantillas heredadas. Si ninguna plantilla que se puede combinar coincide con un índice determinado, una plantilla heredada puede coincidir y ser aplicada. La `_template` operación sigue funcionando en las versiones posteriores de Elasticsearch OSS OpenSearch y en las versiones posteriores, pero las llamadas GET a los dos tipos de plantillas arrojan resultados diferentes.

Elasticsearch versión 7.7

En el caso de Elasticsearch 7.7, OpenSearch Service admite las siguientes operaciones.

- Todas las operaciones de la ruta de índice (como */index-name* /_forcemerge , */index-name* /update/*id* y */index-name* /_close)
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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
- /_delete_by_query [1]
- /_explain
- /_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
- /_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

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 7.4

En el caso de Elasticsearch 7.4, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones de la ruta de índice (como */index-name* `/_forcemerge` , */index-name* `/update/`*id* y */index-name* `/_close`)
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`
- `/_split`
- `/_stats`
- `/_status`

- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_opendistro/_anomaly_detection`
- `/_opendistro/_ism`
- `/_opendistro/_security`
- `/_opendistro/_sql`
- `/_percolate`
- `/_plugin/kibana`
- `/_rank_eval`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 7.1

En el caso de Elasticsearch 7.1, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-name* `/update/`*id*) excepto */index-name* `/_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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`
- `/_ingest/pipeline`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_mtermvectors`
- `/_nodes`
- `/_opendistro/_alerting`
- `/_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`
- `/_update_by_query` [1]
- `/_validate`

- `cluster.max_shards_per_node`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 6.8

En el caso de Elasticsearch 6.8, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-name* `/update/`*id*) excepto */index-name* `/_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_explain`
- `/_field_caps`
- `/_field_stats`
- `/_flush`
- `/_ingest/pipeline`
- `/_mapping`
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`
- `/_scripts` [3]
- `/_search`[2]
- `/_search profile`
- `/_shard_stores`
- `/_shrink`[5]
- `/_snapshot`

- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [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.blocks.read_only
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_opendistro/_alerting
- /_opendistro/_ism
- /_opendistro/_security
- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_rank_eval
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en scroll_id los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar scroll_id los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.

4. Se refiere al método PUT. Para obtener información acerca del método GET, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 6.7

En el caso de Elasticsearch 6.7, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-name* `/update/`*id*) excepto */index-name* `/_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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/_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`
- `/_update_by_query` [1]
- `/_validate`

- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`
- `cluster.max_shards_per_node`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 6.5

En el caso de Elasticsearch 6.5, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-*
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_refresh`
- `/_reindex` [1]
- `/_render`
- `/_rollover`

name /update/*id*) excepto /*index-name* /_close
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
- /_explain
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_opendistro/_alerting
- /_opendistro/_sql
- /_percolate
- /_plugin/kibana
- /_rank_eval
- /_scripts [3]
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en scroll_id

los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.

3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 6.4

En el caso de Elasticsearch 6.4, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-name* `/update/`*id*) excepto */index-name* `/_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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`
- `/_percolate`
- `/_plugin/kibana`
- `/_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.auto_create_index`
- `action.search.shard_count.limit`
- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`

- `/_rank_eval`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 6.3

En el caso de Elasticsearch 6.3, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-*
- `/_cluster/state`
- `/_cluster/stats`
- `/_refresh`
- `/_reindex` [1]

- *name* /update/*id*) excepto /*index-name* /_close
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
- /_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
- /_render
- /_rollover
- /_scripts [3]
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_split
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en scroll_id

los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.

3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método PUT. Para obtener información acerca del método GET, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 6.2

En el caso de Elasticsearch 6.2, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-name* `/update/`*id*) excepto */index-name* `/_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [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`
- `/_percolate`
- `/_plugin/kibana`
- `/_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.auto_create_index`
- `action.search.shard_count.limit`
- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`

- `/_rank_eval`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 6.0

En el caso de Elasticsearch 6.0, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-*
- `/_cluster/state`
- `/_cluster/stats`
- `/_render`
- `/_rollover`

name /update/*id*) excepto */index-name* /_close
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
- /_count
- /_delete_by_query [1]
- /_explain
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_percolate
- /_plugin/kibana
- /_refresh
- /_reindex [1]
- /_scripts [3]
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en scroll_id

los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.

3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 5.6

En el caso de Elasticsearch 5.6, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-name* `/update/`*id*) excepto */index-name* `/_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades [4]:
- `/_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`

- `action.auto_create_index`
- `action.search.shard_count.limit`
- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 5.5

En el caso de Elasticsearch 5.5, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-*
- `/_cluster/state`
- `/_cluster/stats`
- `/_render`
- `/_rollover`

- *name* /update/*id*) excepto /*index-name* /_close
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades [4]:
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
- /_count
- /_delete_by_query [1]
- /_explain
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_percolate
- /_plugin/kibana
- /_refresh
- /_reindex [1]
- /_scripts [3]
- /_search[2]
- /_search profile
- /_shard_stores
- /_shrink[5]
- /_snapshot
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en scroll_id

los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.

3. Consulte the section called “Otros recursos admitidos” para ver consideraciones sobre la utilización de scripts.
4. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
5. Consulte the section called “Reducir”.

Elasticsearch versión 5.3

En el caso de Elasticsearch 5.3, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-name* `/update/`*id*) excepto */index-name* `/_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/allocation/explain`
- `/_cluster/health`
- `/_cluster/pending_tasks`
- `/_cluster/settings` para varias propiedades[3]:
- `/_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`[4]
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_tasks`
- `/_template`
- `/_update_by_query` [1]
- `/_validate`

- `action.auto_create_index`
- `action.search.shard_count.limit`
- `indices.breaker.fielddata.limit`
- `indices.breaker.request.limit`
- `indices.breaker.total.limit`

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación `/_tasks` combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para `/_search/scroll` con un cuerpo de mensaje deben especificar `"Content-Length"` en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en `scroll_id` los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.
3. Se refiere al método `PUT`. Para obtener información acerca del método `GET`, consulte the section called “Diferencias de API destacadas”. Esta lista solo hace referencia a las operaciones genéricas de Elasticsearch que admite OpenSearch Service y no incluye las operaciones compatibles específicas de los complementos para la detección de anomalías, el ISM, etc.
4. Consulte the section called “Reducir”.

Elasticsearch versión 5.1

En el caso de Elasticsearch 5.1, Service admite las siguientes operaciones. OpenSearch

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-*
- `/_cluster/state`
- `/_cluster/stats`
- `/_count`
- `/_delete_by_query` [1]
- `/_render`
- `/_rollover`
- `/_search` [2]
- `/_search profile`

name /update/*id*) excepto /*index-name* /_close
- /_alias
- /_aliases
- /_all
- /_analyze
- /_bulk
- /_cat (excepto /_cat/nodeattrs)
- /_cluster/allocation/explain
- /_cluster/health
- /_cluster/pending_tasks
- /_cluster/settings para varias propiedades (solo PUT):
  - action.auto_create_index
  - action.search.shard_count.limit
  - indices.breaker.fielddata.limit
  - indices.breaker.request.limit
  - indices.breaker.total.limit
- /_explain
- /_field_caps
- /_field_stats
- /_flush
- /_ingest/pipeline
- /_mapping
- /_mget
- /_msearch
- /_mtermvectors
- /_nodes
- /_percolate
- /_plugin/kibana
- /_refresh
- /_reindex [1]
- /_shard_stores
- /_shrink[3]
- /_snapshot
- /_stats
- /_status
- /_tasks
- /_template
- /_update_by_query [1]
- /_validate

1. Los cambios en la configuración de un clúster pueden interrumpir estas operaciones antes de que finalicen. Recomendamos utilizar la operación /_tasks combinada con estas operaciones para verificar que las solicitudes se hayan llevado a cabo correctamente.
2. Las solicitudes DELETE para /_search/scroll con un cuerpo de mensaje deben especificar "Content-Length" en el encabezado HTTP. La mayoría de los clientes agregan este encabezado de forma predeterminada. Para evitar problemas con los = caracteres en scroll_id

los valores, usa el cuerpo de la solicitud, no la cadena de consulta, para pasar `scroll_id` los valores a OpenSearch Service.

3. Consulte the section called “Reducir”.

Elasticsearch versión 2.3

Para Elasticsearch 2.3, OpenSearch Service admite las siguientes operaciones.

- Todas las operaciones en la ruta de índice (como */index-name* `/_forcemerge` y */index-name* `/_recovery`) excepto */index-name* `/_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cache/clear` (solo índice)
- `/_cat` (excepto `/_cat/nodeattrs`)
- `/_cluster/health`
- `/_cluster/settings` para varias propiedades (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`
  - `threadpool.percolate.queue_size`
  - `threadpool.search.queue_size`
- `/_cluster/stats`
- `/_count`
- `/_flush`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_refresh`
- `/_render`
- `/_search`
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_template`

- `threadpool.suggest.queue_size`

Elasticsearch versión 1.5

Para Elasticsearch 1.5, OpenSearch Service admite las siguientes operaciones.

- Todas las operaciones de la ruta de índice, como `/`*`index-name`*` /_optimize` y `/`*`index-name`*` /_warmer`, excepto `/`*`index-name`*` /_close`
- `/_alias`
- `/_aliases`
- `/_all`
- `/_analyze`
- `/_bulk`
- `/_cat`
- `/_cluster/health`
- `/_cluster/settings` para varias propiedades (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`
  - `threadpool.percolate.queue_size`
  - `threadpool.search.queue_size`
  - `threadpool.suggest.queue_size`
- `/_cluster/stats`
- `/_count`
- `/_flush`
- `/_mapping`
- `/_mget`
- `/_msearch`
- `/_nodes`
- `/_percolate`
- `/_plugin/kibana`
- `/_plugin/kibana3`
- `/_plugin/migration`
- `/_refresh`
- `/_search`
- `/_snapshot`
- `/_stats`
- `/_status`
- `/_template`

Cuotas OpenSearch de Amazon Service

Tu AWS cuenta tiene cuotas predeterminadas, anteriormente denominadas límites, para cada AWS servicio. A menos que se indique lo contrario, cada cuota es específica de la región.

Para ver las cuotas de los dominios e instancias de OpenSearch servicio, Amazon OpenSearch Serverless y Amazon OpenSearch Ingestion, consulte las cuotas de Amazon OpenSearch Service en. Referencia general de AWS

Para ver las cuotas de OpenSearch Service en Consola de administración de AWS, abra la consola Service Quotas. En el panel de navegación, elige AWS servicios y selecciona Amazon OpenSearch Service. Para solicitar un aumento de cuota, consulte Solicitud de un aumento de cuota en la Guía de usuario de Service Quotas.

Cuotas de almacenamiento en nodos cálidos

En la siguiente tabla, se enumeran los tipos de instancias de nodo cálido y la cantidad máxima de almacenamiento que puede utilizar cada tipo. En el caso de OI2 las instancias OpenSearch optimizadas, el 80% del almacenamiento local está disponible como caché y el almacenamiento en caliente direccionable máximo es 5 veces el tamaño del almacenamiento en caché.

Por ejemplo, si una OI2 instancia tiene 468 GB de almacenamiento local:

- Almacenamiento en caché = 375 GB (80% del almacenamiento local)
- Almacenamiento en caliente direccionable máximo = 1875 GB (caché de 5 x 375 GB)

Tipo de instancia	Almacenamiento de la instancia (GB)	Tamaño de la caché (GB)	Capacidad máxima de almacenamiento en caliente direccionable (GB)
oi2.large.search	468	375	1875
oi2.xlarge.search	937	750	3750
oi2.2xlarge.search	1875	1500	7500
oi2.4 x large.search	3750	3 000	15000

Tipo de instancia	Almacenamiento de la instancia (GB)	Tamaño de la caché (GB)	Capacidad máxima de almacenamiento en caliente direccionable (GB)
oi2.8xlarge.search	7500	6000	30000

UltraWarm cuotas de almacenamiento

En la siguiente tabla, se muestran los tipos de UltraWarm instancias y la cantidad máxima de almacenamiento que puede utilizar cada tipo. Para obtener más información al respecto UltraWarm, consultethe section called “UltraWarm almacenamiento”.

Tipo de instancia	Almacenamiento máximo
ultrawarm1.medium.search	1,5 TiB
ultrawarm1.large.search	20 TiB

Número de nodos de datos por AZ

En la siguiente tabla, se indica el número total de nodos de datos para la implementación en una zona de disponibilidad. El límite total indica el número de nodos de datos por límite, incluido el recuento de nodos calientes y calientes, el almacenamiento que puede utilizar cada tipo.

Configuración de AZ	Límite de nodos activos	Límite de nodos cálidos	Límite general (activos + cálido)
1: AZ	334	250	334
2: AZ	668	500	668
3: AZ	1002	750	1002

Límite total de nodos por familia de instancias

En la siguiente tabla, se enumera el límite total de nodos por familia de instancias.

Familia de instancias	ElasticSearch OpenSearch hasta 2.15	OpenSearch 2.17 y superior	Límite predeterminado
T2	10	10	10
T3	10	10	10
M3, C4, M4, R4, C5, M5, R5, I2, I3	10	200	80
Gravitón 2, Gravitón 3, Gravitón 4	200	400	80
C7i, R7i, M7i, i4i, i4g, i8g, i7i	200	400	80
OR1.búsqueda media OR1.large.search OR2.medium.search OR2.large.search OM2.large.search OI2.large.search	200	400	80
OR1.xlarge.search y superior OR2.xlarge.search y versiones posteriores OM2.xlarge.search y versiones posteriores OI2.xlarge.search y versiones posteriores	200	1002	80

Familia de instancias	ElasticSearch OpenSearch hasta 2.15	OpenSearch 2.17 y superior	Límite predeterminado
Ultrawarm1	150	750	150

Cuotas de tamaño del volumen de EBS

La siguiente tabla muestra los tamaños mínimo y máximo de los volúmenes de EBS para cada tipo de instancia que admite el servicio. OpenSearch Para obtener información sobre qué tipos de instancias incluyen almacenamiento de instancias y detalles de hardware adicionales, consulta los precios OpenSearch de Amazon Service.

- Si elige el almacenamiento magnético en Tipo de volumen de EBS al crear un dominio, el tamaño de volumen máximo es de 100 GiB para todos los tipos de instancias, excepto t2.small y t2.medium y todas las instancias de Graviton (M6g, C6g, R6g y R6gd) que no admiten almacenamiento magnético. Para el tamaño máximo que se muestra en la tabla siguiente, elija una de las opciones de SSD.
- Algunos tipos de instancias de generaciones anteriores incluyen almacenamiento de la instancia, pero también admiten el almacenamiento de EBS. Si elige el almacenamiento de EBS para uno de estos tipos de instancia, los volúmenes de almacenamiento no se suman. Puede utilizar el almacenamiento de la instancia o un volumen de EBS, no ambos.
- Para los volúmenes gp3 de EBS, los límites máximos admitidos son de 80 000 IOPS y 2 000 de rendimiento. MB/s Tenga en cuenta que los límites efectivos de su dominio pueden ser más bajos según el tipo de instancia. Para ver los límites máximos específicos de cada instancia, consulta los límites por tipo de instancia.

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
t2.micro.search	10 GiB	35 GiB	N/A
t2.small.search	10 GiB	35 GiB	N/A

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
t2.medium.search	10 GiB	35 GiB	N/A
t3.small.search	10 GiB	100 GiB	100 GiB
t3.medium.search	10 GiB	200 GiB	200 GiB
m3.medium.search	10 GiB	100 GiB	N/A
m3.large.search	10 GiB	512 GiB	N/A
m3.xlarge.search	10 GiB	512 GiB	N/A
m3.2xlarge.search	10 GiB	512 GiB	N/A
m4.large.search	10 GiB	512 GiB	N/A
m4.xlarge.search	10 GiB	1 TiB	N/A
m4.2xlarge.search	10 GiB	1,5 TiB	N/A
m4.4xlarge.search	10 GiB	1,5 TiB	N/A
m4.10xlarge.search	10 GiB	1,5 TiB	N/A
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
m6g.large.search	10 GiB	512 GiB	1 TiB
m6g.xlarge.search	10 GiB	1 TiB	2 TiB

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
m6g.2xlarge.search	10 GiB	1,5 TiB	3 TiB
m6g.4xlarge.search	10 GiB	3 TiB	6 TiB
m6g.8xlarge.search	10 GiB	6 TiB	12 TiB
m6g.12xlarge.search	10 GiB	9 TiB	18 TiB
c4.large.search	10 GiB	100 GiB	N/A
c4.xlarge.search	10 GiB	512 GiB	N/A
c4.2xlarge.search	10 GiB	1 TiB	N/A
c4.4xlarge.search	10 GiB	1,5 TiB	N/A
c4.8xlarge.search	10 GiB	1,5 TiB	N/A
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
c6g.2xlarge.search	10 GiB	1 TiB	1 TiB
c6g.4xlarge.search	10 GiB	1,5 TiB	1,5 TiB

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
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/A
r3.xlarge.search	10 GiB	512 GiB	N/A
r3.2xlarge.search	10 GiB	512 GiB	N/A
r3.4xlarge.search	10 GiB	512 GiB	N/A
r3.8xlarge.search	10 GiB	512 GiB	N/A
r4.large.search	10 GiB	1 TiB	N/A
r4.xlarge.search	10 GiB	1,5 TiB	N/A
r4.2xlarge.search	10 GiB	1,5 TiB	N/A
r4.4xlarge.search	10 GiB	1,5 TiB	N/A
r4.8xlarge.search	10 GiB	1,5 TiB	N/A
r4.16xlarge.search	10 GiB	1,5 TiB	N/A
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
r5.12xlarge.search	10 GiB	12 TiB	24 TiB
r6g.large.search	10 GiB	1 TiB	2 TiB

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
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/A	N/A	N/A
r6gd.xlarge.search	N/A	N/A	N/A
r6gd.2xlarge.search	N/A	N/A	N/A
r6gd.4xlarge.search	N/A	N/A	N/A
r6gd.8xlarge.search	N/A	N/A	N/A
r6gd.12xlarge.search	N/A	N/A	N/A
r6gd.16xlarge.search	N/A	N/A	N/A
i2.xlarge.search	10 GiB	512 GiB	N/A
i2.2xlarge.search	10 GiB	512 GiB	N/A
i3.large.search	N/A	N/A	N/A
i3.xlarge.search	N/A	N/A	N/A
i3.2xlarge.search	N/A	N/A	N/A
i3.4xlarge.search	N/A	N/A	N/A
i3.8xlarge.search	N/A	N/A	N/A

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
i3.16xlarge.search	N/A	N/A	N/A
or1.medium.search	20 GiB	N/A	768 GiB
or1.large.search	20 GiB	N/A	1532 GiB
or1.xlarge.search	20 GiB	N/A	3 TiB
or1.2xlarge.search	20 GiB	N/A	6 TiB
or1.4xlarge.search	20 GiB	N/A	12 TiB
or1.8xlarge.search	20 GiB	N/A	16 TiB
or1.12xlarge.search	20 GiB	N/A	24 TiB
or1.16xlarge.search	20 GiB	N/A	36 TiB
or2.medium.search	20 GiB	N/A	768 GiB
or2.large.search	20 GiB	N/A	1532 GiB
or2.xlarge.search	20 GiB	N/A	3 TiB
or2.2xlarge.search	20 GiB	N/A	6 TiB
or2.4xlarge.search	20 GiB	N/A	12 TiB
or2.8xlarge.search	20 GiB	N/A	16 TiB
or2.12xlarge.search	20 GiB	N/A	24 TiB
or2.16xlarge.search	20 GiB	N/A	36 TiB
om2.large.search	20 GiB	N/A	768 GiB
om2.xlarge.search	20 GiB	N/A	2 TiB

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
om2.2xlarge.search	20 GiB	N/A	3 TiB
om2.4xlarge.search	20 GiB	N/A	6 TiB
om2.8xlarge.search	20 GiB	N/A	12 TiB
om2.12xlarge.search	20 GiB	N/A	18 TiB
om2.16xlarge.search	20 GiB	N/A	24 TiB
im4gn.large.search	N/A	N/A	N/A
im4gn.xlarge.search	N/A	N/A	N/A
im4gn.2xlarge.search	N/A	N/A	N/A
im4gn.4xlarge.search	N/A	N/A	N/A
im4gn.8xlarge.search	N/A	N/A	N/A
im4gn.16xlarge.search	N/A	N/A	N/A
C7g.large.search	10 GiB	N/A	256 GiB
C7g.xlarge.search	10 GiB	N/A	512 GiB
C7g.2xlarge.search	10 GiB	N/A	1 TiB
C7g.4xlarge.search	10 GiB	N/A	1,5 TiB
C7g.8xlarge.search	10 GiB	N/A	3 TiB
C7g.12xlarge.search	10 GiB	N/A	4,5 TiB
C7g.16xlarge.search	10 GiB	N/A	6 TiB
M7g.medium.search	10 GiB	N/A	512 GiB

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
M7g.large.search	10 GiB	N/A	768 GiB
M7g.xlarge.search	10 GiB	N/A	2 TiB
M7g.2xlarge.search	10 GiB	N/A	3 TiB
M7g.4xlarge.search	10 GiB	N/A	6 TiB
M7g.8xlarge.search	10 GiB	N/A	12 TiB
M7g.12xlarge.search	10 GiB	N/A	18 TiB
M7g.16xlarge.search	10 GiB	N/A	24 TiB
R7g.medium.search	10 GiB	N/A	768 GiB
R7g.large.search	10 GiB	N/A	1,5 TiB
R7g.xlarge.search	10 GiB	N/A	3 TiB
R7g.2xlarge.search	10 GiB	N/A	6 TiB
R7g.4xlarge.search	10 GiB	N/A	12 TiB
R7g.8xlarge.search	10 GiB	N/A	16 TiB
R7g.12xlarge.search	10 GiB	N/A	24 TiB
R7g.16xlarge.search	10 GiB	N/A	36 TiB
R7gd.large.search	N/A	N/A	N/A
R7gd.xlarge.search	N/A	N/A	N/A
R7gd.2xlarge.search	N/A	N/A	N/A
R7gd.4xlarge.search	N/A	N/A	N/A

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
R7gd.8xlarge.search	N/A	N/A	N/A
R7gd.12xlarge.search	N/A	N/A	N/A
R7gd.16xlarge.search	N/A	N/A	N/A
i4i.large.search	N/A	N/A	N/A
i4i.xlarge.search	N/A	N/A	N/A
i4i.2xlarge.search	N/A	N/A	N/A
i4i.4xlarge.search	N/A	N/A	N/A
i4i.8xlarge.search	N/A	N/A	N/A
i4i.12xlarge.search	N/A	N/A	N/A
i4i.16xlarge.search	N/A	N/A	N/A
i4i.24xlarge.search	N/A	N/A	N/A
i4i.32xlarge.search	N/A	N/A	N/A
i4g.large.search	N/A	N/A	N/A
i4g.xlarge.search	N/A	N/A	N/A
i4g.2xlarge.search	N/A	N/A	N/A
i4g.4xlarge.search	N/A	N/A	N/A
i4g.8xlarge.search	N/A	N/A	N/A
i4g.16xlarge.search	N/A	N/A	N/A
i8g.large.search	N/A	N/A	N/A

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
i8g.xlarge.search	N/A	N/A	N/A
i8g.2xlarge.search	N/A	N/A	N/A
i8g.4xlarge.search	N/A	N/A	N/A
i8g.8xlarge.search	N/A	N/A	N/A
i8g.12xlarge.search	N/A	N/A	N/A
i8g.16xlarge.search	N/A	N/A	N/A
i7i.large.search	N/A	N/A	N/A
i7i.xlarge.search	N/A	N/A	N/A
i7i.2xlarge.search	N/A	N/A	N/A
i7i.4xlarge.search	N/A	N/A	N/A
i7i.8xlarge.search	N/A	N/A	N/A
i7i.12xlarge.search	N/A	N/A	N/A
i7i.16xlarge.search	N/A	N/A	N/A
c7i.large.search	10 GiB	N/A	256 GiB
c7i.xlarge.search	10 GiB	N/A	512 GiB
c7i.2xlarge.search	10 GiB	N/A	1 TiB
c7i.4xlarge.search	10 GiB	N/A	1,5 TiB
c7i.8xlarge.search	10 GiB	N/A	3 TiB
c7i.12xlarge.search	10 GiB	N/A	4,5 TiB

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
c7i.16xlarge.search	10 GiB	N/A	6 TiB
m7i.large.search	10 GiB	N/A	768 GiB
m7i.xlarge.search	10 GiB	N/A	2 TiB
m7i.2xlarge.search	10 GiB	N/A	3 TiB
m7i.4xlarge.search	10 GiB	N/A	6 TiB
m7i.8xlarge.search	10 GiB	N/A	12 TiB
m7i.12xlarge.search	10 GiB	N/A	18 TiB
m7i.16xlarge.search	10 GiB	N/A	24 TiB
r7i.large.search	10 GiB	N/A	1,5 TiB
r7i.xlarge.search	10 GiB	N/A	3 TiB
r7i.2xlarge.search	10 GiB	N/A	6 TiB
r7i.4xlarge.search	10 GiB	N/A	12 TiB
r7i.8xlarge.search	10 GiB	N/A	16 TiB
r7i.12xlarge.search	10 GiB	N/A	24 TiB
r7i.12xlarge.search	10 GiB	N/A	36 TiB
c8g.large	10 GiB	N/A	256 GiB
c8g.xlarge	10 GiB	N/A	512 GiB
c8g.2xlarge	10 GiB	N/A	1 TiB
c8g.4xlarge	10 GiB	N/A	1,5 TiB

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
c8g.8xlarge	10 GiB	N/A	3 TiB
c8g.12xlarge	10 GiB	N/A	4,5 TiB
c8g.16xlarge	10 GiB	N/A	6 TiB
m8g.medium	10 GiB	N/A	512 GiB
m8g.large	10 GiB	N/A	768 GiB
m8g.xlarge	10 GiB	N/A	2 TiB
m8g.2xlarge	10 GiB	N/A	3 TiB
m8g.4xlarge	10 GiB	N/A	6 TiB
m8g.8xlarge	10 GiB	N/A	12 TiB
m8g.12xlarge	10 GiB	N/A	18 TiB
m8g.16xlarge	10 GiB	N/A	24 TiB
r8g.medium	10 GiB	N/A	768 GiB
r8g.large	10 GiB	N/A	1532 GiB
r8g.xlarge	10 GiB	N/A	3 TiB
r8g.2xlarge	10 GiB	N/A	6 TiB
r8g.4xlarge	10 GiB	N/A	12 TiB
r8g.8xlarge	10 GiB	N/A	16 TiB
r8g.12xlarge	10 GiB	N/A	24 TiB
r8g.16xlarge	10 GiB	N/A	36 TiB

Tipo de instancia	Tamaño mínimo de EBS	Tamaño máximo de EBS (gp2)	Tamaño máximo de EBS (gp3)
r8gd.medium	N/A	N/A	N/A
r8gd.large	N/A	N/A	N/A
r8gd.xlarge	N/A	N/A	N/A
r8gd.2xlarge	N/A	N/A	N/A
r8gd.4xlarge	N/A	N/A	N/A
r8gd.8xlarge	N/A	N/A	N/A
r8gd.12xlarge	N/A	N/A	N/A
r8gd.16xlarge	N/A	N/A	N/A
oi2.large.search	N/A	N/A	N/A
oi2.xlarge.search	N/A	N/A	N/A
oi2.2xlarge.search	N/A	N/A	N/A
oi2.4xlarge.search	N/A	N/A	N/A
oi2.8xlarge.search	N/A	N/A	N/A
oi2.12xlarge.search	N/A	N/A	N/A
oi2.16xlarge.search	N/A	N/A	N/A

Cuotas de red

La siguiente tabla muestra el tamaño máximo de las cargas de solicitudes HTTP.

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
t2.micro.search	10 MiB
t2.small.search	10 MiB
t2.medium.search	10 MiB
t3.small.search	10 MiB
t3.medium.search	10 MiB
m3.medium.search	10 MiB
m3.large.search	10 MiB
m3.xlarge.search	100 MiB
m3.2xlarge.search	100 MiB
m4.large.search	10 MiB
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

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
m6g.xlarge.search	100 MiB
m6g.2xlarge.search	100 MiB
m6g.4xlarge.search	100 MiB
m6g.8xlarge.search	100 MiB
m6g.12xlarge.search	100 MiB
c4.large.search	10 MiB
c4.xlarge.search	100 MiB
c4.2xlarge.search	100 MiB
c4.4xlarge.search	100 MiB
c4.8xlarge.search	100 MiB
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

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
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
r4.2xlarge.search	100 MiB
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

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
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
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
i3.2xlarge.search	100 MiB
i3.4xlarge.search	100 MiB
i3.8xlarge.search	100 MiB
i3.16xlarge.search	100 MiB
or1.medium.search	10 MiB

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
or1.large.search	100 MiB
or1.xlarge.search	100 MiB
or1.2xlarge.search	100 MiB
or1.4xlarge.search	100 MiB
or1.8xlarge.search	100 MiB
or1.12xlarge.search	100 MiB
or1.16xlarge.search	100 MiB
or2.medium.search	100 MiB
or2.large.search	100 MiB
or2.xlarge.search	100 MiB
or2.2xlarge.search	100 MiB
or2.4xlarge.search	100 MiB
or2.8xlarge.search	100 MiB
or2.12xlarge.search	100 MiB
or2.16xlarge.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

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
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/A
i4i.xlarge.search	N/A
i4i.2xlarge.search	N/A
i4i.4xlarge.search	N/A
i4i.8xlarge.search	N/A
i4i.12xlarge.search	N/A
i4i.16xlarge.search	N/A
i4i.24xlarge.search	N/A
i4i.32xlarge.search	N/A
i4g.large.search	N/A
i4g.xlarge.search	N/A
i4g.2xlarge.search	N/A

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
i4g.4xlarge.search	N/A
i4g.8xlarge.search	N/A
i4g.16xlarge.search	N/A
i8g.large.search	N/A
i8g.xlarge.search	N/A
i8g.2xlarge.search	N/A
i8g.4xlarge.search	N/A
i8g.8xlarge.search	N/A
i8g.12xlarge.search	N/A
i8g.16xlarge.search	N/A
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
c7i.large.search	100 MiB
c7i.xlarge.search	100 MiB
c7i.2xlarge.search	100 MiB

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
c7i.4xlarge.search	100 MiB
c7i.8xlarge.search	100 MiB
c7i.12xlarge.search	100 MiB
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

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
c8g.2xlarge	100 MiB
c8g.4xlarge	100 MiB
c8g.8xlarge	100 MiB
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

Tipo de instancia	Tamaño máximo de las cargas de solicitud HTTP
r8g.16xlarge	100 MiB
r8gd.medium	100 MiB
r8gd.large	100 MiB
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
oi2.8xlarge.search	100 MiB
oi2.12xlarge.search	100 MiB
oi2.16xlarge.search	100 MiB

Cuotas de tamaño de la partición

En la siguiente sección, se enumeran los tamaños máximos de las particiones para varias familias de instancias.

Tipo de instancia	Multi-AZ sin modo de espera	Multi-AZ con modo de espera
R5, C5, M5, C7i, M7i, R7i	N/A	65 GiB
I3, i4i, i4g, i8g, i7i	N/A	65 GiB
R6g, C6g, M6g, R6gd, C7g,M7g, R7g, R7gd, C8g, M8g, R8g, R8gd	N/A	65 GiB
OR1, OR2, OM2, OI2	N/A	N/A
Im4gn	N/A	65 GiB

Para solicitar un aumento de cuota, póngase en contacto con AWS Support.

Cuotas de recuento de la partición

En la siguiente sección, se muestra el número máximo de fragmentos para las versiones. OpenSearch

Versión del motor	Límite	Notas
Elasticsearch de la versión 1.5 a la 6.x	Sin valor predeterminado	
Elasticsearch 7.x	1 000	El límite predeterminado se puede cambiar mediante la configuración de clúster. max_shards_per_node.
OpenSearch De 1.x a 2.15	1 000	El límite predeterminado se puede cambiar mediante la configuración de clúster. max_shards_per_node.
OpenSearch 2.17 y superiores	1000 por cada 16 GB de heap hasta un máximo de 4000	El límite predeterminados no se puede cambiar.

Cuota de procesamiento de Java

OpenSearch El servicio limita la memoria de procesamiento de Java al 50% de la memoria total disponible (máximo 32 GiB). El límite máximo de 32 GB no se aplica a las instancias r7g ni a las optimizadas. OpenSearch

Para solicitar un aumento del límite, ponte en contacto con AWS Support.

Los usuarios avanzados también pueden especificar el porcentaje del montón utilizado para los datos de campo. Para obtener más información, consulte Configuración avanzada del clúster y OutOfMemoryErrorJVM.

Cuota de políticas de dominio

OpenSearch El servicio limita las políticas de acceso a los dominios a 100 KiB.

Instancias reservadas en Amazon OpenSearch Service

Las instancias reservadas (RIs) de Amazon OpenSearch Service ofrecen descuentos importantes en comparación con las instancias bajo demanda estándar. Las instancias en sí son idénticas; solo RIs se trata de un descuento de facturación que se aplica a las instancias bajo demanda de su cuenta. En el caso de aplicaciones de larga duración con un uso predecible, RIs puede suponer un ahorro considerable a lo largo del tiempo.

OpenSearch El servicio RIs requiere plazos de uno o tres años y tiene tres opciones de pago que afectan a la tasa de descuento:

- Sin pago inicial: no se paga nada por adelantado. Se paga una tarifa por hora con descuento por cada hora dentro del periodo estipulado.
- Pago inicial parcial: se paga una parte del costo por adelantado y después una tarifa por hora con descuento por cada hora dentro del periodo estipulado.
- Pago total anticipado: se paga la totalidad del costo por adelantado. No hay que pagar una tarifa por hora durante el periodo estipulado.

Por regla general, un pago inicial más grande significa un mayor descuento. Las instancias reservadas no se pueden cancelar: cuando se reservan, se compromete a pagar por el periodo completo, y los pagos iniciales no son reembolsables.

RIs no son flexibles; solo se aplican al tipo de instancia exacto que reserve. Por ejemplo, una reserva para ocho instancias `c5.2xlarge.search` no se aplica a dieciséis instancias `c5.xlarge.search` o cuatro instancias `c5.4xlarge.search`. Sin embargo, las cuentas vinculadas que forman parte de una organización AWS Organizations pueden beneficiarse de cualquier solicitud de descuento no utilizada de la cuenta propietaria de la RI, siempre que los tipos de instancia, la región, la familia y el tamaño coincidan. Para obtener más información, consulta los precios y las preguntas frecuentes de Amazon OpenSearch Service.

Comprar instancias reservadas (consola)

La consola le permite ver las instancias reservadas que tiene y adquirir otras nuevas.

Para comprar una reserva

1. Vaya a https://aws.amazon.com, y seleccione Sign In to the Console (Inicie sesión en la consola).
2. En Analytics, selecciona Amazon OpenSearch Service.
3. Seleccione Reserved Instance Leases (Asignaciones de instancias reservadas) en el panel de navegación.

   En esta página, puede ver la reservas que tiene. Si tiene muchas, puede filtrarlas para identificar y ver más fácilmente una reserva en concreto.

   > **Tip**
   >
   > Si no ve el enlace de Reserved Instances Leases (Asignaciones de instancias reservadas), cree un dominio en la Región de AWS.

4. Seleccione Order reserved instance (Pedir instancia reservada).
5. Proporcione un nombre único y descriptivo.
6. Seleccione un tipo de instancia y el número de instancias. Para obtener instrucciones, consulte the section called “Determinación del tamaño de dominios”.
7. Seleccione la duración del periodo y una opción de pago. Revise los datos de pago con cuidado.
8. Seleccione Next (Siguiente).
9. Revise el resumen de compra con cuidado. Las compras de instancias reservadas no son reembolsables.
10. Seleccione Order (Pedir).

Adquisición de instancias reservadas (AWS CLI)

AWS CLI Tiene comandos para ver las ofertas, comprar una reserva y ver tus reservas. El siguiente comando y un ejemplo de respuesta muestran las ofertas de un determinado comando Región de 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"
    }
  ]
}
```

Para ver una explicación de cada valor de retorno, consulte la tabla siguiente.

Campo	Description (Descripción)
`FixedPrice`	El costo inicial de la reserva.
`ReservedInstanceOfferingId`	El ID de la oferta. Anote este valor si desea reservar la oferta.
`RecurringCharges`	La tarifa por hora de la reserva.
`UsagePrice`	Campo heredado. En el OpenSearch caso del servicio, este valor es siempre 0.

Campo	Description (Descripción)
`PaymentOption`	No Upfront (Sin pago inicial), Partial Upfront (Pago inicial parcial) o All Upfront (Pago inicial total).
`Duration`	La duración en segundos: • 31536000 segundos equivale a un año. • 94608000 segundos equivale a tres años.
`InstanceType`	El tipo de instancia de la reserva. Para obtener información sobre los recursos de hardware que se asignan a cada tipo de instancia, consulta los precios OpenSearch de Amazon Service.
`CurrencyCode`	La divisa para `FixedPrice` y `RecurringChargeAmount` .

El siguiente ejemplo adquiere una reserva:

```
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"
}
```

Por último, puede obtener una lista de las reservas para una región determinada mediante el siguiente ejemplo:

```
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` se expresa en formato de tiempo Unix, que es el número de segundos que han transcurrido desde la medianoche UTC del 1 de enero de 1970. Por ejemplo, 1522872571 en formato de tiempo Unix son las 20:09:31 UTC del 4 de abril de 2018. Puede utilizar convertidores online.

Para obtener más información acerca de los comandos utilizados en los ejemplos anteriores, consulte la Referencia de comandos de la AWS CLI.

Compra de instancias reservadas (AWS SDKs)

Las AWS SDKs (excepto Android e iOS SDKs) admiten todas las operaciones que se definen en la referencia de la API de Amazon OpenSearch Service, incluidas las siguientes:

- `DescribeReservedInstanceOfferings`
- `PurchaseReservedInstanceOffering`
- `DescribeReservedInstances`

Este script de ejemplo utiliza el cliente Python de OpenSearchServicebajo nivel del AWS SDK para Python (Boto3) para comprar instancias reservadas. Se debe proporcionar un valor para `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)
```

Para obtener más información sobre la instalación y el uso de AWS SDKs, consulte los kits de desarrollo de AWS software.

Examinar los costos

El Explorador de costos es una herramienta gratuita que puede utilizar para ver sus datos de gasto de los últimos 13 meses. El análisis de estos datos le ayuda a identificar las tendencias y a comprender si RIs se ajustan a su caso de uso. Si ya lo ha hecho RIs, puede agruparlos por opción de compra y mostrar los costos amortizados para compararlos con los gastos en instancias bajo demanda. También puede definir presupuestos de utilización para asegurarse de que aprovecha al máximo sus reservas. Para obtener más información, consulte Análisis de los costos con Cost Explorer en la Guía del usuario de AWS Billing .

Técnicas de optimización de costes para Amazon OpenSearch Service

Las siguientes son algunas de las técnicas más utilizadas para optimizar los costes al utilizar Amazon OpenSearch Service, tanto con clústeres gestionados como sin servidor. Como cada

carga de trabajo es única, evalúe estas estrategias en función de sus patrones de uso específicos y valídelas en un entorno de prueba antes de aplicarlas a la producción.

Temas

- Optimización de costes para los clústeres gestionados por Amazon OpenSearch Service
- Optimización de costes para Amazon OpenSearch Service Serverless

Optimización de costes para los clústeres gestionados por Amazon OpenSearch Service

Fuente derivada: omita el almacenamiento del campo _source

La fuente derivada es una función de optimización del almacenamiento que elimina la sobrecarga que supone almacenar el `_source` campo:

- OpenSearch almacena cada documento ingerido dos veces: una vez en el `_source` campo (documento sin procesar) y otra como campos indexados para la búsqueda.
- El `_source` campo por sí solo puede consumir un espacio de almacenamiento considerable, a menudo entre el 30 y el 50% del total del almacenamiento indexado.
- Con Derived Source, puede omitir el almacenamiento `_source` y, en su lugar, reconstruirlo dinámicamente a partir de campos indexados cuando es necesario (durante las operaciones de búsqueda, obtención, mget, reindexación o actualización).
- Esta opción es opcional y se habilita al crear el índice mediante una configuración de índice compuesto.
- Disponible en todas las regiones en las que se admite la OpenSearch versión 3.1 o posterior.

Ideal para: cargas de trabajo analíticas y de registro en las que no es necesario recuperar el documento original sin procesar con frecuencia, pero sí buscar y agregar campos.

Para obtener más información, consulte la documentación de código abierto y el anuncio sobre las novedades.

OR1 / OR2 / OM2 instances: familias OpenSearch de instancias optimizadas

OR1 y las OM2 instancias más nuevas OR2 utilizan Amazon S3 para el almacenamiento de réplicas mediante la replicación de segmentos:

- OR2: El rendimiento de indexación es hasta un 26% superior al de las instancias OR1 R7g, un 70% más.
- OM2: Un rendimiento de indexación hasta un 15% superior al de las instancias M7g OR1, un 66% más.
- Ambas utilizan la misma arquitectura: EBS local para el almacenamiento principal y S3 para mayor durabilidad.
- Elimina el costo del almacenamiento de réplicas: las réplicas se almacenan en S3 (con una durabilidad de 11 nueves) en lugar de los costosos volúmenes de EBS.
- Hasta un 30% de mejora en la relación precio-rendimiento con respecto a las instancias de la generación anterior.
- Admite la versión Shallow Snapshot v2, es decir, instantáneas casi instantáneas sin sobrecarga. I/O

Ideal para: cargas de trabajo de análisis operativo y análisis de registros con un uso intensivo de indexación.

Para obtener más información, consulte el anuncio OR2 y OM2 las novedades y el tema de la guía para the section called “Instancias optimizadas para OpenSearch ” desarrolladores.

Acumulaciones de índices: agrega datos históricos de series temporales

Los resúmenes de índices resumen y comprimen los datos de series temporales más antiguas en intervalos de tiempo más gruesos, lo que reduce drásticamente el volumen de almacenamiento:

- Datos de sensores y IoT: guarde los datos por segundo en almacenamiento activo de períodos recientes; acumule resúmenes por hora o por día para los datos más antiguos.
- Métricas del sistema: conserve las métricas detalladas de los últimos 30 días y agregue los datos más antiguos en resúmenes diarios o por hora.
- Datos de registro: conserve todos los detalles para el período de solución de problemas activo (por ejemplo, una semana); mantenga los patrones de error resumidos para los períodos anteriores.
- Combínelo con las políticas de ISM para automatizar la migración acumulativa y por niveles en una única política de ciclo de vida.
- Mayores ahorros al sumar segundos a horas en lugar de segundos a minutos.

Para obtener más información, consulta la entrada del blog Index Rollups.

Administración del estado de los índices: automatice todo el ciclo de vida de los datos

Las políticas de ISM automatizan el movimiento de los índices a través de los niveles de almacenamiento y las acciones del ciclo de vida:

- Migre automáticamente los índices: activos UltraWarm a fríos y eliminarlos, en función de la antigüedad, el tamaño o el número de documentos.
- Active las acumulaciones antes de la transición de nivel para reducir el volumen de datos.
- Establezca políticas de renovación (por ejemplo, cuando un índice alcance los 50 GB o tenga 30 días de antigüedad) para controlar el crecimiento del índice.
- Automatice y fuerce la fusión en índices de solo lectura para recuperar el espacio de almacenamiento de los documentos eliminados.
- Combínelo con paquetes acumulativos para ahorrar al máximo en conjuntos de datos de series temporales de gran tamaño.

Instancias reservadas: contrate descuentos predecibles

Para unas cargas de trabajo analíticas estables y predecibles, las instancias reservadas ofrecen importantes descuentos con respecto a los precios bajo demanda:

- Plazos de compromiso de 1 o 3 años, sin opciones de pago por adelantado, parcial o totalmente anticipado.
- Ideal para nodos de datos de nivel avanzado y nodos maestros dedicados que funcionan de forma continua.
- Utilice la calculadora de AWS precios para estimar los ahorros antes de comprometerse.
- Las instancias reservadas son un descuento de facturación que se aplica a las instancias bajo demanda, sin necesidad de realizar cambios en la infraestructura.

Tipos y recuentos de instancias del tamaño correcto

Guía clave de OpenSearch Well-Architected Lens y mejores prácticas para dimensionar correctamente:

- Utilice siempre instancias de la última generación (por ejemplo, las instancias Graviton3 ofrecen un rendimiento hasta un 25% superior al de las instancias basadas en Graviton2).

- Utilice los volúmenes EBS de gp3 en lugar de los de gp2: mejor rendimiento a un coste menor sin coste adicional.
- Haga coincidir el tipo de instancia con la carga de trabajo: optimizada en memoria para búsquedas intensivas, optimizada para procesamiento en caso de alta indexación.
- Evalúe los nodos de administrador de clústeres dedicados: solo se necesitan para 3 o más nodos de datos; evite aprovisionar en exceso el tamaño del nodo maestro.
- Supervise CloudWatch las métricas para detectar el sobreaprovisionamiento: mantener la CPU por debajo del 40%, la JVM por debajo del 50% y el almacenamiento por debajo del 50% son señales de despilfarro.
- Rangos óptimos: entre un 60 y un 80% de CPU, entre un 65 y un 85% de JVM y entre un 70 y un 85% de almacenamiento para cargas de trabajo sostenidas.

Para obtener más información, consulte la entrada del blog Right-Sizing Best Practices.

Optimización de fragmentos: evite la fragmentación excesiva

La fragmentación excesiva es un factor de costes oculto: demasiados fragmentos pequeños desperdician CPU, memoria y pila de JVM:

- Tamaños de partición recomendados: de 10 a 50 GiB por fragmento, según la carga de trabajo.
- No más de 25 fragmentos por GiB de pila de Java, no más de 1000 fragmentos por nodo de datos.
- Utilice las políticas de transferencia de ISM para controlar el crecimiento del índice y evitar la proliferación ilimitada de fragmentos.
- Reduzca el número de réplicas cuando la durabilidad lo permita (o OR2 las instancias eliminan OR1 por completo la necesidad de réplicas).
- Utilice la combinación forzada en índices de solo lectura para reducir el número de segmentos y recuperar espacio de almacenamiento.

Para obtener más información, consulte ¿Cuántos fragmentos necesito?

Cero-ETL/Consulta directa con Amazon S3

Para los datos que rara vez se consultan pero que deben permanecer accesibles, Direct Query (ETL cero con S3) permite consultar datos de S3 directamente desde allí sin ingerirlos: OpenSearch

- Sin coste de ingesta: los datos permanecen en S3.

- El archivado de los datos no supone un coste excesivo de almacenamiento.
- Pay-per-query modelo de cómputo.
- Admite OpenSearch paneles de control para la visualización.
- La latencia de segundos o minutos es aceptable, no en casos de uso en tiempo real.

Toma de muestras y compresión en el momento de la ingestión

Reduzca los costes incluso antes de que los datos lleguen a OpenSearch:

- Muestreo: ingiera solo un subconjunto representativo de flujos de registros de gran volumen (por ejemplo, el 10% de los registros de depuración).
- Compresión de índices: habilite el mejor códec de compresión para reducir el espacio de almacenamiento.
- Filtrado de campos: elimine los campos de alta cardinalidad y bajo valor antes de indexarlos (por ejemplo, los rastreos de pila sin procesar de registros antiguos).
- Políticas de retención: defina los intervalos de retención máximos en función de los requisitos de conformidad; nunca almacene los datos de forma indefinida.

Evite los costos de soporte prolongados: manténgase al día con las versiones del motor

Amazon OpenSearch Service cobra una tarifa fija por hora de instancia normalizada para las versiones de motor de Extended Support:

- El uso de versiones antiguas y no compatibles conlleva cargos adicionales además de los costes de la instancia.
- Actualice a las versiones compatibles actuales para evitar las tarifas de Extended Support.

Asignación de costos, etiquetas y CloudWatch monitoreo

La gobernanza proactiva de los costes evita el derroche:

- Aplica etiquetas de asignación de costes a OpenSearch los dominios para realizar un seguimiento detallado de los costes por equipo o carga de trabajo.
- Establezca CloudWatch alarmas para la utilización del almacenamiento, la presión de la JVM y la CPU para detectar el sobreaprovisionamiento de forma temprana.

- Utilice AWS Cost Explorer para identificar los dominios con una utilización baja y constante.
- Evalúe el ajuste automático: ajusta automáticamente el tamaño del montón de la JVM y otros ajustes para mejorar el rendimiento y reducir el desperdicio de recursos.

Optimización de costes para Amazon OpenSearch Service Serverless

Búsqueda vectorial optimizada en disco (colecciones vectoriales)

La búsqueda vectorial optimizada en disco es una de las técnicas de reducción de costes más eficaces para las cargas de trabajo vectoriales. Ejecuta la búsqueda vectorial a una fracción del coste del modo en memoria, ya que solo mantiene los vectores comprimidos en la RAM y almacena los vectores de máxima precisión en el disco.

Funcionamiento:

- En el modo standard (`in_memory`), el gráfico HNSW completo se carga en la RAM, lo que resulta prohibitivo a gran escala.
- En `on_disk` el modo, solo los vectores comprimidos (cuantificados) se guardan en la memoria para la generación de los candidatos; los vectores de máxima precisión se recuperan del disco únicamente para la fase final de repuntuación (búsqueda bifásica).
- Esto reduce drásticamente los requisitos de RAM y, al mismo tiempo, mantiene una alta calidad de búsqueda.
- El `on_disk` modo predeterminado utiliza una cuantificación binaria de 32 veces, lo que reduce los requisitos de memoria en un 97% en comparación con el modo en memoria.
- Admite los niveles de compresión: 2x (FP16), 4x (byte), 8x, 16x, 32x (binario).
- Latencia P90 en el rango de 100 a 200 ms: adecuada para cargas de trabajo que no requieren tiempos de respuesta de milisegundos de un solo dígito.

Parámetros de ahorro de costes:

Conjunto de datos	Recordemos @100	Latencia del P90	Reducción de costos
TREC-RAG coherente	0,94	104 ms	83%
Tab-1M	0.96	7 ms	67%

Conjunto de datos	Recordemos @100	Latencia del P90	Reducción de costos
Marco-1M	0.99	7 ms	67%

Ideal para: canalizaciones de RAG, búsqueda semántica, recuperación de documentos y cualquier carga de trabajo vectorial en la que una latencia de P90 de 100 a 200 ms sea aceptable y la reducción de costes sea una prioridad.

> **Note**
>
> Para aplicar este cambio a los datos indexados existentes, debe volver a indexarlos. Puede utilizar una herramienta de canalización externa, por ejemplo, para volver a indexar los datos en un índice nuevo.

Para obtener más información, consulte la entrada del blog Disk-Optimized Vector Search, la entrada del blog Quantization Techniques y. the section called “Trabajo con colecciones de búsqueda vectorial”

Optimización automática de vectores (colecciones de vectores)

La optimización automática evalúa automáticamente las configuraciones de los índices vectoriales y recomienda la mejor compensación entre la calidad de la búsqueda, la latencia y el coste de la memoria, sin necesidad de conocimientos sobre vectores:

- Ofrece recomendaciones de optimización en menos de una hora.
- Integrado con las canalizaciones de ingestión de vectores.
- Se puede combinar con la indexación acelerada por GPU para bases de datos vectoriales a escala multimillonaria.

Para obtener más información, consulta la entrada del blog sobre la optimización automática.

Indexación vectorial acelerada por GPU (colecciones de vectores)

La aceleración por GPU convierte la indexación vectorial en HNSW en una indexación sin servidor GPUs, lo que reduce drásticamente el tiempo y el coste de OCU que supone crear índices vectoriales de gran tamaño:

- Los tiempos de creación de índices son entre 6,4 y 13,8 veces más rápidos en comparación con la indexación solo con CPU.
- Hasta un 75% menos de coste de indexación de la OCU para cargas de trabajo vectoriales de escritura intensa.
- GPUs se conectan de forma dinámica: solo se paga OCUs cuando la aceleración de la GPU está activa.
- Permite crear bases de datos vectoriales a escala multimillonaria en menos de una hora.
- Se carga por separado como aceleración vectorial. OCUs

Ideal para: ingestión inicial de vectores a gran escala o escenarios frecuentes de reentrenamiento de modelos en los que la reconstrucción de los índices resulta costosa.

Para obtener más información, consulta la entrada del blog GPU Acceleration.

Establece los límites máximos de OCU (todos los tipos de colecciones)

Amazon OpenSearch Service Serverless escala automáticamente OCUs en función de la demanda. Sin un límite, los costos pueden aumentar inesperadamente. Establezca un límite máximo de OCU a nivel de cuenta o por grupo de cobranza para evitar una expansión descontrolada. El sistema se amplía hasta este límite durante los picos de carga, pero no lo superará.

Valores permitidos:

- Nivel de cuenta: cualquier valor hasta 1700 OCUs (no limitado a múltiplos de 16).
- Grupos de cobro: 1, 2, 4, 8, 16 y múltiplos de 16 a 1696. OCUs

Supervisa CloudWatch las métricas (`OCUUtilization`) para ajustar tu configuración máxima de OCU a lo largo del tiempo.

> **Note**
>
> Si la utilización alcanza el límite máximo de OCU, el rendimiento puede degradarse considerablemente aunque los costes estén contenidos. Alcanzar el límite no resuelve la causa subyacente del aumento de la OCU. En el caso de las colecciones de vectores, la causa principal suelen ser los vectores en memoria, que deben abordarse directamente optimizando la indexación vectorial, reduciendo el tamaño del índice o ajustando las compensaciones entre la recuperación y la precisión.

> **Tip**
>
> Comience con una OCU máxima conservadora y aumente solo cuando CloudWatch muestre una utilización sostenida cerca del límite. Si siempre llegas al límite, investiga la causa principal (especialmente el uso de vectores en memoria para las colecciones de vectores) en lugar de limitarte a aumentar el límite.

Para obtener más información, consulte the section called “Administración de los límites de capacidad”.

Optimice el período de retención (recopilaciones de series temporales)

Las políticas del ciclo de vida de los datos eliminan automáticamente los datos de las recopilaciones de series temporales después de un período de retención específico, lo que impide un crecimiento ilimitado del almacenamiento. Solo las recopilaciones de series temporales respaldan las políticas del ciclo de vida de los datos; las recopilaciones vectoriales y de búsqueda no.

El recuento de OCU para las recopilaciones de series temporales depende directamente de la cantidad de datos recientes que deben guardarse en el almacenamiento local. Las recopilaciones de series temporales guardan solo la parte más reciente de los datos en el almacenamiento local de la OCU; los datos más antiguos se transfieren a S3 y el número de datos se escala en consecuencia: OCUs

OCUs obligatorio = máximo (mínimo OCUs, OCUs necesario para mantener los datos dentro del período de retención)

Configuración de políticas del ciclo de vida de los datos:

- Establezca períodos de retención de 24 horas a 3650 días por índice o patrón de índice.
- Amazon OpenSearch Service Serverless elimina los datos automáticamente haciendo el mejor esfuerzo posible (normalmente en un plazo de 48 horas o el 10% del período de retención).
- Las reglas se pueden aplicar a nivel de recopilación, a nivel de patrón de índice o a nivel de índice individual; prevalecen las reglas más específicas.

Ejemplo de dimensionamiento:

- 1 TiB/day ingesta con una retención de 30 días = aproximadamente 1 TiB de datos calientes = 20 OCUs para indexación + 20 para búsqueda. OCUs

- Reducir a una retención de 7 días = aproximadamente 233 GiB de datos importantes = aproximadamente OCUs 4 para la indexación + OCUs 4 para la búsqueda.

Una retención más corta se traduce en menos datos importantes en el almacenamiento local, se OCUs necesitan menos datos y se reduce la factura informática. Alinee los períodos de retención con los requisitos empresariales y de cumplimiento reales; de forma predeterminada, no conserve los datos indefinidamente.

Para obtener más información, consulte the section called “Uso de políticas de ciclo de vida de los datos”.

Evite almacenar datos innecesarios (todos los tipos de recopilación)

Al reducir el volumen de datos ingeridos, se reducen directamente los costes de cómputo (OCUs) y de almacenamiento:

- Filtre los campos en el momento de la ingesta: utilice canalizaciones para eliminar los campos de bajo valor antes de que lleguen a la colección.
- Evite ingerir datos duplicados o redundantes: deduplique a nivel de canalización.
- Utilice las asignaciones de índice adecuadas: desactive la indexación en los campos que están almacenados pero que nunca se buscan (). `"index": false`
- Para las colecciones de búsquedas: evite almacenar bloques binarios grandes o texto sin procesar, ya que aumenta el espacio de almacenamiento sin valor de búsqueda.

Grupos de recopilación para cargas de trabajo con varios usuarios (todos los tipos de recopilación)

Los grupos de recopilación permiten que varias recopilaciones con diferentes claves de KMS compartan los recursos de la OCU dentro del mismo límite de seguridad, lo que reduce drásticamente los costos de las arquitecturas multiusuario. Aplicable a los clientes que utilizan varias claves KMS por inquilino o por colección:

- Anteriormente, cada clave KMS única requería una clave dedicada OCUs , lo que hacía que el aislamiento por inquilino fuera prohibitivo.
- Con los grupos de recopilación, los inquilinos con claves de cifrado independientes pueden compartir la capacidad de la OCU.

- Ahorro de costes de hasta un 90% para un gran número de cargas de trabajo para inquilinos más pequeños.
- Soporta tanto el mínimo OCUs (línea base garantizada, sin arranques en frío) como el máximo OCUs (límite de costes).
- Las colecciones con diferentes tipos de acceso a la red (pública y VPC) pueden coexistir en el mismo grupo.
- CloudWatch las métricas proporcionan visibilidad por grupo sobre el consumo de recursos y la latencia.

Ideal para: proveedores de SaaS, plataformas multiusuario o cualquier carga de trabajo con muchas colecciones pequeñas, cada una de las cuales requiera su propia clave de KMS.

Para obtener más información, consulte la entrada del blog sobre los grupos de colecciones, el anuncio sobre las novedades y. the section called “Grupos de colecciones”

Otros recursos compatibles en Amazon OpenSearch Service

En este tema se describen los recursos adicionales compatibles con Amazon OpenSearch Service.

bootstrap.memory_lock

OpenSearch El servicio se activa `bootstrap.memory_lock`opensearch.yml, lo que bloquea la memoria de la JVM e impide que el sistema operativo la cambie al disco. Esto se aplica a todos los tipos de instancias admitidos excepto a los siguientes:

- `t2.micro.search`
- `t2.small.search`
- `t2.medium.search`
- `t3.small.search`
- `t3.medium.search`

Módulo de scripting

OpenSearch El servicio admite la creación de scripts para Elasticsearch 5. x y dominios posteriores. No admite scripting para las versiones 1.5 o 2.3.

Las opciones de scripting admitidas incluyen lo siguiente:

- Painless
- Lucene Expressions
- Mustache

Para los dominios de Elasticsearch 5.5 y versiones posteriores, y para todos los OpenSearch dominios, OpenSearch Service admite los scripts almacenados mediante el `_scripts` punto final. Los dominios de Elasticsearch 5.3 y 5.1 solo admiten scripts incorporados.

Transporte TLS

OpenSearch El servicio admite HTTP en el puerto 80 y HTTPS en el puerto 443, pero no admite el transporte TLS.

Tutoriales OpenSearch de Amazon Service

En este capítulo se incluyen varios start-to-finish tutoriales para trabajar con Amazon OpenSearch Service, que incluyen cómo migrar al servicio, crear una aplicación de búsqueda sencilla y crear una visualización en los OpenSearch paneles de control.

Temas

- Tutorial: Creación y búsqueda de documentos en Amazon OpenSearch Service
- Tutorial: Migración a Amazon Service OpenSearch
- Tutorial: Creación de una aplicación de búsqueda con Amazon OpenSearch Service
- Tutorial: Visualización de las llamadas de atención al cliente con el OpenSearch servicio y los paneles OpenSearch

Tutorial: Creación y búsqueda de documentos en Amazon OpenSearch Service

En este tutorial, aprenderás a crear y buscar un documento en Amazon OpenSearch Service. Los datos se añaden a un índice en forma de documento JSON. OpenSearch El servicio crea un índice alrededor del primer documento que se agrega.

Este tutorial explica cómo llevar a cabo solicitudes HTTP para crear documentos, generar automáticamente un ID para un documento y hacer búsquedas básicas y avanzadas en los documentos.

> **Note**
>
> En este tutorial, se utiliza un dominio de acceso abierto. Para obtener el máximo nivel de seguridad, es recomendable que coloque su dominio dentro de una nube privada virtual (VPC).

Requisitos previos

Este tutorial tiene los requisitos previos siguientes:

- Debe tener un Cuenta de AWS.
- Debe tener un dominio de OpenSearch servicio activo.

Adición de un documento a un índice

Para añadir un documento a un índice, puede utilizar cualquier herramienta HTTP, como Postman, cURL o OpenSearch la consola Dashboards. En estos ejemplos se supone que está utilizando la consola de desarrollador en Dashboards. OpenSearch Si utiliza una herramienta diferente, ajústela según corresponda proporcionando la URL completa y las credenciales, si es necesario.

Para agregar un documento a un índice

1. Navega hasta la URL de los OpenSearch paneles de control de tu dominio. Puedes encontrar la URL en el panel de control del dominio, en la consola OpenSearch de servicio. La URL tiene este formato:

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

2. Inicie sesión con su nombre de usuario y contraseña principales.
3. Abra el panel de navegación izquierdo y elija Herramientas para desarrolladores.
4. El verbo HTTP para crear un nuevo recurso es PUT, que es lo que se utiliza para crear un nuevo documento e índice. Ingrese el siguiente comando en la consola:

   ```
   PUT fruit/_doc/1
   {
     "name":"strawberry",
     "color":"red"
   }
   ```

   La solicitud PUT crea un índice llamado fruit (fruta) y agrega un solo documento al índice con un ID de 1. Produce la siguiente respuesta:

   ```
   {
     "_index" : "fruit",
     "_type" : "_doc",
     "_id" : "1",
     "_version" : 1,
     "result" : "created",
     "_shards" : {
   ```

```
    "total" : 2,
    "successful" : 2,
    "failed" : 0
  },
  "_seq_no" : 0,
  "_primary_term" : 1
}
```

La creación se genera automáticamente IDs

OpenSearch El servicio puede generar automáticamente una identificación para sus documentos. El comando que se genera IDs utiliza una solicitud POST en lugar de una solicitud PUT y no requiere ningún identificador de documento (en comparación con la solicitud anterior).

Ingrese la siguiente solicitud en la consola para desarrolladores:

```
POST veggies/_doc
{
  "name":"beet",
  "color":"red",
  "classification":"root"
}
```

Esta solicitud crea un índice denominado veggies (verduras) y agrega el documento al índice. Produce la siguiente respuesta:

```
{
  "_index" : "veggies",
  "_type" : "_doc",
  "_id" : "3WgyS4IB5DLqbRIvLxtF",
  "_version" : 1,
  "result" : "created",
  "_shards" : {
    "total" : 2,
    "successful" : 2,
    "failed" : 0
  },
  "_seq_no" : 0,
  "_primary_term" : 1
}
```

Tenga en cuenta el campo `_id` adicional de la respuesta, que indica que se creó un ID automáticamente.

> **Note**
>
> No proporcione nada después de `_doc` en la URL, donde normalmente va el ID. Como va a crear un documento con un ID generado, aún no debe proporcionar uno. Eso está reservado para actualizaciones.

Actualización de un documento con un comando POST

Para actualizar un documento, se utiliza un comando `POST` de HTTP con el número de ID.

En primer lugar, cree un documento con un ID de `42`:

```
POST fruits/_doc/42
{
  "name":"banana",
  "color":"yellow"
}
```

A continuación, utilice ese ID para actualizar el documento:

```
POST fruits/_doc/42
{
  "name":"banana",
  "color":"yellow",
  "classification":"berries"
}
```

Este comando actualiza el documento con el nuevo campo `classification`. Produce la siguiente respuesta:

```
{
  "_index" : "fruits",
  "_type" : "_doc",
  "_id" : "42",
  "_version" : 2,
  "result" : "updated",
  "_shards" : {
```

```
    "total" : 2,
    "successful" : 2,
    "failed" : 0
  },
  "_seq_no" : 1,
  "_primary_term" : 1
}
```

> **Note**
>
> Si intenta actualizar un documento que no existe, OpenSearch Service lo crea.

Ejecución de acciones por lotes

Puede utilizar la operación de la API de `POST _bulk` para llevar a cabo varias acciones en uno o más índices en una solicitud. Los comandos de acción por lotes tienen el siguiente formato:

```
POST /_bulk
<action_meta>\n
<action_data>\n
<action_meta>\n
<action_data>\n
```

Cada acción requiere dos líneas de JSON. Primero, debe proporcionar la descripción de la acción o los metadatos. En la línea siguiente, debe proporcionar los datos. Cada parte está separada por una nueva línea (\n). La descripción de una acción de una inserción podría verse de la siguiente manera:

```
{ "create" : { "_index" : "veggies", "_type" : "_doc", "_id" : "7" } }
```

La línea siguiente que contiene los datos podría verse de la siguiente manera:

```
{ "name":"kale", "color":"green", "classification":"leafy-green" }
```

En conjunto, los metadatos y los datos representan una sola acción en una operación masiva. Puede llevar a cabo muchas operaciones en una sola solicitud, como esta:

```
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" } }
```

Observe que la última acción es `delete`. No hay datos después de la acción `delete`.

Búsqueda de documentos

Ya que existen datos en el clúster, puede buscarlos. Por ejemplo, es posible que quiera buscar todos los tubérculos, obtener un recuento de todos los vegetales de hoja verde o encontrar la cantidad de errores registrados por hora.

Búsquedas básicas

Una búsqueda básica tiene un aspecto similar a este:

```
GET veggies/_search?q=name:l*
```

La solicitud produce una respuesta JSON que contiene el documento de lechuga.

Búsquedas avanzadas

Puede hacer búsquedas más avanzadas si proporciona las opciones de consulta como JSON en el cuerpo de la solicitud:

```
GET veggies/_search
{
  "query": {
    "term": {
      "name": "lettuce"
    }
  }
}
```

Este ejemplo también produce una respuesta JSON con el documento de lechuga.

Ordenar

Puede llevar a cabo más consultas de este tipo mediante la acción de ordenar. En primer lugar, debe volver a crear el índice porque la asignación automática de campos eligió tipos que no se pueden ordenar de forma predeterminada. Envíe las siguientes solicitudes para eliminar y volver a crear el índice:

```
DELETE /veggies

PUT /veggies
{
   "mappings":{
      "properties":{
         "name":{
            "type":"keyword"
         },
         "color":{
            "type":"keyword"
         },
         "classification":{
            "type":"keyword"
         }
      }
   }
}
```

A continuación, rellene el índice con datos:

```
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" }
```

Ya puede buscar con un orden. Esta solicitud agrega un orden ascendente por clasificación:

```
GET veggies/_search
{
```

```
  "query" : {
    "term": { "color": "green" }
  },
  "sort" : [
      "classification"
  ]
}
```

Recursos relacionados

Para obtener más información, consulte los siguientes recursos:

- Introducción
- Indexación de datos
- Buscar datos

Tutorial: Migración a Amazon Service OpenSearch

Las instantáneas de índice son una forma popular de migrar de un clúster de Elasticsearch heredado OpenSearch o autogestionado a Amazon Service. OpenSearch En términos generales, el proceso consta de los siguientes pasos:

1. Tome una instantánea del clúster existente y cargue la instantánea en un bucket de Amazon S3.
2. Crea un dominio de servicio. OpenSearch
3. Conceda permisos al OpenSearch Servicio para acceder al depósito y asegúrese de tener permisos para trabajar con las instantáneas.
4. Restaure la instantánea en el dominio del OpenSearch servicio.

Este tutorial proporciona pasos más detallados y opciones alternativas, cuando corresponda.

Tomar y cargar la instantánea

Si bien puedes usar el complemento repository-s3 para tomar instantáneas directamente en S3, debes instalar el complemento en cada nodo, realizar ajustes `opensearch.yml` (o `elasticsearch.yml` si usas un clúster de Elasticsearch), reiniciar cada nodo, añadir tus AWS credenciales y, finalmente, tomar la instantánea. El complemento es una gran opción para uso continuo o para migrar clústeres más grandes.

En el caso de los clústeres más pequeños, una solución única consiste en tomar una instantánea del sistema de archivos compartido y, a continuación, utilizarla para subirla a S3. AWS CLI Si ya tiene una instantánea, vaya al paso 4.

Cómo tomar una instantánea y cargarla a Amazon S3

1. Agregue la configuración `path.repo` a `opensearch.yml` (o `Elasticsearch.yml`) en todos los nodos y, a continuación, reinicie cada nodo.

   ```
   path.repo: ["/my/shared/directory/snapshots"]
   ```

2. Registre un repositorio de instantáneas, necesario antes de tomar una instantánea. Un repositorio es solo una ubicación de almacenamiento: un sistema de archivos compartido, Amazon S3, un sistema de archivos distribuido de Hadoop (HDFS), etc. En este caso, utilizaremos un sistema de archivos compartido (“fs”):

   ```
   PUT _snapshot/my-snapshot-repo-name
   {
     "type": "fs",
     "settings": {
       "location": "/my/shared/directory/snapshots"
     }
   }
   ```

3. Tomar la instantánea:

   ```
   PUT _snapshot/my-snapshot-repo-name/my-snapshot-name
   {
     "indices": "migration-index1,migration-index2,other-indices-*",
     "include_global_state": false
   }
   ```

4. Instale la AWS CLI y ejecute `aws configure` para agregar sus credenciales.

5. Desplácese hasta el directorio de instantáneas. A continuación, ejecute los siguientes comandos para crear un nuevo bucket de S3 y cargar el contenido del directorio de instantáneas en ese bucket:

   ```
   aws s3 mb s3://amzn-s3-demo-bucket --region us-west-2
   aws s3 sync . s3://amzn-s3-demo-bucket --sse AES256
   ```

En función del tamaño de la instantánea y la velocidad de su conexión a internet, esta operación puede tardar un tiempo.

Crear un dominio

Aunque la consola es la forma más sencilla de crear un dominio, en este caso ya tienes el terminal abierto e AWS CLI instalado. Modifique el siguiente comando para crear un dominio que se ajuste a sus necesidades:

```
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
```

Tal como está, el comando crea un dominio accesible a Internet con dos nodos de datos, cada uno con 100 GiB de almacenamiento. También habilita el control de acceso detallado con autenticación básica HTTP y toda la configuración de cifrado. Utilice la consola de OpenSearch servicio si necesita una configuración de seguridad más avanzada, como una VPC.

Antes de ejecutar el comando, cambie el nombre de dominio, las credenciales del usuario maestro y el número de cuenta. Especifique la misma Región de AWS que utilizó para el bucket de S3 y una OpenSearch/Elasticsearch versión que sea compatible con la instantánea.

Important

Las instantáneas solo son compatibles con versiones posteriores y solo con una versión principal. Por ejemplo, no puede restaurar una instantánea a partir de un OpenSearch

1. x clúster en un Elasticsearch 7. x clúster, solo un OpenSearch 1. x o 2. cúmulo x. La versión secundaria también importa. No puede restaurar una instantánea de un clúster 5.3.3 autogestionado en un dominio de servicio 5.3.2 OpenSearch . Te recomendamos que elijas la versión más reciente de Elasticsearch OpenSearch compatible con tu instantánea. Para obtener una tabla de versiones compatibles, consulte the section called “Uso de una instantánea para migrar datos”.

Conceda permisos al bucket de S3.

En la consola AWS Identity and Access Management (IAM), cree un rol con los siguientes permisos y relación de confianza. Al crear un rol, elija S3 como el Servicio de AWS . Asigne al rol el nombre `OpenSearchSnapshotRole` para que sea fácil de encontrar.

Permisos

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/*"
      ]
    }
  ]
```

```
}
```

Relación de confianza

JSON

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

Después, conceda permisos a su rol de IAM personal para asumir `OpenSearchSnapshotRole`. Cree la siguiente política y adjúntela a su identidad.

Permisos

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
      "Effect": "Allow",
      "Action": "iam:PassRole",
      "Resource": "arn:aws:iam::123456789012:role/OpenSearchSnapshotRole"
    }
  ]
}
```

Asigne la función de instantánea en los OpenSearch paneles de control (si utiliza un control de acceso detallado)

Si ha habilitado el control de acceso detallado, incluso si utiliza autenticación HTTP básica para todos los demás fines, debe asignar el rol `manage_snapshots` al rol de IAM para poder trabajar con instantáneas.

Para otorgar permisos a la identidad a fin de trabajar con instantáneas

1. Inicie sesión en Dashboards con las credenciales de usuario maestro que especificó al crear el dominio del servicio. OpenSearch Puede encontrar la URL de los paneles en la consola de OpenSearch servicio. Adopta el formato `https://domain-endpoint/_dashboards/`.
2. En el menú principal, seleccione Seguridad, Roles y seleccione el rol manage_snapshots.
3. Seleccione Usuarios asignados, Administrar mapeo.
4. Agregue el ARN del dominio de su rol de IAM personal en el campo apropiado. El ARN está en uno de los siguientes formatos:

   ```
   arn:aws:iam::123456789123:user/user-name
   ```

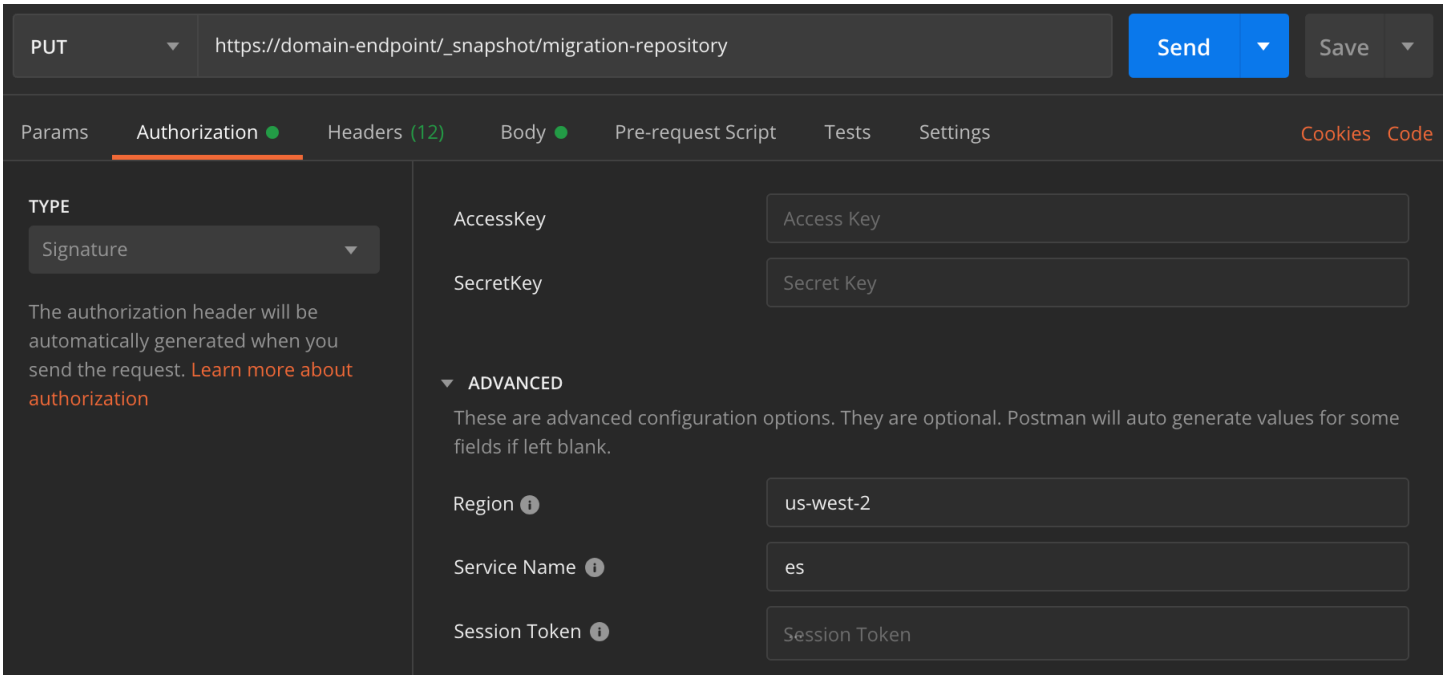
   ```
   arn:aws:iam::123456789123:role/role-name
   ```

5. Seleccione Asignar y confirme que el rol aparece en Usuarios asignados.

Restaurar la instantánea

En este momento, tiene dos formas de acceder a su dominio de OpenSearch servicio: mediante la autenticación básica HTTP con sus credenciales de usuario maestro o mediante la AWS autenticación con sus credenciales de IAM. Como las instantáneas utilizan Amazon S3, que no tiene el concepto de usuario maestro, debe utilizar sus credenciales de IAM para registrar el repositorio de instantáneas en su dominio de OpenSearch servicio.

La mayoría de los lenguajes de programación tienen bibliotecas para ayudar con las solicitudes de firma, pero el enfoque más simple es utilizar una herramienta como Postman y poner sus credenciales de IAM en la sección Autorización.



Para restaurar la instantánea

1. Independientemente de cómo elija firmar sus solicitudes, el primer paso es registrar el repositorio:

```
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. A continuación, enumere las instantáneas en el repositorio y busque la que desea restaurar. En este punto, puede continuar utilizando Postman o cambiar a una herramienta como curl.

   Abreviado

   ```
   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. Restaurare la instantánea.

   Abreviado

   ```
   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. Finalmente, verifique que sus índices se restablecieron como se esperaba.

   Abreviado

   ```
   GET _cat/indices?v
   ```

   curl

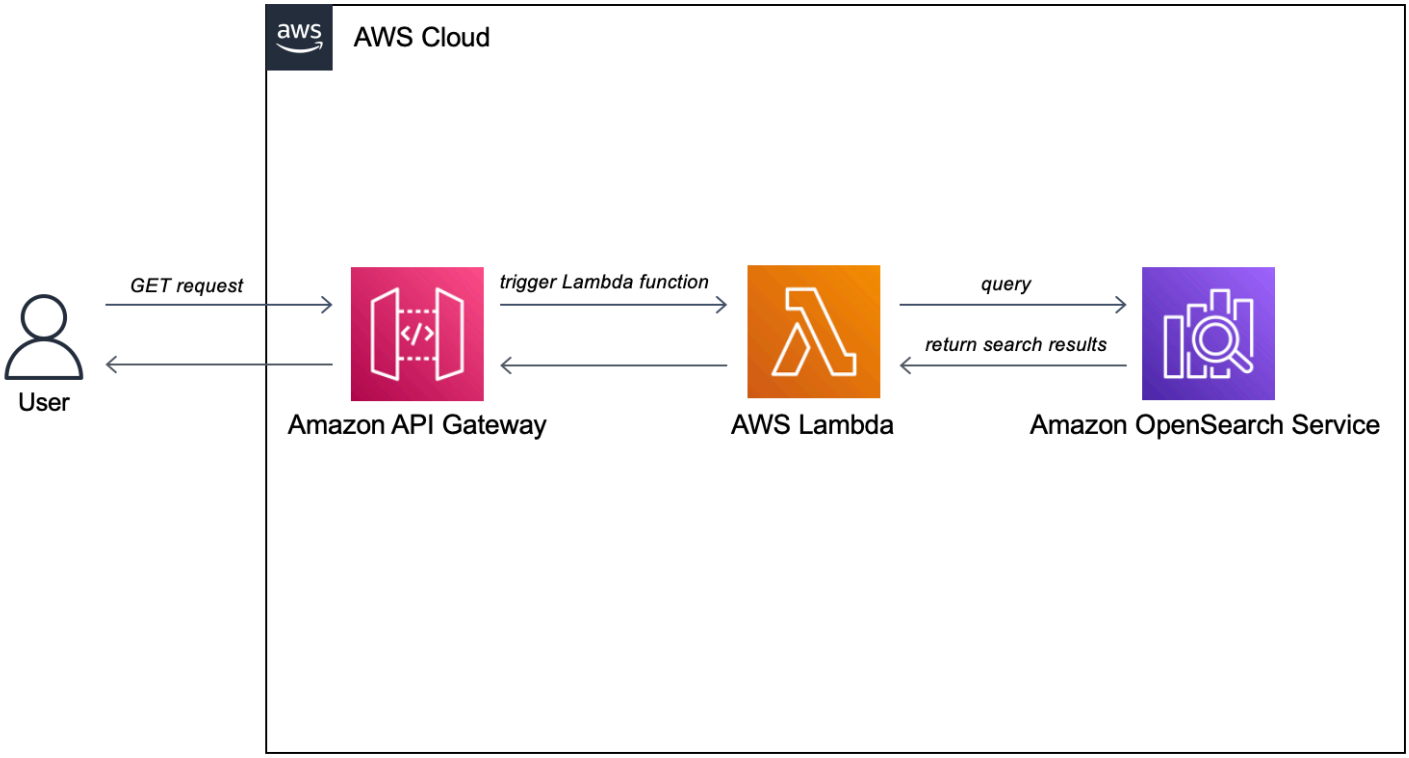
   ```
   curl -XGET -u 'master-user:master-user-password' https://domain-endpoint/_cat/
   indices?v
   ```

En este punto, la migración está completa. Puede configurar sus clientes para que usen el nuevo punto de enlace del OpenSearch servicio, cambiar el tamaño del dominio para adaptarlo a su carga de trabajo, comprobar el número de fragmentos de sus índices, cambiar a un usuario maestro de IAM o empezar a crear visualizaciones en los paneles. OpenSearch

Tutorial: Creación de una aplicación de búsqueda con Amazon OpenSearch Service

Una forma habitual de crear una aplicación de búsqueda con Amazon OpenSearch Service consiste en utilizar formularios web para enviar las consultas de los usuarios a un servidor. Luego, puedes autorizar al servidor a que lo llame OpenSearch APIs directamente y hacer que el servidor envíe las solicitudes al OpenSearch Servicio. Pero si desea escribir un código del lado del cliente que no dependa de un servidor, debe compensar los riesgos de seguridad y rendimiento. No es aconsejable permitir el acceso público y sin firma al OpenSearch APIs . Los usuarios pueden acceder a puntos de conexión no protegidos o influir en el rendimiento del clúster a través de consultas demasiado complejas (o demasiadas consultas).

En este capítulo se presenta una solución: utilice Amazon API Gateway para restringir a los usuarios a un subconjunto de ellos OpenSearch APIs y AWS Lambda para firmar las solicitudes de API Gateway al OpenSearch servicio.



> **Note**
>
> Se aplican precios estándar de API Gateway y Lambda, pero dentro del uso limitado de este tutorial, los costos deben ser insignificantes.

Requisitos previos

Un requisito previo para este tutorial es un dominio OpenSearch de servicio. Si aún no tiene uno, siga los pasos que se indican en Crear un dominio de OpenSearch servicio para crear uno.

Paso 1: Indexe los datos de muestra

Descargue sample-movies.zip, descomprímalo y utilice la operación _bulk de la API para agregar 5000 documentos al índice `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"}
...
```

Tenga en cuenta que lo anterior es un comando de ejemplo con un pequeño subconjunto de datos disponibles. Para realizar la operación `_bulk`, debe copiar y pegar todo el contenido del archivo `sample-movies`. Para obtener más instrucciones, consulte the section called “Opción 2: cargar varios documentos”.

También puede utilizar el siguiente comando curl para conseguir el mismo resultado:

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

Paso 2: crear e implementar una función de Lambda

Antes de crear su API en API Gateway, cree la función de Lambda a la que pasa las solicitudes.

Crear la función de Lambda

En esta solución, API Gateway pasa las solicitudes a una función de Lambda, que consulta el OpenSearch servicio y devuelve los resultados. Dado que esta función de ejemplo utiliza bibliotecas externas, debe crear un paquete de implementación y cargarlo en Lambda.

Creación del paquete de implementación

1. Abra un símbolo del sistema y cree un directorio del proyecto `my-opensearch-function`. Por ejemplo, en macOS:

   ```
   mkdir my-opensearch-function
   ```

2. Desplácese hasta el directorio del proyecto `my-sourcecode-function`.

   ```
   cd my-opensearch-function
   ```

3. Copie el contenido del siguiente código Python de ejemplo y guárdelo en un nuevo archivo denominado `opensearch-lambda.py`. Añada su región y el punto de conexión del host al archivo.

   ```
   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. Instale las bibliotecas externas en un nuevo directorio `package`.

```
pip3 install --target ./package boto3
pip3 install --target ./package requests
```

```
pip3 install --target ./package requests_aws4auth
```

5. Cree un paquete de implementación con las bibliotecas instaladas en la raíz. El siguiente comando genera un archivo `my-deployment-package.zip` en el directorio de su proyecto.

```
cd package
zip -r ../my-deployment-package.zip .
```

6. Agregue el archivo `opensearch-lambda.py` a la raíz del archivo zip.

```
cd ..
zip my-deployment-package.zip opensearch-lambda.py
```

Para obtener más información sobre cómo crear funciones de Lambda y paquetes de implementación, consulte Deploy Python Lambda functions with .zip file archives en la Guía para desarrolladores de AWS Lambda y the section called “Crear el paquete de implementación de Lambda” en esta guía.

Cómo crear la función usando la consola de Lambda

1. Diríjase a la consola Lambda de su https://console.aws.amazon.com/lambda/ casa. En el panel de navegación izquierdo, seleccione Funciones.
2. Seleccione Crear función.
3. Configure los siguientes campos:

   - Nombre de función: opensearch-function
   - Tiempo de ejecución: Python 3.9
   - Arquitectura: x86_64

   Mantenga todas las demás opciones predeterminadas y elija Crear función.
4. En la sección Código fuente de la página de resumen de la función, seleccione el menú desplegable Cargar desde y seleccione el archivo .zip. Busque el archivo `my-deployment-package.zip` que ha creado y seleccione Guardar.
5. El controlador es el método de su código de función que procesa eventos. En Configuración de tiempo de ejecución, seleccione Editar y cambie el nombre del controlador de acuerdo con el nombre del archivo en el paquete de implementación donde se encuentra la función de Lambda. Como el archivo se denomina `opensearch-lambda.py`, cambie el nombre del controlador

a *opensearch-lambda*.lambda_handler. Para más información, consulte Controlador de funciones de Lambda en Python.

Paso 3: crear la API en API Gateway

El uso de API Gateway le permite crear una API más limitada y simplifica el proceso de interacción con la OpenSearch _search API. API Gateway le permite habilitar las características de seguridad, como la autenticación de Amazon Cognito y la limitación controlada. Realice los siguientes pasos para crear e implementar una API:

Crear y configurar la API

Para crear una API mediante la consola de API Gateway

1. Dirígete a la consola de API Gateway en https://console.aws.amazon.com/apigateway/casa. En el panel de navegación izquierdo, selecciona APIs.
2. Localice API de REST (no privado) y elija Crear.
3. En la siguiente página, busque la sección Crear nueva API y asegúrese de que Nueva API esté seleccionada.
4. Configure los siguientes campos:
   - Nombre de la API: opensearch-api
   - Descripción: API pública para buscar un dominio OpenSearch de Amazon Service
   - Tipo de punto de conexión: Regional
5. Seleccione Crear API.
6. Seleccione Acciones y Crear método.
7. Seleccione GET (Obtener) en el menú desplegable y haga clic en la marca de verificación para confirmar.
8. Configure los siguientes ajustes y, a continuación, elija Guardar:

Opción	Valor
Tipo de integración	Función de Lambda
Usar integración de proxy de Lambda	Sí

Opción	Valor
Región de Lambda	*us-west-1*
Función de Lambda	opensearch-lambda
Usar tiempo de espera predeterminado	Sí

Configuración de la solicitud de método

Seleccione Solicitud de método y configure los siguientes ajustes:

Opción	Valor
Autorización	NONE
Validador de solicitud	Validar parámetros de cadena de consulta y encabezados
Clave de API requerida	falso

En Parámetros de cadena de consulta de URL, seleccione Añadir cadena de consulta y configure el siguiente parámetro:

Opción	Valor
Name	q
Obligatorio	Sí

Implemente la API y configure una etapa

La consola de API Gateway le permite implementar una API creando una implementación y asociándola a una etapa nueva o una existente.

1. Seleccione Acciones e Implementar API.
2. Para Etapa de implementación elija Nueva etapa y nombre la etapa `opensearch-api-test`.

3. Seleccione Implementar.
4. Configure los siguientes ajustes en el editor de etapas y, a continuación, elija Guardar los cambios:

Opción	Valor
Habilitar limitación controlada	Sí
Tarifa	1 000
Ráfagas	500

Estos ajustes configuran una API que solo tiene un método: una solicitud `GET` al punto de conexión raíz (`https://`*`some-id`*`.execute-api.`*`us-west-1`*`.amazonaws.com/search-es-api-test`). La solicitud requiere un solo parámetro (q), la cadena de consulta que se busca. Cuando se llama, el método pasa la solicitud a Lambda, que ejecuta la función `opensearch-lambda`. Para más información, consulte Creación de una API en Amazon API Gateway e Implementación de una API de REST en Amazon API Gateway.

Paso 4: (opcional) modificar la política de acceso a dominios

Su dominio de OpenSearch servicio debe permitir que la función Lambda realice `GET` solicitudes al `movies` índice. Si su dominio tiene una política de acceso abierto con control de acceso detallado habilitado, puede dejarlo tal cual:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "*"
      },
      "Action": "es:*",
      "Resource": "arn:aws:es:us-west-1:123456789012:domain/domain-name/*"
    }
```

```
  ]
}
```

Como alternativa, puede elegir hacer su política de acceso a dominio más detallada. Por ejemplo, la siguiente política mínima proporciona acceso de lectura a `opensearch-lambda-role` (creado a través de Lambda) al índice `movies`. Para obtener el nombre exacto del rol que Lambda crea automáticamente, vaya a la consola AWS Identity and Access Management (IAM), elija Roles y busque «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**
>
> Si ha habilitado un control de acceso detallado para el dominio, también debe asignar la función a un usuario en los OpenSearch paneles; de lo contrario, aparecerán errores de permisos.

Configure los permisos del rol de ejecución de Lambda

Además de configurar la política de acceso al dominio, también debe asegurarse de que la función de ejecución de Lambda tenga los permisos de IAM necesarios para acceder a su OpenSearch

dominio de servicio. La función Lambda requiere permisos específicos en función de si se utiliza un dominio gestionado o una colección de OpenSearch Service Serverless.

Para los dominios de OpenSearch servicios gestionados:

Adjunte la siguiente política de IAM a su función de ejecución de Lambda para que pueda realizar solicitudes a OpenSearch su dominio de servicio:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "es:ESHttpGet",
        "es:ESHttpPost"
      ],
      "Resource": "arn:aws:es:us-west-1:123456789012:domain/domain-name/*"
    }
  ]
}
```

Para las colecciones de OpenSearch Service Serverless:

Si utiliza OpenSearch Service Serverless, adjunte la siguiente política de IAM a su función de ejecución de Lambda:

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "aoss:*",
      "Resource": "arn:aws:aoss:us-west-1:123456789012:collection/collection-id"
    }
  ]
```

```
}
```

Para adjuntar estas políticas a su rol de ejecución de Lambda:

1. Vaya a la consola de IAM en https://console.aws.amazon.com/iam/.
2. Elija Funciones y busque su función de ejecución de Lambda (normalmente denominada `opensearch-lambda-role-xxxxxxxx`).
3. Elija Agregar permisos y luego Crear política insertada.
4. Elija la pestaña JSON y pegue la política correspondiente de arriba, sustituyendo los valores de los marcadores de posición por su recurso real. ARNs
5. Elija Revisar política, escriba un nombre como `OpenSearchAccess` y seleccione Crear política.

> **Note**
>
> Sin estos permisos de IAM, la función Lambda recibirá errores de «Acceso denegado» al intentar consultar OpenSearch el dominio del servicio, incluso si la política de acceso al dominio permite las solicitudes.

Para obtener más información sobre las políticas de acceso, consulte the section called “Configurar políticas de acceso”.

Asigne el rol de Lambda (si utiliza un control de acceso detallado)

El control de acceso detallado presenta un paso adicional antes de poder probar la aplicación. Incluso si utiliza la autenticación básica HTTP para todos los demás fines, es necesario asignar el rol de Lambda a un usuario, de lo contrario verá errores de permisos.

1. Navegue hasta la URL del OpenSearch panel de control del dominio.
2. En el menú principal, seleccione Seguridad, Roles y elija el enlace a `all_access`, el rol al que necesita asignar el rol de Lambda.
3. Seleccione Usuarios asignados, Administrar mapeo.
4. En Roles de backend, agregue el nombre de recurso de Amazon (ARN) del rol de Lambda. El ARN debe adoptar la forma de `arn:aws:iam::123456789123:role/service-role/opensearch-lambda-role-1abcdefg`.

5. Seleccione Asignar y confirme que el usuario o el rol aparecen en Usuarios asignados.

Paso 5: Pruebe la aplicación web

Para probar la aplicación web

1. Descargue sample-site.zip, descomprímalo y abra `scripts/search.js` en el editor de texto de su elección.
2. Actualice la variable `apigatewayendpoint` para que apunte al punto de conexión de API Gateway y añada una barra oblicua al final de la ruta en cuestión. Puede encontrar rápidamente el punto de conexión en API Gateway seleccionando Etapas y seleccionando el nombre de la API. La variable `apigatewayendpoint` debe adoptar la forma de `https://`*`some-id`*`.execute-api.`*`us-west-1`*`.amazonaws.com/opensearch-api-test/`.
3. Abra `index.html` e intente ejecutar búsquedas de thor, house y algunos otros términos.

Movie Search

thor

Found 7 results.



Thor

2011 — 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.



Thor: The Dark World

2013 — Faced with an enemy that even Odin and Asgard cannot withstand, Thor must embark on his most perilous and personal journey yet, one that will reunite him with Jane Foster and force him to sacrifice everything to save us all.



Vikingdom

2013 — A forgotten king, Eirick, is tasked with the impossible odds to defeat Thor, the God of Thunder.

Solucionar errores de CORS

Aunque la función de Lambda incluye contenido en la respuesta para admitir CORS, es posible que aparezca el siguiente error:

```
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.
```

Si esto sucede, pruebe lo siguiente:

1. Habilitar CORS en el recurso GET. En Avanzado, configure Access-Control-Allow-Credentials en `'true'`.
2. Vuelva a implementar su API en API Gateway (Acciones, Implementar API).
3. Elimine y vuelva a agregar el activador de la función de Lambda. Agregue volver a añadirlo, seleccione Agregar desencadenador y cree el punto de conexión HTTP que invoca su función. El desencadenador debe tener la siguiente configuración:

Desencadenador	API	Etapa de implementación	Seguridad
API Gateway	opensearch-api	opensearch-api-test	Abra

Siguientes pasos

Este capítulo es solo un punto de partida para demostrar un concepto. Podría considerar las siguientes modificaciones:

- Agregue sus propios datos al dominio del OpenSearch servicio.
- Agregue métodos a su API.
- En la función de Lambda, modifique la consulta de búsqueda o aumente los diferentes campos.
- Aplique un estilo diferente a los resultados o modifique `search.js` para mostrar diferentes campos al usuario.

Tutorial: Visualización de las llamadas de atención al cliente con el OpenSearch servicio y los paneles OpenSearch

Este capítulo es un tutorial completo de la siguiente situación: el servicio de atención al cliente de una empresa recibe llamadas y quiere analizarlas. ¿Cuál es el tema de cada llamada? ¿Cuántas fueron positivas? ¿Cuántas negativas? ¿Cómo los gerentes pueden buscar o revisar las transcripciones de estas llamadas?

Un flujo de trabajo manual puede implicar que los empleados escuchen las grabaciones, apunten el tema de cada llamada y decidan si la interacción con el cliente fue positiva.

Dicho proceso requeriría mucho esfuerzo y trabajo. Suponiendo un tiempo medio de 10 minutos por llamada, cada empleado podría escuchar solo 48 llamadas al día. Salvo sesgo humano, los datos que generan sería de alta precisión, pero la cantidad de datos sería mínima: tan solo el tema de la llamada y un valor booleano para determinar si el cliente quedó satisfecho. Cualquier información más compleja, como, por ejemplo, una transcripción completa, requeriría gran cantidad de tiempo.

Con Amazon S3, Amazon Transcribe, Amazon Comprehend y OpenSearch Amazon Service, puede automatizar un proceso similar con muy poco código y obtener muchos más datos. Por ejemplo, puede obtener una transcripción completa de la llamada, palabras clave de la transcripción, y un "sentimiento" general de la llamada (positiva, negativa, neutra o mixta). Luego, puede usar OpenSearch los OpenSearch paneles de control para buscar y visualizar los datos.

Si bien puedes usar este tutorial tal como está, la intención es generar ideas sobre cómo enriquecer tus documentos JSON antes de indexarlos en Service. OpenSearch

Costo estimado:

En general, seguir los pasos que se indican en este tutorial debería costar menos de 2 USD. En este tutorial se utilizan los siguientes recursos:

- bucket de S3 con menos de 100 MB transferidos y almacenados

  Para obtener más información, consulte Precios de Amazon S3.
- OpenSearch Dominio de servicio con una `t2.medium` instancia y 10 GiB de almacenamiento de EBS durante varias horas

  Para obtener más información, consulta los precios OpenSearch de Amazon Service.
- Varias llamadas a Amazon Transcribe

Para obtener más información, consulte Precios de Amazon Transcribe.

- Varias llamadas de procesamiento de lenguaje natural a Amazon Comprehend

  Para obtener más información, consulte Precios de Amazon Comprehend.

Temas

- Paso 1: configure los requisitos previos
- Paso 2: copie el código de muestra
- (Opcional) Paso 3: indexe los datos de ejemplo
- Paso 4: analice y visualice sus datos
- Paso 5: elimine recursos y pasos siguientes

Paso 1: configure los requisitos previos

Antes de continuar, debe contar con los siguientes recursos.

Requisito previo	Description (Descripción)
Bucket de Amazon S3	Para obtener más información, consulte Creación de un bucket en la Guía del usuario de Amazon Simple Storage Service.
OpenSearch Dominio de servicio	El destino de los datos. Para obtener más información, consulte Creación de dominios OpenSearch de servicio.

Si aún no tiene estos recursos, puede crearlos utilizando los siguientes comandos de la AWS CLI :

```
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**
>
> Estos comandos utilizan la región `us-west-2`, pero puede utilizar cualquier región que admita Amazon Comprehend. Para obtener más información, consulte Referencia general de AWS.

Paso 2: copie el código de muestra

1. Copie y pegue el siguiente código de muestra de Python 3 en un archivo nuevo denominado `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. Actualice las seis variables iniciales.
3. Instale los paquetes necesarios mediante los siguientes comandos:

```
pip install boto3
pip install requests
pip install requests_aws4auth
```

4. Colóquelo MP3 en el mismo directorio `call-center.py` y ejecute el script. A continuación se indica un ejemplo de salida:

```
$ 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` realiza una serie de operaciones:

1. El script carga un archivo de audio (en este caso, un MP3, pero Amazon Transcribe admite varios formatos) a su bucket de S3.
2. Envía la URL del archivo de audio a Amazon Transcribe y espera que realice la tarea de transcripción.

   El tiempo necesario para terminar el trabajo de transcripción depende de la duración del archivo de audio. Supongamos que tomará minutos, no segundos.

   > **Tip**
   >
   > Para mejorar la calidad de la transcripción, puede configurar un vocabulario personalizado para Amazon Transcribe.

3. Una vez que el trabajo de transcripción termina, el script extrae la transcripción, la recorta a 5 000 caracteres y envía a Amazon Comprehend para identificar palabras clave y analizar el sentimiento.
4. Por último, el script agrega la transcripción completa, las palabras clave, el sentimiento y la marca de tiempo actual a un documento JSON y lo indexa en Service. OpenSearch

> **Tip**
>
> LibriVoxtiene audiolibros de dominio público que puede usar para realizar pruebas.

(Opcional) Paso 3: indexe los datos de ejemplo

Si no tiene a mano varias grabaciones de llamadas, y no es algo muy frecuente, puede indexar los documentos de muestra en sample-calls.zip, que son comparables a lo que produce `call-center.py`.

1. Cree un archivo denominado `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. Actualice las dos variables iniciales para `host` y `region`.
3. Instale el paquete necesario mediante el siguiente comando:

```
pip install opensearch-py
```

4. Descargue y descomprima sample-calls.zip.
5. Guarde `sample-calls.bulk` en el mismo directorio que `bulk-helper.py` y ejecute el ayudante. A continuación se indica un ejemplo de salida:

```
$ python bulk-helper.py
{
  "errors": false,
  "items": [
    {
      "index": {
        "_id": "1",
        "_index": "support-calls",
        "_primary_term": 1,
        "_seq_no": 42,
        "_shards": {
          "failed": 0,
          "successful": 1,
```

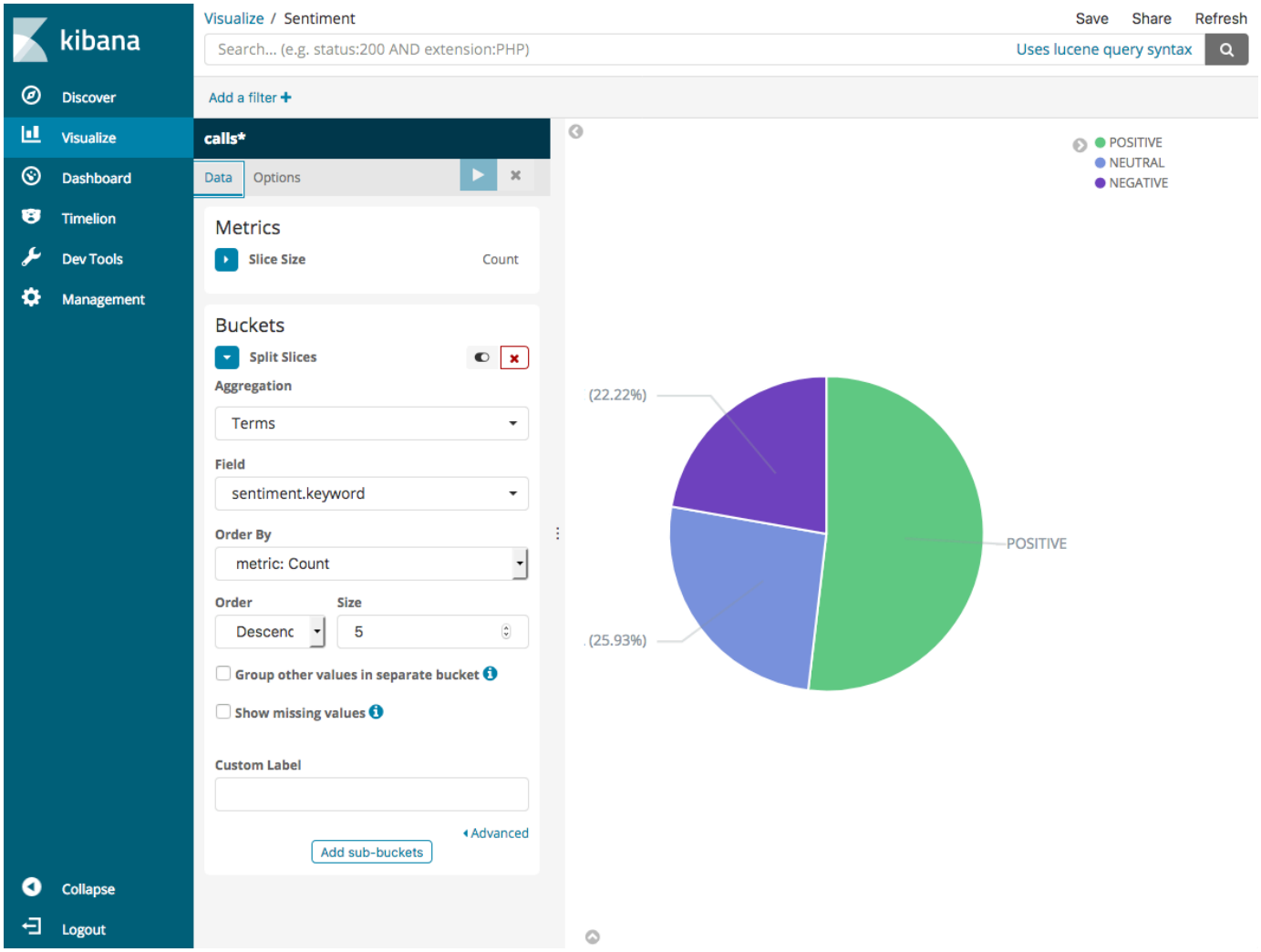
```
          "total": 2
        },
        "_type": "_doc",
        "_version": 9,
        "result": "updated",
        "status": 200
      }
    },
    ...
  ],
  "took": 27
}
```

Paso 4: analice y visualice sus datos

Ahora que tiene algunos datos en OpenSearch Service, puede visualizarlos mediante los OpenSearch paneles de control.

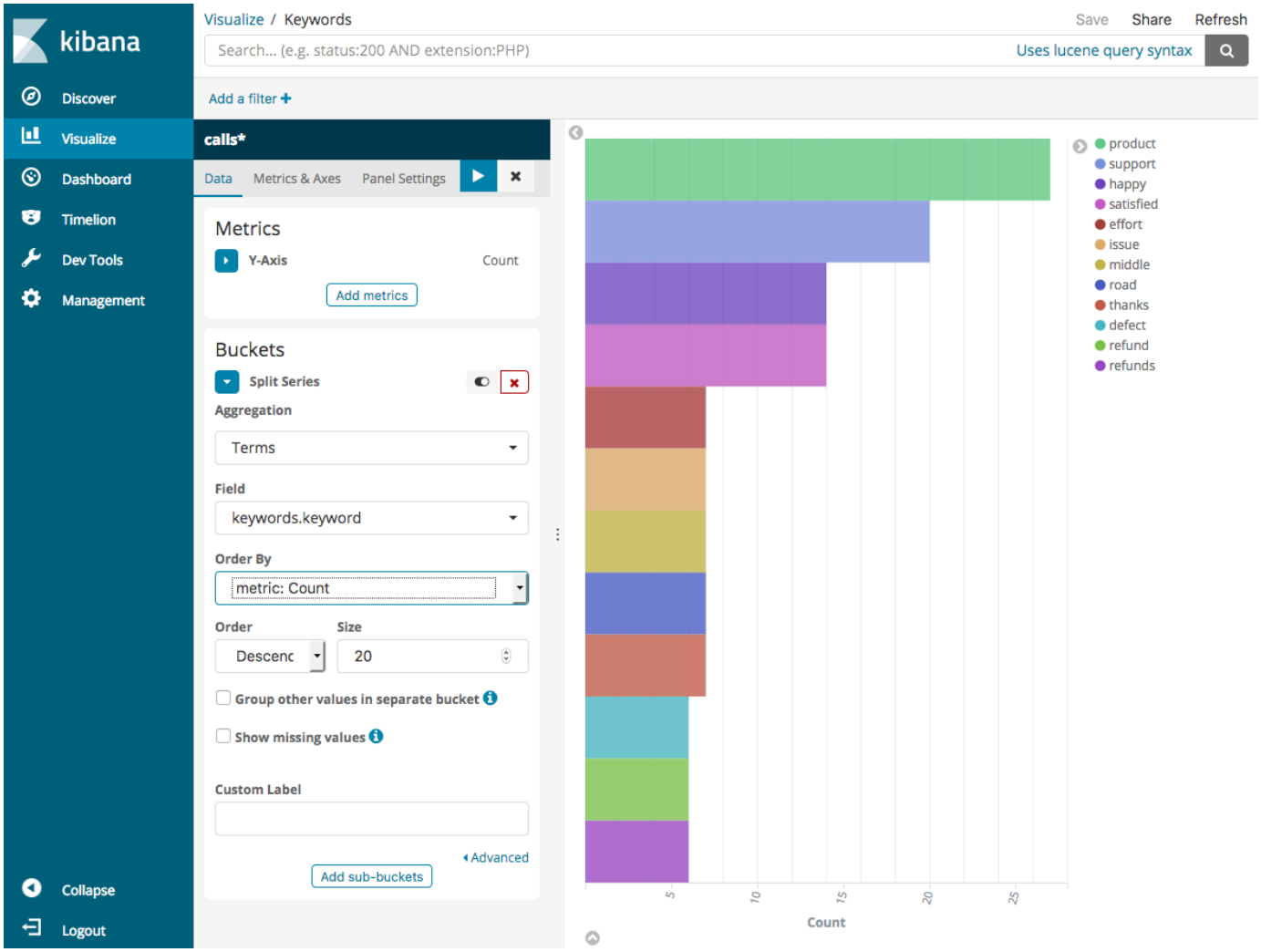
1. Vaya a `https://search-`*`domain`*`.`*`region`*`.es.amazonaws.com/_dashboards`.
2. Para poder usar los OpenSearch paneles, necesita un patrón de índice. Dashboards utiliza patrones de índice para limitar su análisis a uno o más índices. Para que coincida con el `support-calls`índice que`call-center.py` cree, diríjase a Pila de administración, Patrones, y defina un patrón de índice de `support*`. y luego elija Paso siguiente.
3. En Time Filter field name (Nombre de campo Filtro de tiempo), elija timestamp (marca temporal).
4. Ahora pude comenzar a crear visualizaciones. Seleccione Visualize (Visualizar) y añada una nueva visualización.
5. Seleccione el gráfico circular y el patrón de índice `support*`.
6. El valor predeterminado de visualización es básico. Si quiere crear una visualización más interesante, elija Split Slices (Sectores divididos).

   Para Aggregation (Agregación), elija Terms (Términos). En Field (Campo), elija sentiment.keyword. Seleccione Apply changes (Aplicar cambios) y Save (Guardar).

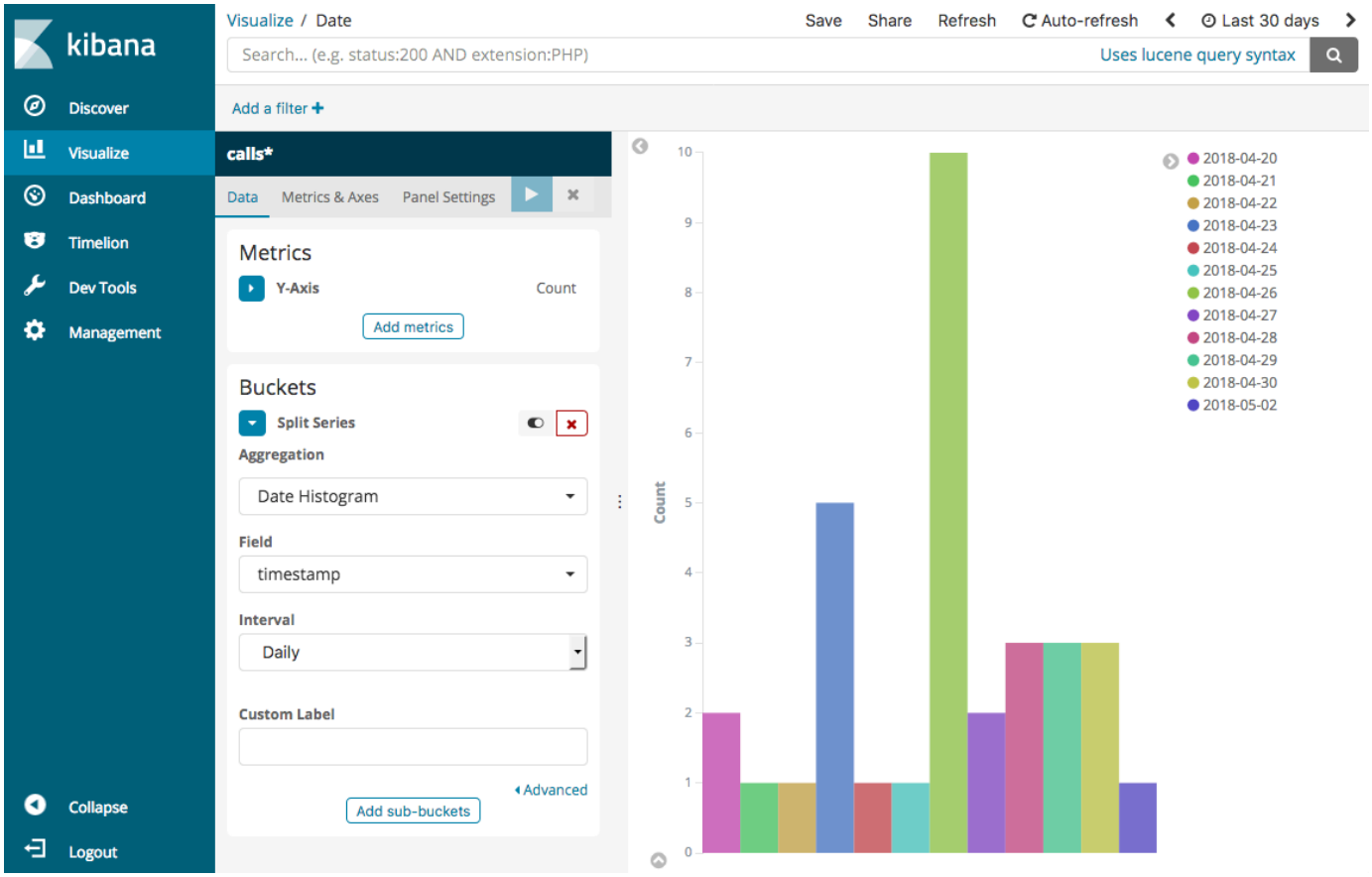


7. Vuelva a la página Visualizar y agregue otra visualización. Esta vez, elija el gráfico de barras horizontales.
8. Seleccione Split Series (Series divididas).

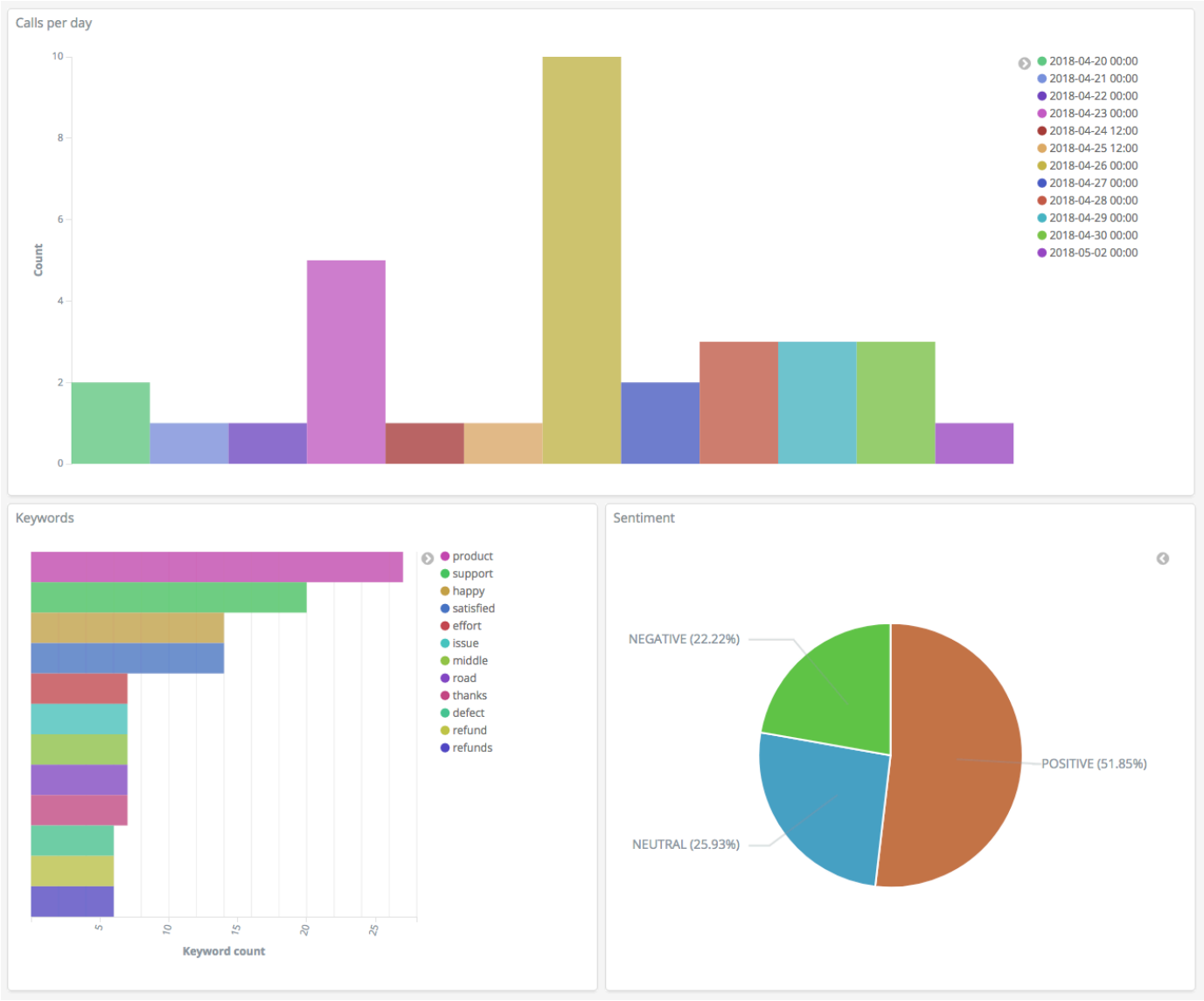
   Para Aggregation (Agregación), elija Terms (Términos). En Field (Campo), elija keywords.keyword y cambie el valor de Size (Tamaño) a 20. Seleccione Apply Changes (Aplicar cambios) y Save (Guardar).



9. Vuelva a la página Visualizar y agregue una última visualización, un gráfico de barras verticales.
10. Seleccione Split Series (Series divididas). En Aggregation (Agregación), elija Date Histogram (Histograma de fecha). En Field (Campo), elija timestamp (marca temporal) y cambie Interval (Intervalo) a Daily (Diario).
11. Seleccione Metrics & Axes (Métricas y ejes) y cambie Mode (Modo) a normal.
12. Seleccione Apply Changes (Aplicar cambios) y Save (Guardar).



13. Ahora que tiene tres visualizaciones, puede añadirlas a un panel de Dashboards. Seleccione Dashboard (Panel), cree un panel y añada sus visualizaciones.



Paso 5: elimine recursos y pasos siguientes

Para evitar cargos innecesarios, elimine el bucket de S3 y el dominio OpenSearch de servicio. Para obtener más información, consulta Eliminar un bucket en la Guía del usuario de Amazon Simple Storage Service y Eliminar un dominio de OpenSearch servicio en esta guía.

Las transcripciones requieren mucho menos espacio en disco que MP3 los archivos. Es posible que pueda acortar el período de MP3 retención (por ejemplo, de tres meses de grabaciones de llamadas a un mes), conservar años de transcripciones y, aun así, ahorrar en costos de almacenamiento.

También puede automatizar el proceso de transcripción con AWS Step Functions Lambda, añadir metadatos adicionales antes de la indexación o crear visualizaciones más complejas que se adapten a su caso de uso exacto.

Cambio OpenSearch de nombre de Amazon Service: resumen de los cambios

El 8 de septiembre de 2021, nuestro paquete de búsquedas y análisis pasó a llamarse Amazon OpenSearch Service. OpenSearch El servicio es compatible con OpenSearch el sistema operativo Elasticsearch anterior, así como con el sistema operativo anterior. En las secciones siguientes, se describen las diferentes partes del servicio que sufrieron modificaciones con el cambio de nombre y las acciones que se deben realizar para garantizar que los dominios sigan funcionando correctamente.

Algunos de estos cambios solo se aplican cuando actualizas tus dominios de Elasticsearch a. OpenSearch En otros casos, como en la consola de Billing and Cost Management, la experiencia cambia inmediatamente.

Tenga en cuenta que esta lista no es exhaustiva. Mientras que otras partes del producto también cambiaron, estas actualizaciones son las más relevantes.

Nueva versión de la API

La nueva versión de la API de configuración del OpenSearch servicio (01/01/2021) funciona OpenSearch igual de bien con el OSS de Elasticsearch heredado. Se sustituyeron 21 operaciones de la API por nombres más concisos e independientes del motor (por ejemplo, se `CreateElasticsearchDomain` cambiaron a`CreateDomain`), pero OpenSearch Service sigue siendo compatible con ambas versiones de la API.

Recomendamos utilizar las nuevas operaciones de la API para crear y administrar dominios en el futuro. Tenga en cuenta que cuando utilice las nuevas operaciones de la API para crear un dominio, debe especificar el parámetro `EngineVersion` en el formato `Elasticsearch_X.Y` o `OpenSearch_X.Y`, en lugar de solo el número de versión. Si no especificas una versión, se utilizará de forma predeterminada la versión más reciente de. OpenSearch

Actualice AWS CLI la suya a la versión 1.20.40 o posterior para utilizarla `aws opensearch ...` para crear y gestionar sus dominios. Para conocer el nuevo formato de CLI, consulte la referencia de OpenSearch CLI.

Tipos de instancia con cambio de nombre

Los tipos de instancia de Amazon OpenSearch Service ahora tienen el formato, `<type>.<size>.search` por ejemplo, `m6g.large.search` en lugar de `m6g.large.elasticsearch`. No es necesario realizar ninguna acción. Los dominios existentes comenzarán a hacer referencia automáticamente a los nuevos tipos de instancias dentro de la API y en la consola de Billing and Cost Management.

Si tienes instancias reservadas (RIs), el cambio no afectará a tu contrato. La versión antigua de la API de configuración aún es compatible con el formato de nomenclatura anterior, pero si desea utilizar la nueva versión de la API, debe utilizar el nuevo formato.

Cambios en las políticas de acceso

En las secciones siguientes, se describen las acciones que debe realizar para actualizar las políticas de acceso.

Políticas de IAM

Recomendamos que actualice sus Políticas de IAM para utilizar las operaciones de la API con cambio de nombre. Sin embargo, OpenSearch Service seguirá respetando las políticas existentes mediante la replicación interna de los permisos de API anteriores. Por ejemplo, si actualmente tiene permiso para realizar la operación `CreateElasticsearchDomain`, ahora puede realizar llamadas a ambos `CreateElasticsearchDomain` (operación antigua de la API) y `CreateDomain` (nueva operación de la API). Lo mismo se aplica a las denegaciones explícitas. Para obtener una lista de las operaciones de la API actualizadas, consulte la referencia de elementos de política.

Políticas de SCP

Las políticas de control de servicios (SCPs) introducen un nivel adicional de complejidad en comparación con la IAM estándar. Para evitar que las políticas de SCP se rompan, debe agregar tanto las operaciones de la API antiguas como las nuevas a cada una de sus políticas de SCP. Por ejemplo, si actualmente un usuario tiene permisos para `CreateElasticsearchDomain`, también debe concederles permisos para `CreateDomain` a fin de que puedan retener la capacidad de crear dominios. Lo mismo se aplica a las denegaciones explícitas.

Por ejemplo:

```
"Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "es:CreateElasticsearchDomain",
        "es:CreateDomain"
         ...
      ],
    },
      "Effect": "Deny",
      "Action:" [
        "es:DeleteElasticsearchDomain",
        "es:DeleteDomain"
         ...
```

Nuevos tipos de recursos

OpenSearch El servicio presenta los siguientes tipos de recursos nuevos:

Recurso	Description (Descripción)
`AWS::OpenSearchService::Domain`	Representa un dominio OpenSearch de Amazon Service. Este recurso existe en el nivel de servicio y no es específico del software que se ejecuta en el dominio. Se aplica a servicios como CloudFormation y AWS Resource Groups, en los que crea y administra recursos para el servicio en su conjunto. Para obtener instrucciones sobre cómo actualizar los dominios definidos en CloudForm ation Elasticsearch a OpenSearch, consulte las observaciones de la guía del CloudFormation usuario.
`AWS::OpenSearch::Domain`	Representa el OpenSearch/Elasticsearch software que se ejecuta en un dominio. Este recurso se aplica a servicios como AWS CloudTraily AWS Config, que hacen referenci

Recurso	Description (Descripción)
	a al software que se ejecuta en el dominio y no OpenSearch al servicio en su conjunto. Estos servicios ahora contienen tipos de recursos independientes para los dominios que ejecutan Elasticsearch (`AWS::Elasticsearch ::Domain`) y los dominios que ejecutan OpenSearch (`AWS::OpenSearch::D omain`).

> **Note**
>
> En AWS Config, seguirás viendo tus datos con el tipo de `AWS::Elasticsearch::Domain` recurso existente durante varias semanas, incluso si actualizas uno o más dominios a ellos. OpenSearch

Kibana ha cambiado su nombre a Dashboards OpenSearch

OpenSearch Los paneles, la AWS alternativa a Kibana, son una herramienta de visualización de código abierto diseñada para funcionar con ella. OpenSearch Después de actualizar un dominio de Elasticsearch a OpenSearch, el punto final cambia a. `/_plugin/kibana` `/_dashboards` OpenSearch El servicio redirigirá todas las solicitudes al nuevo punto final, pero si utilizas el punto final de Kibana en alguna de tus políticas de IAM, actualiza esas políticas para incluir también el nuevo `/_dashboards` punto final.

Si lo usasthe section called “Autenticación SAML para paneles OpenSearch”, antes de actualizar tu dominio a OpenSearch, debes cambiar todos los Kibana URLs configurados en tu proveedor de identidad (IdP) `/_plugin/kibana` de a. `/_dashboards` Los más comunes URLs son el servicio de confirmación al consumidor (ACS) URLs y el destinatario. URLs

Se cambió el nombre del `kibana_read_only` rol predeterminado de los OpenSearch paneles a `opensearch_dashboards_read_only` y el `kibana_user` rol a. `opensearch_dashboards_user` El cambio se aplica a todos los 1 recién creados OpenSearch . x dominios que ejecutan el software de servicio R20211203 o posterior. Si se actualiza un dominio existente al software de servicio R20211203, los nombres de los roles siguen siendo los mismos.

CloudWatch Métricas renombradas

Varias CloudWatch métricas cambian para los dominios en ejecución OpenSearch. Cuando actualizas un dominio a OpenSearch, las métricas cambian automáticamente y CloudWatch las alarmas actuales no funcionan. Antes de actualizar el clúster de una versión de Elasticsearch a una OpenSearch versión, asegúrate de actualizar CloudWatch las alarmas para usar las nuevas métricas.

Cambiaron las siguientes métricas:

Nombre de métrica original	Nombre nuevo
KibanaHealthyNodes	OpenSearchDashboardsHealthyNodes
KibanaConcurrentConnections	OpenSearchDashboardsConcurrentConnections
KibanaHeapTotal	OpenSearchDashboardsHeapTotal
KibanaHeapUsed	OpenSearchDashboardsHeapUsed
KibanaHeapUtilization	OpenSearchDashboardsHeapUtilization
KibanaOS1MinuteLoad	OpenSearchDashboardsOS1MinuteLoad
KibanaRequestTotal	OpenSearchDashboardsRequestTotal
KibanaResponseTimesMaxInMillis	OpenSearchDashboardsResponseTimesMaxInMillis
ESReportingFailedRequestSysErrCount	KibanaReportingFailedRequestSysErrCount
ESReportingRequestCount	KibanaReportingRequestCount
ESReportingFailedRequestUserErrCount	KibanaReportingFailedRequestUserErrCount

Nombre de métrica original	Nombre nuevo
`ESReportingSuccessCount`	`KibanaReportingSuccessCount`
`ElasticsearchRequests`	`OpenSearchRequests`

Para ver una lista completa de las métricas que OpenSearch Service envía a Amazon CloudWatch, consultathe section called “Monitoreo de métricas del clúster”.

Cambios en la consola de Billing and Cost Management

Los datos históricos de la consola de Billing and Cost Management y de los informes de costo y uso seguirán utilizando el nombre de servicio anterior, por lo que tendrás que empezar a usar filtros tanto para Amazon OpenSearch Service como para el nombre heredado de Elasticsearch al buscar datos. Si ya tienes informes guardados, actualiza los filtros para asegurarte de que también incluyan OpenSearch el servicio. Al principio, es posible que recibas una alerta cuando tu uso de Elasticsearch disminuya y aumente OpenSearch, pero desaparecerá al cabo de unos días.

Además del nombre del servicio, los siguientes campos cambiarán para todos los informes, facturas y operaciones de la API de lista de precios:

Campo	Formato antiguo	Formato nuevo
Tipo de instancia	`m5.large.elasticsearch`	`m5.large.search`
Familia de productos	Instancia de Elasticsearch Volumen de Elasticsearch	Instancia OpenSearch de servicio de Amazon Volumen de OpenSearch servicios de Amazon
Descripción del precio	5,098 USD por hora de instancia c5.18xlarge.elasticsearch (u hora parcial), UE	5,098 USD por hora de instancia c5.18xlarge.search (u hora parcial), UE
Familia de instancias	`ultrawarm.elasticsearch`	`ultrawarm.search`

Nuevo formato de evento

El formato de los eventos que OpenSearch Service envía a Amazon EventBridge y Amazon CloudWatch ha cambiado, específicamente el `detail-type` campo. El campo fuente (`aws.es`) aún es el mismo. Para obtener el formato completo de cada tipo de evento, consulte the section called "Supervisión de eventos". Si tiene reglas de evento existentes que dependen del formato antiguo, asegúrese de actualizarlas para que se ajusten al nuevo formato.

¿Qué permanece igual?

Las siguientes características y funcionalidades, entre otras no enumeradas, permanecerán iguales:

- Entidad principal de servicio (`es.amazonaws.com`)
- Código de proveedor
- Dominio ARNs
- Puntos de conexión del dominio

Comience: actualice sus dominios a OpenSearch 1.x

OpenSearch 1. x admite actualizaciones desde las versiones 6.8 y 7 de Elasticsearch. x. Para obtener instrucciones a fin de actualizar su dominio, consulte the section called "Actualización de un dominio (consola)". Si utilizas la API de configuración AWS CLI o para actualizar tu dominio, debes especificar el `TargetVersion` as`OpenSearch_1.x`.

OpenSearch 1. x introduce una configuración de dominio adicional denominada Activar el modo de compatibilidad. Debido a que algunos clientes y complementos de Elasticsearch OSS comprueban la versión del clúster antes de conectarse, el modo de compatibilidad establece OpenSearch que su versión es 7.10 para que estos clientes sigan funcionando.

Puedes habilitar el modo de compatibilidad al crear OpenSearch dominios por primera vez o al actualizar OpenSearch desde una versión de Elasticsearch. Si no está establecido, el parámetro predeterminado es `false` al crear un dominio y `true` cuando actualiza un dominio.

Para habilitar el modo de compatibilidad mediante la API de configuración, establezca `override_main_response_version` en `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"
   }
}
```

Para habilitar o deshabilitar el modo de compatibilidad en los OpenSearch dominios existentes, debes usar la operación de la API OpenSearch _cluster/settings:

```
PUT /_cluster/settings
{
  "persistent" : {
    "compatibility.override_main_response_version" : true
  }
}
```

Solución de problemas de Amazon OpenSearch Service

En este tema se describe cómo identificar y resolver problemas comunes OpenSearch de Amazon Service. Consulte la información de esta sección antes de ponerse en contacto con el Soporte de AWS.

¿No puedes acceder a los OpenSearch paneles

El punto final de OpenSearch Dashboards no admite solicitudes firmadas. Si la política de control de acceso del dominio solo concede acceso a determinados roles de IAM y no se ha configurado la autenticación de Amazon Cognito, es posible que aparezca el siguiente error cuando se intente acceder a Dashboards:

```
"User: anonymous is not authorized to perform: es:ESHttpGet"
```

Si tu dominio de OpenSearch servicio usa el acceso a la VPC, es posible que no recibas este error, pero puede que se agote el tiempo de espera de la solicitud. Para obtener más información acerca de cómo solucionar este problema y las distintas opciones de configuración disponibles the section called “Control del acceso a Dashboards”, consulte the section called “Acerca de las políticas de acceso en los dominios de VPC” y the section called “Gestión de identidad y acceso”.

No se puede obtener acceso al dominio de VPC

Consulte the section called “Acerca de las políticas de acceso en los dominios de VPC” y the section called “Probar los dominios de VPC”.

Clúster en estado de solo lectura

En comparación con las versiones anteriores de Elasticsearch OpenSearch y Elasticsearch 7. x utilizan un sistema diferente para la coordinación de clústeres. En este nuevo sistema, cuando el clúster pierde quórum, el clúster no estará disponible hasta que realice alguna acción. La pérdida de quórum puede adoptar dos formas:

- Si el clúster utiliza nodos principales dedicados, la pérdida de cuórum se produce cuando la mitad o más no están disponibles.

- Si el clúster no utiliza nodos principales dedicados, la pérdida de cuórum se produce cuando la mitad o más de los nodos de datos no están disponibles.

Si se produce una pérdida de quórum y el clúster tiene más de un nodo, OpenSearch Service restaura el quórum y coloca el clúster en un estado de solo lectura. Tiene dos opciones:

- Elimine el estado de solo lectura y utilice el clúster tal y como está.
- Restaure el clúster o los índices individuales a partir de una instantánea.

Si prefiere utilizar el clúster tal y como está, verifique que el estado del clúster sea verde mediante la siguiente solicitud:

```
GET _cat/health?v
```

Si el estado del clúster es rojo, le recomendamos que restaure el clúster a partir de una instantánea. También puede consultar the section called “Estado rojo del clúster” para ver los pasos de solución de problemas. Si el estado del clúster es verde, compruebe que todos los índices esperados estén presentes utilizando la siguiente solicitud:

```
GET _cat/indices?v
```

A continuación, ejecute algunas búsquedas para verificar que los datos esperados están presentes. Si es, puede eliminar el estado de solo lectura utilizando la siguiente solicitud:

```
PUT _cluster/settings
{
  "persistent": {
    "cluster.blocks.read_only": false
  }
}
```

Si se produce una pérdida de quórum y el clúster solo tiene un nodo, OpenSearch Service reemplaza el nodo y no coloca el clúster en un estado de solo lectura. De lo contrario, las opciones son las mismas: utilice el clúster tal y como está o restaure a partir de una instantánea.

En ambas situaciones, OpenSearch Service le envía dos eventos. Panel de estado El primero le informa de la pérdida de cuórum. El segundo ocurre después de que OpenSearch Service

restablezca correctamente el quórum. Para obtener más información sobre el uso del Panel de estado, consulte la Guía del AWS Health usuario.

Estado rojo del clúster

Un estado de clúster rojo significa que al menos un fragmento principal y sus réplicas no están asignadas a un nodo. OpenSearch El servicio sigue intentando realizar instantáneas automatizadas de todos los índices, independientemente de su estado, pero las instantáneas fallan mientras persiste el estado del clúster rojo.

Las causas más comunes de un estado rojo del clúster son que los nodos del clúster han fallado y que el proceso OpenSearch se ha bloqueado debido a una carga de procesamiento elevada continua.

> **Note**
>
> OpenSearch El servicio almacena las instantáneas automatizadas durante 14 días, independientemente del estado del clúster. Por lo tanto, si el estado del clúster rojo persiste durante más de dos semanas, se eliminará la última instantánea automatizada correcta y podría perder permanentemente los datos del clúster. Si su dominio de OpenSearch servicio pasa a tener un estado de clúster rojo, Soporte es posible que se ponga en contacto con usted para preguntarle si desea solucionar el problema usted mismo o si desea que el equipo de soporte lo ayude. Puedes configurar una CloudWatch alarma para que te avise cuando aparezca un estado de clúster rojo.

En última instancia, las particiones rojas causan clústeres rojos y los índices rojos causan particiones rojas. OpenSearch Tiene algunos indicadores útiles APIs para identificar los índices que causan el estado de conglomerado rojo.

- `GET /_cluster/allocation/explain` elige la primera partición no asignada que encuentra y explica por qué no se puede asignar 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` muestra el estado, el número de documentos y el uso de disco de cada índice:

```
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
```

Eliminar los índices rojos es la forma más rápida de solucionar el estado rojo del clúster. En función del motivo por el que el clúster esté en rojo, puede escalar su dominio de OpenSearch servicio para utilizar tipos de instancias más grandes, más instancias o más almacenamiento basado en EBS e intentar recrear los índices problemáticos.

Si eliminar un índice problemático no es factible, puede restaurar una instantánea, eliminar documentos del índice, cambiar la configuración del índice, reducir el número de réplicas o eliminar otros índices para liberar espacio en el disco. El paso importante es resolver el estado del clúster rojo antes de volver a configurar el dominio de servicio. OpenSearch Volver a configurar un dominio con un estado rojo del clúster puede agravar el problema y el dominio podría quedarse bloqueado en el estado de configuración Processing (Procesando) hasta que se resuelva el estado.

Corrección automática de clústeres en rojo

Si el estado del clúster está en rojo continuo durante más de una hora, OpenSearch Service intentará corregirlo automáticamente redireccionando los fragmentos no asignados o restaurándolos a partir de instantáneas anteriores.

Si no corrige uno o más índices rojos y el estado del clúster permanece en rojo durante un total de 14 días, OpenSearch Service solo tomará medidas adicionales si el clúster cumple al menos uno de los siguientes criterios:

- Tiene solo una zona de disponibilidad
- No tiene nodos maestros dedicados
- Contiene tipos de instancias de ráfaga (T2 o T3)

En este momento, si su clúster cumple uno de estos criterios, OpenSearch Service le enviará notificaciones diarias durante los próximos 7 días explicándole que si no corrige estos índices, se eliminarán todos los fragmentos no asignados. Si el estado del clúster sigue siendo rojo después de 21 días, OpenSearch Service eliminará los fragmentos no asignados (almacenamiento y procesamiento) de todos los índices rojos. Recibirá notificaciones en el panel de notificaciones de la consola de OpenSearch servicio para cada uno de estos eventos. Para obtener más información, consulte the section called “Eventos del estado del clúster”.

Recuperación de una carga de procesamiento elevada continua

Para determinar si el estado rojo de un clúster se debe a una carga de procesamiento elevada continua en un nodo de datos, monitorice las siguientes métricas de clúster.

Métrica relevante	Description (Descripción)	Recuperación
JVMMemoryPresión	Especifica el porcentaje del montón de Java que se emplea para todos los nodos de datos de un clúster. Vea la estadística Máximo de esta métrica y busque caídas cada vez más pequeñas en la presión de memoria a medida que el recolector de elementos no utilizados de Java no consigue reclamar suficiente memoria. Probablemente, este patrón se deba a consultas complejas o a campos de datos de gran tamaño. Los tipos de instancias x86 utilizan el recolector de basura Concurrent Mark Sweep (CMS), que se ejecuta junto a subprocesos de aplicaciones	Establezca interruptores de memoria para la JVM. Para obtener más información, consulte the section called “JVM OutOfMemoryError”. Si el problema persiste, elimine los índices innecesarios, reduzca el número o la complejidad de las solicitudes al dominio, agregue instancias o utilice tipos de instancia más grandes.

Métrica relevante	Description (Descripción)	Recuperación
	para que las pausas sigan siendo breves. Si CMS no puede obtener memoria suficiente durante sus recolecciones normales, desencadena una recolección de basura completa, lo que puede provocar pausas prolongadas de las aplicaciones y afectar a la estabilidad del clúster. Los tipos de instancias Graviton basados en ARM utilizan el recolector de basura Garbage-First (G1), que es similar a CMS, pero utiliza pausas breves adicionales y desfragmentación del montón para reducir aún más la necesidad de recolecciones de basura completas. En cualquier caso, si el uso de memoria sigue aumentando más de lo que el recolector de basura puede recuperar durante la recolección completa de elementos no utilizados, OpenSearch se bloquea y se produce un error de memoria insuficiente. En todos los tipos de instancias, una buena regla es mantener el uso por debajo del 80 %. La API `_nodes/stats/jvm` ofrece un resumen útil sobre estadísticas de JVM, el uso del grupo de memoria e información de recopilación de elementos no utilizados:	

Métrica relevante	Description (Descripción)	Recuperación
	`GET domain-endpoint /_nodes/stats/jvm?pretty`	
CPUUtilization	Especifica el porcentaje de recursos de CPU usados para los nodos de datos de un clúster. Consulte la estadística Máximo para esta métrica y busque un patrón continuo de uso elevado.	Agregue nodos de datos o aumente el tamaño de los tipos de instancia de los nodos de datos existentes.
Nodos	Especifica el número de nodos de un clúster. Vea la estadística Mínimo para esta métrica. Este valor fluctúa cuando el servicio implementa una nueva flota de instancias para un clúster.	Agregue nodos de datos.

Estado amarillo del clúster

Un estado amarillo del clúster significa que los fragmentos principales de todos los índices están asignados a nodos de un clúster, pero los fragmentos de réplica de al menos un índice no lo están. Los clústeres de un solo nodo siempre se inicializan con un estado de clúster amarillo porque no hay ningún otro nodo al que OpenSearch Service pueda asignar una réplica. Para conseguir el estado verde del clúster, aumente el número de nodos. Para obtener más información, consulte the section called “Determinación del tamaño de dominios”.

Los clústeres de varios nodos pueden tener brevemente un estado de clúster amarillo después de crear un nuevo índice o después de un error de nodo. Este estado se resuelve automáticamente a medida que se OpenSearch replican los datos en todo el clúster. La falta de espacio en disco también puede causar el estado de clúster amarillo; el clúster sólo puede distribuir fragmentos de réplica si los nodos tienen espacio en disco para acomodarlos.

ClusterBlockException

Puede que reciba un error `ClusterBlockException` por los siguientes motivos.

Falta de espacio de almacenamiento disponible

Si uno o más nodos del clúster tienen un espacio de almacenamiento inferior al valor mínimo del 1) 20% del espacio de almacenamiento disponible o 2) 20 GiB de espacio de almacenamiento, las operaciones de escritura básicas, como añadir documentos y crear índices, pueden empezar a fallar. the section called “Cálculo de requisitos de almacenamiento”proporciona un resumen de cómo OpenSearch Service utiliza el espacio en disco.

Para evitar problemas, supervise la `FreeStorageSpace` métrica en la consola de OpenSearch servicio y cree CloudWatch alarmas que se activen cuando `FreeStorageSpace` caiga por debajo de un umbral determinado. `GET /_cat/allocation?v`también proporciona un resumen útil de la asignación de particiones y el uso del disco. Para resolver los problemas relacionados con la falta de espacio de almacenamiento, amplíe su dominio de OpenSearch servicio para utilizar tipos de instancias más grandes, más instancias o más almacenamiento basado en EBS.

Presión alta de memoria de JVM

Cuando la métrica de JVMMemorypresión supera el 92% durante 30 minutos, OpenSearch Service activa un mecanismo de protección y bloquea todas las operaciones de escritura para evitar que el clúster alcance el estado rojo. Cuando la protección está activada, se produce un error `ClusterBlockException` en las operaciones de escritura, los nuevos índices no se pueden crear y se genera el error `IndexCreateBlockException`.

Cuando la métrica de JVMMemorypresión vuelve al 88% o menos durante cinco minutos, la protección se desactiva y las operaciones de escritura en el clúster se desbloquean.

La presión alta de memoria de JVM puede deberse a picos en el número de solicitudes al clúster, asignaciones de particiones desequilibradas en los nodos, demasiadas particiones en un clúster, explosiones de mapeo de índices o datos de campo, o tipos de instancias que no pueden gestionar las cargas entrantes. También puede deberse al uso de agregaciones, comodines o intervalos de tiempo amplios en las consultas.

Para reducir el tráfico al clúster y resolver problemas de presión alta de memoria de JVM, pruebe una o más de las siguientes acciones:

- Escale el dominio para que el tamaño máximo de la pila por nodo sea de 32 GB.
- Reduzca el número de particiones mediante la eliminación de los índices antiguos o no utilizados.

- Borre la memoria caché de datos con la Operación de la API de `POST` *`index-name`*`/_cache/clear?fielddata=true`. Tenga en cuenta que borrar la memoria caché puede interrumpir las consultas en curso.

En general, para evitar una presión alta de memoria de JVM en el futuro, siga estas prácticas recomendadas:

- Evite la agregación en los campos de texto o cambie el tipo de mapeo para sus índices `keyword`.
- Optimice las solicitudes de búsqueda e indexación mediante la selección del número correcto de particiones.
- Configure políticas de Index State Management (ISM) paraeliminar los índices que no se utilizan periódicamente.

Error al migrar a multi-AZ con modo de espera

Se pueden producir los siguientes problemas al migrar un dominio existente a Multi-AZ con modo de espera.

Creación de un índice, una plantilla de índice o una política de ISM durante la migración de dominios sin modo de espera a dominios con modo de espera

Si crea un índice al migrar un dominio de Multi-AZ sin modo de espera a uno con modo de espera, y la plantilla de índice o la política de ISM no siguen las pautas de copia de datos recomendadas, esto puede provocar una incoherencia en los datos y es posible que la migración no se realice correctamente. Para evitar esta situación, cree el nuevo índice con un recuento de copias de datos (incluidos los nodos principales y las réplicas) que sea múltiplo de tres. Puede comprobar el progreso de la migración mediante la API de `DescribeDomainChangeProgress`. Si encuentra un error en el recuento de réplicas, corrija el error y, a continuación, póngase en contacto con Soporte de AWS para volver a intentar la migración.

Número incorrecto de copias de datos

Si no tiene el número correcto de copias de datos en su dominio, la migración a Multi-AZ con modo de espera fallará.

JVM OutOfMemoryError

`OutOfMemoryError` de una JVM significa normalmente que se ha alcanzado uno de los siguientes interruptores de la JVM.

Interruptor	Description (Descripción)	Propiedad de configuración del clúster
Interruptor principal	Porcentaje total de memoria del montón de la JVM permitido para todos los interruptores. El valor predeterminado es 95 %.	`indices.breaker.total.limit`
Interruptor de datos de campo	Porcentaje de memoria del montón de la JVM permitido para cargar en memoria un único campo de datos. El valor predeterminado es 40 %. Si carga datos con campos de gran tamaño, probablemente deba aumentar este límite.	`indices.breaker.fielddata.limit`
Interruptor de solicitud	Porcentaje de memoria del montón de la JVM permitido para las estructuras de datos que se usan para responder a una solicitud de servicio. El valor predeterminado es 60 %. Si sus solicitud es de servicio implican el cálculo de agregaciones, es posible que deba aumentar este límite.	`indices.breaker.request.limit`

Nodos de clúster defectuosos

Las instancias de Amazon EC2 pueden experimentar terminaciones y reinicios inesperados. Por lo general, el OpenSearch servicio reinicia los nodos por usted. Sin embargo, es posible que uno o varios nodos de un clúster de OpenSearch permanezcan en una condición de error.

Para comprobar este estado, abre el panel de control de tu dominio en la consola OpenSearch de servicio. Vaya a la pestaña Estado del clúster y busque la métrica Nodos totales. Compruebe si el número de nodos que aparece es menor que el número que configuró para el clúster. Si la métrica indica que uno o varios nodos no funcionan durante más de un día, contacte con AWS Support.

También puedes configurar una CloudWatch alarma para que te avise cuando se produzca este problema.

> **Note**
>
> La métrica Nodos totales no es precisa durante los cambios en la configuración del clúster y durante el mantenimiento rutinario del servicio. Este es el comportamiento esperado. La métrica informará en breve del número correcto de nodos del clúster. Para obtener más información, consulte the section called “Cambios de configuración”.

Para proteger sus clústeres de las terminaciones y reinicios inesperados de los nodos, cree al menos una réplica para cada índice de su dominio de OpenSearch servicio.

Límite máximo de fragmentos superado

OpenSearch así como 7. Las versiones x de Elasticsearch tienen una configuración predeterminada de no más de 1000 fragmentos por nodo. OpenSearch/Elasticsearch arroja un error si una solicitud, como la creación de un índice nuevo, provoca que superes este límite. Si detecta este error, dispone de varias opciones:

- Agregue más nodos de datos al clúster.
- Aumente la configuración `_cluster/settings/cluster.max_shards_per_node`.
- Utilice la API _shrink para reducir el número de particiones en el nodo.

Dominio atascado en estado de procesamiento

Tu dominio OpenSearch de servicio pasa al estado de «Procesamiento» cuando se encuentra en medio de un cambio de configuración. Al iniciar un cambio de configuración, el estado del dominio cambia a «Procesando», mientras que el OpenSearch Servicio crea un nuevo entorno. En el nuevo entorno, OpenSearch Service lanza un nuevo conjunto de nodos aplicables (como nodos de datos, maestros o UltraWarm). Una vez que finaliza la migración, se terminan los nodos más antiguos.

El clúster puede quedarse atascado en el estado “Processing” (Procesando) si se produce una de estas situaciones:

- No se lanza un nuevo conjunto de nodos de datos.
- La migración de particiones al nuevo conjunto de nodos de datos no se realiza correctamente.
- La comprobación de validación ha fallado con errores.

Para ver los pasos de resolución detallados en cada una de estas situaciones, consulta ¿Por qué mi dominio de Amazon OpenSearch Service está atascado en el estado «Procesando»? .

Bajo balance de ráfaga EBS

OpenSearch El servicio le envía una notificación de consola cuando el saldo de ráfagas de EBS de uno de sus volúmenes de uso general (SSD) es inferior al 70% y una notificación de seguimiento si el saldo cae por debajo del 20%. Para solucionar este problema, puede escalar verticalmente el clúster o reducir las IOPS de lectura y escritura para que se pueda acreditar el balance de ráfaga. El balance de ráfagas se mantiene en 0 para los dominios con tipos de volúmenes gp3, así como para los dominios con volúmenes gp2 que tengan un tamaño de volumen superior a 1000 GiB. Para obtener más información, consulte Volúmenes de SSD de uso general (gp2). Puede monitorizar el saldo de ráfagas de EBS con la `BurstBalance` CloudWatch métrica.

La métrica de EBS aumenta durante el redimensionamiento del volumen

Al modificar los tamaños de volumen de Amazon Elastic Block Store, es posible que observe aumentos temporales en varias métricas de EBS como `Write Throughput`, `Write Throughput`

`Micro bursting`, `Disk Queue Depth` y `IOPS`. Este es el comportamiento esperado durante la operación de cambio de tamaño y, por lo general, dura unos segundos. Sin embargo, la duración puede variar en función del tamaño del volumen que se modifique.

Para evitar problemas de latencia y rechazos de solicitudes, realice cambios de tamaño del volumen de EBS solo cuando el clúster esté en buen estado y el tráfico de red sea bajo.

No se pueden habilitar los registros de auditoría

Es posible que aparezca el siguiente error al intentar habilitar la publicación del registro de auditoría mediante la consola de OpenSearch servicio:

La política de acceso a los recursos especificada para el grupo de CloudWatch registros no concede permisos suficientes para que Amazon OpenSearch Service cree un flujo de registros. Compruebe la política de acceso a recursos.

Si detecta este error, compruebe que el `resource`elemento de su política incluye el ARN del grupo de registro correcto. Si lo hace, siga estos pasos:

1. Espere varios minutos.
2. Actualice la página en el navegador web.
3. Seleccione Seleccionar grupo existente.
4. Para Grupo de registros existente, seleccione el grupo de registro que ha creado antes de recibir el mensaje de error.
5. En la sección de políticas de acceso, seleccione Seleccionar política existente.
6. Para Política existente, elija la política que ha creado antes de recibir el mensaje de error.
7. Seleccione Habilitar.

Si el error persiste después de repetir el proceso varias veces, póngase en contacto con el servicio de Soporte de AWS.

No se puede cerrar el índice

OpenSearch El servicio admite la `_close`API solo para las versiones 7.4 OpenSearch y posteriores de Elasticsearch. Si está usando una versión anterior está restaurando un índice desde una instantánea, puede eliminar el índice existente (antes o después de volver a indexarlo).

Verificaciones de licencias del cliente

Las distribuciones predeterminadas de Logstash y Beats incluyen una comprobación de la licencia de propiedad y no se conectan a la versión de código abierto de. OpenSearch Asegúrese de utilizar las distribuciones Apache 2.0 (OSS) de estos clientes con Service. OpenSearch

Limitación controlada de solicitudes

Si recibe errores persistentes de `403 Request throttled due to too many requests` o `429 Too Many Requests`, considere la posibilidad de escalar verticalmente. Amazon OpenSearch Service limita las solicitudes si la carga útil puede provocar que el uso de memoria supere el tamaño máximo del montón de Java.

No se puede usar SSH en nodo

No puedes usar SSH para acceder a ninguno de los nodos de tu OpenSearch clúster ni modificarlos directamente. `opensearch.yml` En su lugar, usa la consola o SDKs para configurar tu dominio. AWS CLI También puede especificar algunos ajustes a nivel de clúster mediante el OpenSearch REST APIs. Para obtener más información, consulta la referencia de la API de Amazon OpenSearch Service ythe section called “Operaciones admitidas”.

Si necesita más información sobre el rendimiento del clúster, puede publicar los registros de errores y los registros lentos en CloudWatch.

Error de instantánea "No válido para la clase de almacenamiento del objeto"

OpenSearch Las instantáneas de servicio no son compatibles con la clase de almacenamiento Amazon Glacier. Es posible que aparezca este error al intentar mostrar las instantáneas si el bucket de S3 incluye una regla de ciclo de vida que traslada los objetos a la clase de almacenamiento de Amazon Glacier.

Si tiene que restaurar una instantánea desde el bucket, restaure los objetos de Amazon Glacier, copie los objetos a un nuevo bucket y registre el nuevo bucket como repositorio de instantáneas.

Encabezado de host no válido

OpenSearch El servicio requiere que los clientes lo especifiquen `Host` en los encabezados de las solicitudes. Un valor de `Host` válido es el punto de conexión del dominio sin `https://`, como:

```
Host: search-my-sample-domain-ih2lhn2ew2scurji.us-west-2.es.amazonaws.com
```

Si recibes un `Invalid Host Header` error al realizar una solicitud, comprueba que tu cliente o proxy incluya el punto final del dominio del OpenSearch servicio (y no, por ejemplo, su dirección IP) en el `Host` encabezado.

Tipo de instancia M3 no válido

OpenSearch El servicio no permite añadir ni modificar instancias M3 a dominios existentes que estén ejecutando Elasticsearch OpenSearch o a versiones 6.7 o posteriores. Puede seguir utilizando instancias M3 con Elasticsearch 6.5 y versiones anteriores.

Recomendamos elegir un tipo de instancia más reciente. Para los dominios que ejecutan OpenSearch Elasticsearch 6.7 o versiones posteriores, se aplica la siguiente restricción:

- Si el dominio existente no utiliza instancias M3, ya no podrá cambiar a ellas.
- Si cambia un dominio existente de un tipo de instancia M3 a otro tipo de instancia, no puede volver a cambiar.

Las consultas activas dejan de funcionar después de habilitarlas UltraWarm

Cuando se habilita UltraWarm en un dominio, si no hay ninguna anulación previa de la `search.max_buckets` configuración, OpenSearch Service establece automáticamente el valor en `10000` para evitar que las consultas que consumen mucha memoria saturen los nodos calientes. Si tus consultas más frecuentes utilizan más de 10 000 cubos, es posible que dejen de funcionar cuando las habilites. UltraWarm

Como no puedes modificar esta configuración debido a la naturaleza gestionada de Amazon OpenSearch Service, necesitas abrir un caso de soporte para aumentar el límite. Los aumentos de límite no requieren una suscripción de premium support.

No se puede cambiar a una versión anterior después de una actualización

Las actualizaciones in situ son irreversibles, pero si se pone en contacto con AWS Support, pueden ayudarle a restaurar la instantánea automática previamente actualizada en un nuevo dominio. Por ejemplo, si actualizas un dominio de Elasticsearch 5.6 a 6.4, AWS Support puede ayudarte a restaurar la instantánea previa a la actualización en un nuevo dominio de Elasticsearch 5.6. Si ha realizado una instantánea manual del dominio original, puede realizar ese paso usted mismo.

¿Necesitas un resumen de los dominios de todos Regiones de AWS

El siguiente script utiliza el AWS CLI comando describe-regions de Amazon EC2 para crear una lista de todas las regiones en las que OpenSearch el servicio podría estar disponible. A continuación, llama a cada región 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
```

Recibe el siguiente resultado para cada región:

```
Listing domains in region:'us-west-2'...
[
  {
    "DomainName": "sample-domain"
  }
]
```

Las regiones en las que el OpenSearch servicio no está disponible muestran el mensaje «No se ha podido conectar a la URL del punto final».

Error del navegador al usar los OpenSearch paneles

Su navegador incluye los mensajes de error del servicio en los objetos de respuesta HTTP cuando utiliza los paneles de control para ver los datos de su OpenSearch dominio de servicio. Puede usar

las herramientas para desarrolladores que suele haber disponibles en los navegadores web, como el Modo desarrollador de Chrome, para ver los errores de servicio subyacentes y como ayuda en las tareas de depuración.

Para ver errores de servicio en Chrome

1. Desde la barra del menú superior de Chrome, seleccione Ver, Desarrollador, Herramientas de desarrollador.
2. Seleccione la pestaña Red.
3. En la columna Estado, seleccione cualquier sesión HTTP que tenga un estado de 500.

Para ver errores de servicio en Firefox

1. Desde el menú, seleccione Herramientas, Desarrollador de web, Red.
2. Seleccione cualquier sesión HTTP que tenga un estado de 500.
3. Seleccione la pestaña Respuesta para ver la respuesta de servicio.

Partición de nodos y sesgo de almacenamiento

El sesgo de partición de nodo es cuando uno o más nodos de un clúster tienen significativamente más particiones que los demás nodos. El sesgo de almacenamiento de nodo es cuando uno o más nodos de un clúster tienen mucho más almacenamiento (`disk.indices`) que los demás nodos. Si bien estas dos condiciones pueden ocurrir temporalmente, por ejemplo, cuando un dominio ha reemplazado un nodo y todavía le está asignando particiones, debe abordarlas si persisten.

Para identificar ambos tipos de sesgo, ejecute la operación de la API _cat/allocation y compare las entradas `shards` y `disk.indices` en la respuesta:

```
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
```

Si bien es normal que haya algún sesgo de almacenamiento, cualquier valor por encima del 10 % del promedio es significativo. Cuando la distribución de las particiones está sesgada, el uso de la CPU, la red y el ancho de banda del disco también puede verse sesgado. Debido a que una mayor cantidad de datos generalmente implica más operaciones de indexación y búsqueda, los nodos más pesados también tienden a ser los nodos con mayor carga de recursos, mientras que los nodos más livianos representan una capacidad infrautilizada.

Corrección: utilice recuentos de particiones que sean múltiplos del recuento de nodos de datos para garantizar que cada índice se distribuya de manera uniforme entre los nodos de datos.

Partición del índice y el sesgo de almacenamiento

El sesgo de partición del índice es cuando uno o más nodos contienen más particiones de un índice que los otros nodos. El sesgo de almacenamiento de índice es cuando uno o más nodos contienen una cantidad desproporcionadamente grande del almacenamiento total de un índice.

El sesgo del índice es más difícil de identificar que el sesgo del nodo porque requiere cierta manipulación de la salida de la API _cat/shards. Investigue el sesgo del índice si hay algún indicio de sesgo en las métricas de clúster o nodo. A continuación, se indican algunos indicios comunes del sesgo de índice:

- Errores HTTP 429 que se producen en un subconjunto de nodos de datos
- Colocación en cola de operaciones de búsqueda o índice desigual en los nodos de datos
- Utilización desigual de la and/or CPU apilada de JVM en todos los nodos de datos

Corrección: utilice recuentos de particiones que sean múltiplos del recuento de nodos de datos para garantizar que cada índice se distribuya de manera uniforme entre los nodos de datos. Si el almacenamiento de índices o los fragmentos siguen sesgados, es posible que tengas que forzar una reasignación de los fragmentos, lo que ocurre con cada despliegue azul o verde de tu dominio de servicio. OpenSearch

Operación no autorizada tras seleccionar acceso a la VPC

Al crear un dominio nuevo mediante la consola de OpenSearch servicio, tiene la opción de seleccionar VPC o acceso público. Si seleccionas el acceso a la VPC, el OpenSearch servicio consulta la información de la VPC y produce un error si no tienes los permisos adecuados:

```
You are not authorized to perform this operation. (Service: AmazonEC2; Status Code:
 403; Error Code: UnauthorizedOperation
```

Para poder realizar esta consulta, debe tener acceso a las operaciones `ec2:DescribeVpcs`, `ec2:DescribeSubnets` y `ec2:DescribeSecurityGroups`. Este requisito solo es para la consola. Si usa la AWS CLI para crear y configurar un dominio con un punto final de VPC, no necesita acceder a esas operaciones.

Bloqueo al cargar después de crear el dominio de la VPC

Después de crear un dominio que utiliza el acceso a la VPC, es posible que la operación Configuration state (estado de configuración) del dominio no pase del estado Loading (Cargando). Si se produce este problema, es probable que tengas AWS Security Token Service (AWS STS) deshabilitado en tu región.

Para añadir puntos de enlace de VPC a su VPC, el OpenSearch servicio debe asumir la función. `AWSServiceRoleForAmazonOpenSearchService` Por lo tanto, AWS STS debe estar habilitado para crear nuevos dominios que utilicen el acceso a la VPC en una región determinada. Para obtener más información sobre cómo habilitar y deshabilitar AWS STS, consulta la Guía del usuario de IAM.

Solicitudes denegadas a la API OpenSearch

Con la introducción del control de acceso basado en etiquetas para la OpenSearch API, es posible que empieces a ver errores de acceso denegado que antes no aparecían. Esto puede deberse a que una o más de sus políticas de acceso contienen `Deny` y utilizan la condición `ResourceTag`, y ahora se están respetando esas condiciones.

Por ejemplo, la política siguiente solo denegaba el acceso a la `CreateDomain` de la API de configuración, si el dominio tenía la etiqueta `environment=production`. Aunque la lista de acciones también incluye `ESHttpPut`, la declaración de denegación no se aplicó a esa acción ni a ninguna otra acción `ESHttp*`.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [{
    "Action": [
      "es:CreateDomain",
      "es:ESHttpPut"
    ],
    "Effect": "Deny",
    "Resource": "*",
    "Condition": {
      "ForAnyValue:StringEquals": {
        "aws:ResourceTag/environment": [
          "production"
        ]
      }
    }
  }]
}
```

Con la incorporación de etiquetas para los métodos OpenSearch HTTP, una política de IAM basada en la identidad como la anterior provocará que se deniegue al usuario adjunto el acceso a la acción. `ESHttpPut` Anteriormente, en ausencia de validación de etiquetas, el usuario adjunto habría podido enviar solicitudes PUT.

Si empieza a ver errores de acceso denegado después de actualizar sus dominios al software de servicio R20220323 o posterior, compruebe las políticas de acceso basadas en identidad para ver si es así y actualícelas si es necesario para permitir el acceso.

No se puede conectar desde Alpine Linux

Alpine Linux limita el tamaño de respuesta de DNS a 512 bytes. Si intenta conectarse a su dominio de OpenSearch servicio desde la versión 3.18.0 o inferior de Alpine Linux, la resolución de DNS puede fallar si el dominio está en una VPC y tiene más de 20 nodos. Si utiliza una versión de Alpine Linux superior a la 3.18.0, debería poder resolver más de 20 hosts. Para obtener más información, consulte las notas de la versión de Alpine Linux 3.18.0.

Si su dominio está en una VPC, le recomendamos que utilice otras distribuciones de Linux, como Debian, Ubuntu, CentOS, Red Hat Enterprise Linux o Amazon Linux 2, para conectarse a él.

Hay demasiadas solicitudes de Search Backpressure

El control de admisión basado en la CPU es un mecanismo de control que limita de forma proactiva el número de solicitudes a un nodo en función de su capacidad actual, tanto en caso de incrementos orgánicos como de picos de tráfico. Las solicitudes excesivas devuelven un código de estado HTTP 429 “Demasiadas solicitudes” en caso de rechazo. Este error indica que los recursos del clúster son insuficientes, que las solicitudes de búsqueda consumen muchos recursos o que se ha producido un aumento imprevisto de la carga de trabajo.

Search Backpressure proporciona el motivo del rechazo, lo que puede ayudar a ajustar las solicitudes de búsqueda que consumen muchos recursos. En caso de picos de tráfico, recomendamos volver a intentarlo desde el lado del cliente, con un retroceso y una fluctuación exponenciales.

Error de certificado cuando se utiliza un SDK

Como AWS SDKs utiliza los certificados de CA de su ordenador, los cambios en los certificados de los AWS servidores pueden provocar errores de conexión al intentar utilizar un SDK. Los mensajes de error varían, pero suelen contener el siguiente texto:

```
Failed to query OpenSearch
...
SSL3_GET_SERVER_CERTIFICATE:certificate verify failed
```

Puede evitar estos errores conservando los certificados de CA y el sistema operativo de su equipo up-to-date. Si detecta este problema en un entorno corporativo y no administra su propio equipo, es posible que deba pedir ayuda a un administrador con el proceso de actualización.

La siguiente lista muestra las versiones mínimas necesarias del sistema operativo y Java:

- Las versiones de Microsoft Windows que tienen instaladas actualizaciones de enero de 2005 o posteriores contienen al menos una de las obligatorias CAs en su lista de confianza.
- Mac OS X 10.4 con Java para Mac OS X 10.4 versión 5 (febrero de 2007), Mac OS X 10.5 (octubre de 2007) y versiones posteriores incluyen al menos una de las obligatorias CAs en su lista de confianza.

- Red Hat Enterprise Linux 5 (marzo de 2007), 6 y 7 y CentOS 5, 6 y 7 contienen al menos uno de los requisitos CAs en su lista de CA de confianza predeterminada.
- Java 1.4.2_12 (mayo de 2006), 5 Update 2 (marzo de 2005) y todas las versiones posteriores, incluidas Java 6 (diciembre de 2006), 7 y 8, contienen al menos uno de los elementos obligatorios CAs en su lista de CA de confianza predeterminada.

Las tres entidades de certificación son:

- Amazon Root CA 1
- Starfield Services Root Certificate Authority - G2
- Starfield Class 2 Certification Authority

Los certificados raíz de las dos primeras autoridades están disponibles en Amazon Trust Services, pero conservar el ordenador up-to-date es la solución más sencilla. Para obtener más información sobre los certificados proporcionados por ACM, consulte. AWS Certificate Manager FAQs

> **Note**
>
> Actualmente, los dominios de OpenSearch servicio de la región us-east-1 utilizan certificados de una autoridad diferente. Tenemos previsto actualizar la región para utilizar estas nuevas entidades de certificación en un futuro próximo.

La instalación del complemento personalizado falla debido a la compatibilidad de las versiones

Problema: la instalación del complemento falló debido a una discordancia de versiones entre el complemento y la instancia en ejecución. OpenSearch La función devuelve el siguiente error:

```
PluginValidationFailureReason : The provided plugin could not be loaded.
```

Causa: el complemento se compiló para OpenSearch ${*MAJOR*}. ${*MINOR*}.{*PATCH*}, pero su entorno ejecuta OpenSearch ${*MAJOR*}. ${*MINOR*} .0. OpenSearch requiere una coincidencia exacta de versiones entre los complementos y la OpenSearch instalación principal por motivos de estabilidad y seguridad.

Posible solución: crea el plugin con la OpenSearch versión ${*MAJOR*}. $ *{MINOR}* .0 para que coincida con la versión de tu clúster.

Para comprobar y actualizar la versión de OpenSearch

1. Utilice la API o el panel de control del clúster para ejecutar el siguiente comando. Sustituya *default placeholder values* por su propia información.

   Solicitud de API:

   ```
   curl -X GET your-opensearch-endpoint/
   ```

   Consola de Dev Tools en el panel de control:

   ```
   GET /
   ```

   El comando devuelve información similar a la siguiente.

   ```
   {
     "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. Si el número de versión no es ${*MAJOR*}. $ *{MINOR}* .0, reconstruya el complemento de la siguiente manera:

   a. Actualiza el plugin `descriptor.properties` para especificar la versión ${*MAJOR*}. $ *{MINOR}* .0.

b. Reconstruya el complemento con el comando correspondiente a su tipo de proyecto.

c. Ejecute el comando update-package con el archivo recién creado `.zip`.

d. Ejecute el comando associate-package para asociar la última versión del complemento que se creó al ejecutar el comando `update-package` en el paso anterior.

Historial de documentos de Amazon OpenSearch Service

En este tema se describen los cambios importantes en Amazon OpenSearch Service. Las actualizaciones de software de servicio agregan compatibilidad con nuevas características, parches de seguridad, correcciones de errores y otras mejoras. Para utilizar las nuevas características, tal vez deba actualizar el software de servicio de su dominio. Para obtener más información, consulte the section called “Actualizaciones del software del servicio”.

Las funciones del servicio se implementan de forma gradual según el Regiones de AWS lugar donde esté disponible el servicio. Actualizamos esta documentación solo para la primera versión. No proporcionamos información sobre la disponibilidad en regiones ni anunciamos lanzamientos posteriores en regiones. Para obtener información sobre la disponibilidad regional de las funciones del servicio y suscribirse a las notificaciones sobre actualizaciones, consulte ¿Qué hay de nuevo en el mercado? AWS

Para recibir notificaciones sobre las actualizaciones, puede suscribirse a la fuente RSS.

> **Note**
>
> Versiones de revisión: las versiones de software de servicio que terminan en “-P” y un número, como R20211203-P4, son versiones de revisión. Es probable que las revisiones incluyan mejoras de rendimiento, correcciones de errores menores y correcciones de seguridad o mejoras de posición. Dado que las revisiones no incluyen nuevas características ni cambios extraordinarios, por lo general no tienen un impacto directo en el usuario ni la documentación, y por ese motivo no se incluyen los detalles específicos de cada revisión en este historial de documentos.

Cambio	Descripción	Fecha
Support para OpenSearch 3.5	Amazon OpenSearch Service es compatible con las OpenSearch versiones 3.4 y 3.5. Para obtener más información, consulta las notas de la versión OpenSearch 3.4 y las notas de la versión	17 de marzo de 2026

	OpenSearch 3.5. Para obtener información sobre las operaciones OpenSearch 3.4 y 3.5, consulta la referencia de la API OpenSearch REST o la referencia de la API para el complemento específico. Para obtener una lista completa de los cambios más important es, consulta la OpenSearch documentación aquí.	
Nuevas funciones: aceleración por GPU para índices vectorial es, optimización automática e ingestión de vectores	Amazon OpenSearch Service ahora admite la aceleración por GPU para índices vectorial es, lo que permite realizar operaciones de búsqueda vectoriales más rápidas. Amazon OpenSearch Service también presenta la optimizac ión automática para ajustar automáticamente el rendimien to de la recopilación y la ingesta de vectores para simplificar la carga de datos. Para obtener más informaci ón, consulte Aceleración por GPU para índices vectorial es, Optimización automática e Ingesta de vectores.	2 de diciembre de 2025

Support para OpenSearch 3.3	Amazon OpenSearch Service es compatible con las OpenSearch versiones 3.2 y 3.3. Para obtener más información, consulte las notas de la versión OpenSearch 3.2 y las notas de la versión OpenSearch 3.3. Para obtener información sobre las operaciones OpenSearch 3.2 y 3.3, consulta la referencia de la API OpenSearch REST o la referencia de la API para el complemento específico.	19 de noviembre de 2025
Compatibilidad con familias de instancias basadas en Graviton4 (c8g, m8g, r8g, r8gd)	Amazon OpenSearch Service ahora admite familias de instancias basadas en Graviton4 (c8g, m8g, r8g, r8gd). Para obtener más información, consulta Tipos de instancias compatibles en Amazon OpenSearch Service.	17 de octubre de 2025

Nueva función: aprendizaje automático OpenSearch de ingestión (inferencia por lotes offline)	Las canalizaciones de OpenSearch Amazon Ingestion (OSI) admiten el procesamiento de inferencias por lotes fuera de línea con aprendizaje automático (ML) para enriquecer de manera eficiente grandes volúmenes de datos a bajo costo. Utilice la inferencia en lotes sin conexión siempre que tenga conjuntos de datos grandes que puedan procesarse de forma asíncrona. La inferencia de lotes sin conexión funciona con Amazon Bedrock y SageMaker modelos. Esta función está disponible en todos los dominios Regiones de AWS compatibles con OpenSearch Ingestion with OpenSearch Service 2.17 o más. Para obtener más información, consulte Uso de una canalización de OpenSearch ingestión con la inferencia por lotes fuera de línea con aprendizaje automático.	2 de octubre de 2025

Nueva función: CloudFormation plantillas para configurar la inferencia remota para el resaltado semántico	El resaltado semántico es una característica de búsqueda avanzada que mejora la relevancia de los resultados al analizar el significado y el contexto de las consultas en lugar de basarse únicamente en las coincidencias exactas de las palabras clave. Esta función utiliza modelos de machine learning para evaluar la similitud semántica entre las consultas de búsqueda y el contenido del documento, identificando y resaltando las frases o pasajes más relevantes desde el punto de vista contextual de los documentos. Amazon OpenSearch Service ahora incluye una CloudFormation plantilla que configura automáticamente los SageMaker modelos de Amazon para el resaltado semántico. Para obtener más información, consulte Inferencia remota para plantillas de resaltado semántico.	26 de septiembre de 2025

Se actualizó la política AmazonOpenSearchIngestionServiceRolePolicy gestionada	La actualización otorga permiso a OpenSearch Ingestion para modificar los puntos de enlace de VPC creados OpenSearch por para compartir canalizaciones entre sí. VPCs	19 de septiembre de 2025
Nueva función: configure las canalizaciones de OpenSearch ingestión para la ingesta entre cuentas	En el caso de fuentes basadas en push, como HTTP y OTel, Amazon OpenSearch Ingestion le permite compartir canalizaciones Cuentas de AWS desde una nube privada virtual (VPC) a un punto final de canalización en una VPC independiente. Para obtener más información, consulte Configuración OpenSearch de canalizaciones de ingestión para la ingestión entre cuentas	19 de septiembre de 2025
Amazon OpenSearch Serverless admite la búsqueda vectorial basada en disco	La búsqueda vectorial basada en disco reduce considera blemente los costos operativo s de las cargas de trabajo de vector en entornos con poca memoria. Para obtener más información, consulte Búsqueda vectorial basada en discos.	18 de septiembre de 2025

Support para OpenSearch 3.1	Amazon OpenSearch Service es compatible con OpenSearch la versión 3.1. Para obtener información sobre la compatibilidad de las actualizaciones, consulte Rutas de actualización compatibles. Para obtener más información sobre los cambios en las nuevas versiones, consulta la documentación de las versiones OpenSearch 3.0 y 3.1. Para obtener información sobre las operaciones de la OpenSearch versión 3.1, consulta la OpenSearch referencia de la API de REST o la referencia de la API del complemento específico.	9 de septiembre de 2025

Amazon OpenSearch Serverless es compatible con los estándares federales de procesamiento de información (FIPS) 140-2	La norma FIPS 140-2 es un estándar del Gobierno estadounidense y canadiense que especifica los requisitos de seguridad para los módulos criptográficos que protegen la información confidencial. Los puntos de enlace FIPS OpenSearch sin servidor de Amazon están disponibles en los lugares Regiones de AWS donde se admite FIPS. Estos puntos de conexión utilizan TLS 1.2 o una versión posterior y algoritmos criptográficos validados FIPS para todas las comunicaciones.	5 de septiembre de 2025

Nueva característica: control de acceso detallado basado en atributos SAML	Amazon OpenSearch Service admite un control de acceso detallado con SAML para asignar dinámicamente los usuarios y grupos de su proveedor de identidad a los usuarios y roles de control de acceso OpenSearch detallados. Puede asignar estas funciones a OpenSearch dominios específicos y colecciones sin servidor, y definir los permisos a nivel de índice y la seguridad a nivel de los documentos. Para obtener más información, consulte Configuración del control de acceso detallado basado en atributos de SAML.	8 de agosto de 2025

Tema nuevo: puntos finales y cuotas de la interfaz de usuario OpenSearch	Amazon OpenSearch UI amplía su experiencia de análisis operativo moderniza da a nuevas Regiones de AWS, incluidas Asia Pacífico (Hyderabad), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Europa (Milán), Europa (Zúrich), Europa (España) y EE. UU. Oeste (Norte de California), lo que le permite obtener información sobre los datos que abarcan dominios gestionados y colecciones sin servidor desde un único punto final. OpenSearch La interfaz de usuario puede conectarse a OpenSearc h dominios superiores a la versión 1.3 y a colecciones sin servidor. OpenSearch Ahora está disponible en 22 Region es de AWS.	7 de agosto de 2025

Nuevas características: enriquecimiento semántico automático	El enriquecimiento semántico automático simplifica la implementación y las capacidades de la búsqueda semántica en Amazon Service. OpenSearch La búsqueda semántica arroja resultados de consulta que incorporan no solo la coincidencia de palabras clave, sino también la intención y el significado contextual de la búsqueda del usuario. Para obtener más información, consulte Acerca del enriquecimiento semántico automático. El enriquecimiento semántico automático ofrece los siguientes beneficios: Implementación simplificada - No necesita conocimiento experto en machine learning (ML) ni integraciones complejas. Proceso automatizado - El enriquecimiento semántico se produce automáticamente durante la ingesta de datos. Relevancia de búsqueda mejorada	6 de agosto de 2025

	El enriquecimiento semántico mejora la calidad y la precisión contextual de los resultados de búsqueda.
Escalabilidad	El enriquecimiento semántico aplica la búsqueda semántica a grandes conjuntos de datos sin intervención manual.

Nueva función: Machine Learning en Amazon OpenSearch Serverless	OpenSearch Serverless ahora admite las capacidades de Machine Learning para flujos de trabajo de IA, como RAG (Retrieval-Augmented Generation) y búsqueda semántica. Utilice conectores para acceder a los algoritmos de aprendizaje automático y a los modelos remotos alojados en plataformas de terceros o en Servicios de AWS. Para obtener más información, consulte Configurar Machine Learning en Amazon OpenSearch Serverless. Machine Learning on OpenSearch Serverless incluye las siguientes funciones: Búsqueda neuronal y búsqueda híbrida Realice operaciones de búsqueda semántica que comprendan el contexto y el significado de la consulta para obtener resultados más relevantes. Combine la búsqueda semántica y por palabra clave con la búsqueda híbrida para mejorar la relevancia. Para	6 de agosto de 2025

obtener más informaci
ón, consulte Configurar
la búsqueda neuronal y
la búsqueda híbrida en
sistemas OpenSearch sin
servidor.

Conectores y modelos

Conéctese a modelos
alojados en plataformas
de aprendizaje automátic
o de terceros y Servicios
de AWS. Configure los
conectores con modelos y
despliéguelos para ejecutar
predicciones y flujos de
trabajo de IA.

Flujos de trabajo

Automatice las tareas
complejas de configura
ción y preprocesamiento
del aprendizaje automátic
o encadenando varias
llamadas a la API. Los
flujos de trabajo simplifican
la creación de aplicaciones
de IA. OpenSearch Para
obtener más información,
consulte Configurar flujos
de trabajo en entornos
OpenSearch sin servidor.

Nuevas funciones: instantáneas OpenSearch sin servidor	Las instantáneas son point-in-time copias de seguridad de sus colecciones de Amazon OpenSearch Serverless que proporcionan funciones de recuperación ante desastres. OpenSearch Serverless crea y administra automáticamente las instantáneas de sus colecciones, lo que garantiza la continuidad empresarial y la protección de los datos. Ventajas principales • Copias de seguridad automáticas cada hora sin necesidad de configuración manual • Gastos generales de mantenimiento cero • Sin costos de almacenamiento adicionales • Recuperación rápida en caso de pérdida accidental de datos • Posibilidad de restaurar índices específicos a partir de una instantánea Para obtener más información, consulte Copia de seguridad de las colecciones mediante instantáneas.	30 de julio de 2025

OpenSearch Ingestion admite dos fuentes nuevas	OpenSearch La ingestión ahora admite las siguientes fuentes: • Amazon Aurora • Amazon RDS	17 de julio de 2025

Nuevas funciones: integración OpenSearch de servicios con Amazon S3 Vectors	Esta integración ofrece las siguientes funciones para mejorar considerablemente la búsqueda vectorial en OpenSearch Service. • Amazon S3 Vectors ofrece el primer almacén de objetos en la nube con compatibilidad nativa para almacenar y consultar vectores. S3 Vectors proporciona un almacenamiento vectorial rentable, elástico y duradero que se puede consultar en función del significado y la similitud semánticos. Ofrece tiempos de respuesta a las consultas inferiores a un segundo y reduce los costos de carga, almacenamiento y consulta de vectores hasta un 90 %. • Amazon OpenSearch Service ofrece la posibilidad de utilizar Amazon S3 como motor vectorial para índices vectoriales. Esta característica le permite descargar datos vectoriales a Amazon S3 y, al mismo tiempo, mantener las capacidades de búsqueda vectorial en menos de un segundo a un bajo costo.	15 de julio de 2025

	Para obtener más información, consulte (Vista previa) de la integración del OpenSearch servicio con Amazon S3 Vectors.	
Tema nuevo: Integración de Atlassian con Ingestion OpenSearch	Ahora puedes usar las canalizaciones de Amazon OpenSearch Ingestion para procesar datos de los servicios de Atlassian Jira y Confluence. Esta integración te permite crear una base de conocimientos unificada con capacidad de búsqueda al sincronizar proyectos completos de Jira y espacios de Confluence en ellos. OpenSearch La integración mantiene la relevancia en tiempo real mediante la supervisión continua y la sincronización automática de las actualizaciones, al tiempo que ofrece opciones de filtrado flexibles para proyectos, tipos de problemas y contenido específicos. Para obtener más información, consulta Cómo utilizar una canalización de OpenSearch ingestión con Atlassian Services.	5 de junio de 2025

Gestión personalizada de complementos con AWS CLI	Ahora puedes gestionar los complementos personalizados para Amazon OpenSearch Service mediante AWS CLI. Esta característica le permite instalar, actualizar, proteger y desinstalar complementos personalizados mediante un flujo de trabajo de línea de comandos simplificado. También puedes proteger tus complementos con AWS KMS claves y gestionar las actualizaciones de los complementos cuando los actualices a versiones más recientes de OpenSearch. Para obtener más informaci ón, consulta Administrar complementos personali zados mediante la AWS KMS integración AWS CLI de paquetes personalizados OpenSearch de Amazon Service.	27 de mayo de 2025

Support para OpenSearch 2.19	Amazon OpenSearch Service es compatible con OpenSearch la versión 2.19. Para obtener más información, consulta las notas de la versión OpenSearch 2.19. Para obtener información sobre las operaciones de la versión OpenSearch 2.19, consulta la referencia de la API de OpenSearch REST o la referencia de la API del complemento específico.	30 de abril de 2025

Nuevo soporte para especificar las funciones de canalización en Ingestion OpenSearch	Al crear o actualizar una canalización en Amazon OpenSearch Ingestion, deja de tener un rol de canalización específico en una configuración de canalización en formato YAML. En su lugar, en la consola, debe especificar el rol en un nuevo campo de rol de Pipeline. Si utiliza el AWS CLI, utilizará un parámetro nuevo `--pipeline-role-arn` . Para obtener más información, consulte los temas siguientes: • Configuración de roles y usuarios en Amazon OpenSearch Ingestion • Creación de canalizaciones OpenSearch de Amazon Ingestion • Tutorial: Ingerir datos en un dominio mediante Amazon OpenSearch Ingestion • Tutorial: Ingerir datos en una colección mediante Amazon OpenSearch Ingestion • CreatePipeline • UpdatePipeline	22 de abril de 2025

Nuevo tema	El tema de resolución de problemas La instalación de complementos personalizados falla debido a la compatibilidad de versiones aborda los pasos que debe seguir cuando encuentre el error `PluginValidationFailureReason : The provided plugin could not be loaded` . Este error suele ser el resultado de una discordancia de versiones entre el complemento y la OpenSearch instancia en ejecución.	21 de abril de 2025
Nueva función: compatibilidad de la OpenSearch interfaz de usuario con la búsqueda entre clústeres	OpenSearch La interfaz de usuario ahora es compatible con la búsqueda entre clústeres. Esto le permite utilizar la OpenSearch interfaz de usuario en uno para acceder Región de AWS a los clústeres de una región diferente. Esto se hace configurándolo como un clúster remoto conectado a un clúster de la misma región. Para obtener más información, consulte Acceso a datos entre regiones y cuentas con búsqueda entre clústeres.	16 de abril de 2025

Contenido revisado: «Uso de la OpenSearch interfaz de usuario en Amazon OpenSearch Service»	Hemos revisado y ampliado el capítulo Uso de la OpenSearch interfaz de usuario en Amazon OpenSearch Service, incluyendo procedimientos ampliados, detalles de configuración adicionales y un nuevo historial de versiones del soporte de Amazon OpenSearch Service para OpenSearch la interfaz de usuario.	15 de abril de 2025
Desarrollador de Amazon Q para Amazon OpenSearch Service	La integración de Amazon Q con Amazon OpenSearch Service ofrece las siguientes capacidades generativas: • Genere visualizaciones utilizando un lenguaje natural • Ver resúmenes e información sobre alertas • Consulta los resúmenes de los resultados de las consultas generados por Amazon Q en la página Discover • Vea los detectores de anomalías recomendados • Accede al chat de Amazon Q para preguntas OpenSearch relacionadas	31 de marzo de 2025

Se actualizó la política AmazonOpenSearchServiceRolePolicy gestionada	La política permite OpenSearch actualizar el alcance de acceso de cualquier AWS IAM Identity Center aplicación que solo sea administrada por OpenSearch.	28 de marzo de 2025
OR2 e OM2 instancias para Amazon OpenSearch Service	Amazon OpenSearch Service ahora admite OR2 e OM2 instancias. En los análisis comparativos internos, OR2 mostró un rendimiento de indexación hasta un 26% mejor que el de R7 OR1 y un 70% más que él. OM2 se mostró hasta un 15% mejor que el M7g OR1 y un 66% más que él.	25 de marzo de 2025
Consulta directa para CloudWatch Logs y Security Lake	Amazon OpenSearch Service ahora admite consultas directas para consultar datos en CloudWatch Logs y Security Lake.	1 de diciembre de 2024
índices k-NN	A partir de OpenSearch la versión 2.17, puede mover los índices K-nn a los niveles UltraWarm y a los de almacenamiento en frío.	13 de noviembre de 2024

OpenSearch Compatibilidad con la versión 2.17	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 2.17. Esta versión incluye todas las funciones que formaban parte de las versiones 2.16 y 2.17. Para obtener más informaci ón, consulta las notas de la versión 2.16 y 2.17.	13 de noviembre de 2024
OpenSearch Compatibilidad con la versión 2.15	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 2.15. Esta versión incluye todas las funciones que formaban parte de las versiones 2.14 y 2.15. Para obtener más informaci ón, consulta las notas de la versión 2.14 y 2.15.	11 de octubre de 2024
Nuevo rol vinculado a servicio	Amazon OpenSearch Service añade una función vinculada a un servicio llamada`AWSServic eRoleForOpensearch IngestionSelfManag edVpce` , que permite a Amazon OpenSearch Ingestion enviar datos de métricas a canalizaciones con puntos de enlace de Amazon CloudWatch VPC autogesti onados.	12 de junio de 2024

Consultas directas	Amazon OpenSearch Service ahora permite ejecutar consultas directamente sobre los datos almacenados en Amazon S3, sin necesidad de ingerir los datos en un OpenSearch índice.	22 de mayo de 2024
OpenSearch Compatibilidad con la versión 2.13	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 2.13. Esta versión incluye todas las características que formaban parte de las versiones 2.12 y 2.13. Para obtener más información, consulte las notas de las versiones 2.12 y 2.13.	21 de mayo de 2024
Soporte OpenSearch de Amazon Ingestion para Data Prepper versión 2.7	Amazon OpenSearch Ingestion añade compatibi lidad con la versión 2.7 de Data Prepper. Para obtener más información, consulte las notas de la versión 2.7.	4 de abril de 2024
Servicio de AWS acceso privado para colecciones sin servidor OpenSearch	Ahora puede conceder acceso específico Servicios de AWS, como Amazon Bedrock, a sus colecciones OpenSearc h Serverless dentro de una política de acceso a la red.	28 de marzo de 2024
Actualizaciones locales de EBS	Ahora puede realizar algunos cambios de EBS en sus dominios sin una implement ación azul/verde en Amazon Service. OpenSearch	14 de febrero de 2024

Visibilidad del cambio de configuración	Ahora puedes realizar un seguimiento de los cambios en la configuración del dominio en la consola de Amazon OpenSearch Service y mediante la API de configuración.	6 de febrero de 2024
Disponibilidad general de las colecciones de búsquedas vectoriales	Las colecciones de búsquedas vectoriales de Amazon OpenSearch Serverless ya están disponibles de forma general. Durante la fase de versión preliminar se realizaron las siguientes mejoras notables: • Las colecciones de búsquedas vectoriales ahora admiten cargas de trabajo con miles de millones de vectores, cada uno con hasta 128 dimensiones. • OpenSearch Los paneles ahora admiten las colecciones de búsquedas vectoriales.	29 de noviembre de 2023
OR1 instances	Amazon OpenSearch Service ahora admite tipos de OR1 instancias.	29 de noviembre de 2023

Consultas directas con Amazon S3 (versión preliminar)	Las consultas directas proporcionan una solución totalmente gestionada para que los datos transacci onales estén disponibles en Amazon OpenSearch Service en cuestión de segundos después de haberlos escrito en un bucket de Amazon S3.	29 de noviembre de 2023
Capacidad de 10 TiB para colecciones de series temporales	Amazon OpenSearch Serverless añade compatibi lidad con hasta 10 TiB de datos de índice para recopilac iones de series temporale s. Esta versión también admite una capacidad máxima permitida de 200 OCUs para todos los tipos de colecciones y la posibilidad de deshabilitar las réplicas en espera al crear una colección.	29 de noviembre de 2023
OpenSearch Compatibilidad con la versión 2.1.1	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 2.11. Esta versión incluye todas las características que formaban parte de las versiones 2.10 y 2.11. Para obtener más información, consulte las notas de las versiones 2.10 y 2.11.	17 de noviembre de 2023

Soporte OpenSearch de Amazon Ingestion para Data Prepper versión 2.6	Amazon OpenSearch Ingestion añade compatibilidad con la versión 2.6 de Data Prepper. Para obtener más información, consulte las notas de la versión 2.6. Además, puede especificar Amazon DynamoDB como origen de la canalización. Para obtener más información, consulte Uso de una canalización OpenSearch de ingestión con Amazon DynamoDB.	17 de noviembre de 2023
Soporte OpenSearch de Amazon Ingestion para Data Prepper versión 2.5	Amazon OpenSearch Ingestion añade compatibilidad con la versión 2.5 de Data Prepper. Para obtener más información, consulte las notas de la versión 2.5. Además, ahora puede especificar un dominio de OpenSearch servicio o una colección OpenSearch sin servidor como fuente de canalización. Para obtener más información, consulte el complemento OpenSearch fuente en la documentación de Data Prepper.	17 de noviembre de 2023

CloudFormation plantilla para inferencia remota	Para facilitar la configuración de la inferencia remota para la búsqueda semántica, Amazon OpenSearch Service proporciona una AWS CloudFormation plantilla en la consola que automatiza el proceso de aprovisionamiento de modelos.	7 de noviembre de 2023
Actualización de la política de roles vinculados a servicios	Añade los permisos necesarios para que la política `AmazonOpenSearchServiceRolePolicy` de funciones vinculadas al servicio asigne y desasigne direcciones. IPv6 La política `AmazonElasticsearchServiceRolePolicy` de Elasticsearch en desuso también se ha actualizado para garantizar la compatibilidad con versiones anteriores.	26 de octubre de 2023

Políticas de ciclo de vida de Amazon OpenSearch Serverless	Amazon OpenSearch Serverless presenta políticas de ciclo de vida de indexación para agilizar la administración de la retención y eliminación de datos. Ahora puede usar APIs o una interfaz de configuración en la consola para establecer políticas de retención de datos para las recopilaciones de series temporales, lo que elimina la necesidad de crear índices o scripts diarios para eliminar los datos antiguos.	25 de octubre de 2023
Compatibilidad con instancias Im4gn	Amazon OpenSearch Service ahora admite los tipos de instancias de IM4GN. Las instancias iM4GN están optimizadas para cargas de trabajo que gestionan grandes conjuntos de datos y necesitan una alta densidad de almacenamiento por vCPU.	20 de octubre de 2023

Opciones administrativas	Amazon OpenSearch Service ahora ofrece varias opciones administrativas que proporcio nan un control detallado si necesitas solucionar problemas con tu dominio. Estas opciones incluyen la posibilidad de reiniciar el OpenSearch proceso en un nodo de datos y la posibilidad de reiniciar un nodo de datos.	17 de octubre de 2023
Complementos opcionales	Amazon OpenSearch Service añade compatibilidad con cuatro nuevos complementos de análisis de idiomas: Nori (coreano), Sudachi (japonés) , Pinyin (chino) y STConvert Analysis (chino), así como el complemento Amazon Personalize Search Ranking.	16 de octubre de 2023
OpenSearch Soporte para la versión 2.9	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 2.9. Esta versión incluye todas las características que formaban parte de las versiones 2.8 y 2.9. Para obtener más información, consulte las notas de las versiones 2.8 y 2.9.	2 de octubre de 2023

Conectores de ML	Amazon OpenSearch Service añade soporte para conectores de aprendizaje automático (ML). Los conectores facilitan el acceso a los modelos de aprendizaje automático alojados en otras Servicios de AWS plataformas de aprendizaje automático (ML) o en plataformas de aprendizaje automático (ML) de terceros.	6 de septiembre de 2023
Amazon OpenSearch Ingestion añade compatibilidad con la versión 2.4 de Data Prepper	Amazon OpenSearch Ingestion añade compatibilidad con la versión 2.4 de Data Prepper. Para obtener más información, consulte las notas de la versión 2.4. Además, ahora puede especificar Amazon Managed Streaming for Apache Kafka (Amazon MSK) como origen de la canalización.	31 de agosto de 2023
Capacidad de 6 TiB para colecciones de series temporales	Amazon OpenSearch Serverless añade compatibilidad con hasta 6 TiB de datos de índice para recopilaciones de series temporales. Esta versión también admite una capacidad máxima permitida de 100 OCUs para las recopilaciones de series temporales y de búsqueda.	15 de agosto de 2023

Colecciones de búsqueda vectorial	Amazon OpenSearch Serverless añade la opción de crear una colección de búsquedas vectoriales, que puede utilizar para almacenar incrustaciones vectoriales para potenciar las búsquedas semánticas y de similitud.	26 de julio de 2023
OpenSearch Soporte para 2.7	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 2.7. Esta versión incluye todas las características que formaban parte de las versiones 2.6 y 2.7. Para obtener más información, consulte las notas de las versiones 2.6 y 2.7.	10 de julio de 2023
Compatibilidad con Data Prepper 2.3	Amazon OpenSearch Ingestion añade compatibilidad con la versión 2.3 de Data Prepper. Para obtener más información, consulte las notas de la versión 2.3. Además, ahora puede especificar Amazon Security Lake como origen de la canalización.	26 de junio de 2023

Multi-AZ con modo de espera	Amazon OpenSearch Service añade la opción de implement ar un dominio en tres zonas de disponibilidad (AZs), cada zona de disponibilidad contiene una copia completa de los datos y los nodos de una de ellas AZs actúan como zona de espera. La opción de implementación Multi-AZ con modo de espera ofrece una disponibilidad del 99,99 % y un rendimiento uniforme en caso de que se produzca un fallo en la infraestructura.	3 de mayo de 2023
Nuevo rol vinculado a servicio	Amazon OpenSearch Service añade un rol vinculado a un servicio denominad o`AWSServiceRoleForA mazonOpenSearchIng estionService` , que permite a Amazon OpenSearc h Ingestion enviar datos de métricas a. Amazon CloudWatch	26 de abril de 2023

OpenSearch Ingestión de Amazon	Amazon OpenSearch Ingestion es un recopilador de datos totalmente gestionado que proporciona datos de registro y rastreo en tiempo real a los dominios de OpenSearch servicio y a las colecciones OpenSearch sin servidor. OpenSearch Ingestion elimina la necesidad de utilizar soluciones de terceros, como Logstash o Jaeger, para incorporar datos a sus dominios y colecciones.	26 de abril de 2023
OpenSearch Soporte 2.5	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 2.5. Esta versión incluye todas las características que formaban parte de las versiones 2.4 y 2.5. Para obtener más información, consulte las notas de las versiones 2.4 y 2.5.	13 de marzo de 2023

Ventanas de menor actividad para mantenimiento	Amazon OpenSearch Service añade ventanas fuera de las horas pico, que son bloques de tiempo diarios de 10 horas y poco tráfico durante los cuales puede programar las actualizaciones del software del servicio y las optimizac iones de Auto-Tune que requieren una implementación. blue/green Las actualiza ciones en ventanas de menor actividad ayudan a minimizar la carga de los nodos maestros dedicados de un clúster durante los períodos de mayor tráfico. En el caso de los dominios nuevos creados después del 16 de febrero, la ventana de menor actividad se configura automáticamente para que funcione entre las 22:00 y las 8:00, hora local. Para los dominios existentes, debe habilitar la ventana de forma explícita.	16 de febrero de 2023

Configurar la autenticación SAML durante la creación del dominio	Amazon OpenSearch Service ahora admite la configuración de la autenticación SAML durante la creación del dominio. Anteriormente, había que configurar las opciones de SAML una vez que ya se había creado el dominio.	1 de febrero de 2023
Reindexación remota para dominios de VPC	Amazon OpenSearch Service añade la opción de una conexión de punto final de VPC entre dos dominios. Ahora se puede utilizar reindexación remota para copiar índices de un dominio de VPC a otro sin un proxy inverso. Los dominios de VPC deben ejecutar el software de servicio R20221114 o posterior para utilizar esta característica.	31 de enero de 2023

Disponibilidad general de Amazon OpenSearch Serverless	Amazon OpenSearch Serverless ya está disponible de forma general. Durante la fase de versión preliminar se realizaron las siguientes mejoras notables: • La capacidad ahora se puede reducir al mínimo configurado OCUs cuando hay una disminución del tráfico en el punto final de recolección. • El máximo permitido tanto OCUs para la indexación como para la búsqueda se incrementó de 20 a 50. Cada OCU incluye suficiente almacenamiento efímero en caliente para 120 GiB de datos de índice. • Ahora puede configurar los ajustes de acceso a datos mientras crea colecciones, en lugar de tener que configurarlos en un flujo de trabajo independiente.	25 de enero de 2023

Ejecución de prueba asíncrona	Amazon OpenSearch Service ahora admite la ejecución en seco asíncrona, lo que te permite realizar una comprobación de validación antes de realizar un cambio de configuración y te notifica si los cambios van a provocar una implementación. blue/gree n	19 de enero de 2023
Nuevo rol vinculado a servicio	Amazon OpenSearch Service añade un rol vinculado a un servicio denominad o `AWSServiceRoleForA mazonOpenSearchSer verless` , que permite a OpenSearch Serverless enviar datos de métricas a. Amazon CloudWatch	29 de noviembre de 2022
Vista previa de Amazon OpenSearch Serverless	Amazon OpenSearch Serverless es una configura ción bajo demanda, de escalado automático y sin servidores para Amazon OpenSearch Service. Serverless elimina las complejidades operativas del aprovisionamiento, la configuración y el ajuste de los clústeres. OpenSearch	29 de noviembre de 2022

OpenSearch 2.3 soporte	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 2.3. Esta versión incluye todas las características que formaban parte de las versiones 2.0, 2.1 y 2.2. Para obtener más información, consulte las notas de las versiones 2.0, 2.1, 2.2 y 2.3. La versión 2.3 contiene un cambio importante. Para obtener más información, consulte Rutas de actualización admitidas.	15 de noviembre de 2022
Compatibilidad con el complemento de Notificaciones	Amazon OpenSearch Service ahora es compatible con el complemento Notifications, que ofrece una ubicación central para todas las notificaciones de los OpenSearch complementos. A partir de la versión 2.0, los destinos de alerta quedaron obsoletos y se sustituyeron por canales de notificación.	15 de noviembre de 2022

Compatibilidad con Kibana 7.1.1	Los dominios OpenSearch de Amazon Service que ejecutan Elasticsearch 7.1 ahora admiten la última versión del parche para Kibana 7.1.1, que añade correccio nes de errores y mejora la seguridad. Cuando actualice sus dominios 7.1 al software de servicio R20221114, OpenSearch Service los actualizará automáticamente a esta versión de parche.	15 de noviembre de 2022
Compatibilidad con Kibana 6.8.13	Los dominios OpenSearch de Amazon Service que ejecutan Elasticsearch 6.8 ahora admiten la última versión del parche para Kibana 6.8.13, que añade correccio nes de errores y mejora la seguridad. Cuando actualice sus dominios 6.8 al software de servicio R20221114, OpenSearch Service los actualizará automáticamente a esta versión de parche.	15 de noviembre de 2022

Compatibilidad con Kibana 6.3.2	Los dominios OpenSearch de Amazon Service que ejecutan Elasticsearch 6.3 ahora admiten la última versión del parche para Kibana 6.3.2, que añade correccio nes de errores y mejora la seguridad. Cuando actualice sus dominios 6.3 al software de servicio R20221114, OpenSearch Service los actualizará automáticamente a esta versión de parche.	15 de noviembre de 2022

AWS PrivateLink	Con los puntos de enlace OpenSearch de VPC gestionados por Amazon Service, puede conectarse directamente a los dominios de VPC de servicio mediante un punto de enlace de OpenSearch VPC de interfaz en lugar de conectarse a través de Internet. Solo se puede acceder a un punto final de OpenSearch VPC gestionado por el servicio desde la VPC en la que se aprovisiona el punto de conexión o desde cualquier punto de enlace VPCs con la VPC en el que se aprovisio na el punto final, según lo permitan las tablas de enrutamiento y los grupos de seguridad. Su dominio de VPC debe ejecutar el software de servicio R20220928 o posterior para conectarse a un punto de conexión de VPC de interfaz.	7 de noviembre de 2022
Correcciones de errores y mejoras de rendimiento	El software de servicio R20220928 incluye correccio nes de errores y mejoras de rendimiento, incluido un registro SAML mejorado. La actualización también cambia el inquilino predeterminado a `Global` en lugar de `Private`.	3 de octubre de 2022

Referencia de la API mejorada	Amazon OpenSearch Service ofrece una referencia de API de configuración mejorada y completa. Las nuevas referencias contienen todos los tipos de datos y acciones disponibles, ejemplos de sintaxis de solicitudes y respuestas, y enlaces a las referencias de SDK correspon dientes para todos los idiomas admitidos.	13 de septiembre de 2022
Validación azul-verde	Amazon OpenSearch Service ahora realiza una comprobac ión de validación antes de las blue/green implementaciones y muestra errores de validació n si tu dominio no es apto para una actualización.	16 de agosto de 2022
OpenSearch 1.3 soporte	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 1.3. Para obtener más informaci ón, consulte las notas de la versión 1.3.	27 de julio de 2022

Compatibilidad con el complemento ML Commons	Amazon OpenSearch Service añade compatibilidad con el complemento ML Commons, que proporciona un conjunto de algoritmos de aprendizaje automático comunes mediante llamadas a la API de transport e y REST. También puede interactuar con el complemen to ML Commons a través de comandos de PPL.	27 de julio de 2022
Compatibilidad con los volúmenes gp3	Amazon OpenSearch Service añade compatibilidad con el tipo de volumen SSD de uso general de gp3 EBS. Puede especificar las IOPS aprovisionadas y el rendimien to adicionales al crear o modificar el dominio.	26 de julio de 2022
Documentación de prácticas recomendadas mejoradas	La documentación OpenSearc h de Amazon Service proporciona mejores prácticas operativas mejoradas y recomendaciones generales para crear y operar dominios OpenSearch de servicio.	6 de julio de 2022
Integración con Service Quotas	Ahora puedes ver las cuotas de Amazon OpenSearch Service y solicitar aumentos de cuota desde la consola Service Quotas.	29 de junio de 2022

Control de acceso a la API basado en OpenSearch etiquetas	Ahora puede usar etiquetas para controlar el acceso a. OpenSearch APIs Anteriorm ente, solo podía utilizar etiquetas para controlar el acceso a la API de configura ción.	16 de junio de 2022
Búsqueda entre clústeres en regiones	Ahora se admite la búsqueda entre clústeres Regiones de AWS siempre que ambos dominios ejecuten la versión 7.10 o posterior de Elasticse arch, o cualquier versión de. OpenSearch	14 de junio de 2022
Compatibilidad con Kibana 5.6 individual	Amazon OpenSearch Service añade soporte para una versión única de Kibana 5.6.16. Con Kibana 5.6.16 individual, puede utilizar Kibana 5.6 como front-end mientras se conecta a las versiones 5.1, 5.3, 5.5 y 5.6 de Elasticsearch. Debe estar en el software del servicio R20220323 o posterior para utilizar Kibana 5.6 individual.	4 de abril de 2022

R20220323-P1	Amazon OpenSearch Service publicó recientemente la actualización del software de servicio R20220323, pero la actualización se anuló posteriormente debido a un problema. Se recomiend a actualizar los dominios a la versión de revisión R20220323-P1 o posterior, que soluciona el problema.	4 de abril de 2022
OpenSearch 1.2 soporte	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 1.2. Para obtener más informaci ón, consulte las notas de la versión 1.2.	4 de abril de 2022
Observabilidad	La instalación predeterminada de OpenSearch Dashboard s for Amazon OpenSearch Service incluye el complemen to Observability, que puedes usar para visualizar eventos basados en datos mediante el lenguaje de procesami ento canalizado (PPL) para explorar y consultar tus datos. El complemento requiere la versión OpenSearch 1.2 o una versión posterior y el software de servicio R20220323 o una versión posterior.	4 de abril de 2022

Compatibilidad con Kibana 7.7.1	Los dominios OpenSearch de Amazon Service que ejecutan Elasticsearch 7.7 ahora admiten la última versión del parche para Kibana 7.7, que añade correccio nes de errores y mejora la seguridad. Cuando actualice sus dominios 7.7 al software de servicio R20220323 o posterior, OpenSearch Service los actualizará automátic amente a esta versión de parche.	4 de abril de 2022

Cambios en la métrica de presión de memoria de JVM	Amazon OpenSearch Service ha cambiado la lógica de las `JVMMemoryPressure` CloudWatch métricas para reflejar con mayor precisión el uso de la memoria. Anteriormente, las métricas solo tenían en cuenta el grupo de memoria de antigua generación del montón de JVM. Con este cambio, la métrica también tiene en cuenta el grupo de memoria de nueva generación. Después de actualizar el dominio al software del servicio R20220323, es posible que observe un incremento en las métricas `JVMMemoryPressure` , `MasterJVMMemoryPressure` o `WarmJVMMemoryPressure` .	4 de abril de 2022
Diccionarios personalizados con el complemento IK (Chinese) Analysis	Amazon OpenSearch Service ahora admite el uso de diccionarios personalizados con el complemento IK (chino) Analysis.	4 de abril de 2022

Replicación entre clústeres en dominios existentes	Amazon OpenSearch Service eliminó la limitación de que solo se pueden implementar la búsqueda y la replicación entre clústeres en dominios creados a partir del 3 de junio de 2020. Ahora puede habilitar estas caracterí sticas en todos los dominios, independientemente de cuándo se hayan creado. Ambos dominios deben estar en el software del servicio R20220323 o posterior.	4 de abril de 2022
Visibilidad de implementación azul-verde	Amazon OpenSearch Service ahora ofrece más visibilidad del progreso de las implement aciones azul/verde. Puede supervisar estos detalles en la consola o mediante la API de configuración.	27 de enero de 2022

Control de acceso detallado en dominios existentes	Ahora puede habilitar el control de acceso detallado en los dominios existentes. Puede habilitar un período de migración temporal para las políticas de acceso abiertas o basadas en IP para garantizar que los usuarios puedan seguir accediendo a su dominio mientras crea y asigna roles. La habilitación del control de acceso detallado en los dominios existentes requiere el software de servicio R20211203 o posterior.	6 de enero de 2022
Funciones de Dashboards renombradas OpenSearch	Con el software de servicio R20211203, se le ha cambiado el nombre al rol `kibana_user` por `opensearch_dashboards_user`, y a `kibana_read_only` se le ha cambiado el nombre por `opensearch_dashboards_read_only`. Este cambio se aplica a todos los 1 recién creados OpenSearch . x dominios. En el caso de OpenSearch los dominios existentes que se actualicen al software de servicio R20211203, las funciones siguen siendo las mismas.	4 de enero de 2022

OpenSearch 1.1 soporte	Amazon OpenSearch Service ahora es compatible con OpenSearch la versión 1.1. Para obtener más informaci ón, consulte las notas de la versión 1.1.	4 de enero de 2022
Editor visual de ISM	La instalación predeterminada de OpenSearch Dashboard s for Amazon OpenSearc h Service ahora admite el editor visual de las políticas de ISM. Esta función requiere la versión OpenSearch 1.1 o una versión posterior.	4 de enero de 2022
Actualización de la prevenció n del suplente confuso entre servicios	Amazon OpenSearch Service admite el uso de las claves de contexto de condición `aws:SourceAccount` global `aws:SourceArn` y las claves de contexto en las políticas de recursos de IAM para evitar el confuso problema de los diputados . Debe estar en el software de servicio R20211203 o posterior para utilizar estas claves de condición.	4 de enero de 2022

Parche de Log4j	El software de servicio R20211203-P2 actualiza la versión de Log4j utilizada en Service, tal como se recomienda en las recomendaciones de los documentos CVE-2021-44228 y OpenSearch CVE-2021-45046. El parche se aplica a los dominios que ejecutan todas las versiones OpenSearch de Elasticsearch y de Elasticsearch. OpenSearch El servicio seguirá actualizando varias versiones de Log4j internamente y no se restringirán necesariamente a la última versión de Log4j. La versión de Log4j de su dominio depende de la versión de software que ejecute el dominio. Sin embargo, independientemente de la versión de Log4j, siempre que ejecute R20211203-P2 o posterior, sus dominios contienen la actualización de Log4j necesaria para abordar CVE-2021-44228 y CVE-2021-45046.	15 de diciembre de 2021

Replicación entre clústeres	La replicación entre clústeres le permite replicar índices, mapeos y metadatos de un dominio de servicio a otro. OpenSearch La replicación entre clústeres requiere un dominio que ejecute Elasticsearch 7.10 o 1.1 o una versión posterior. OpenSearch	5 de octubre de 2021
AWS Nuevas políticas administradas	El lanzamiento de Amazon OpenSearch Service incluye nuevas políticas AWS gestionadas y la obsolescencia de las políticas antiguas.	8 de septiembre de 2021
Compatibilidad con Kibana 6.4.3	Los dominios OpenSearch de Amazon Service que ejecutan la versión 6.4 antigua de Elasticsearch ahora son compatibles con la última versión del parche para Kibana 6.4, que añade correcciones de errores y mejora la seguridad. OpenSearch El servicio actualizará automáticamente los dominios a esta versión del parche.	8 de septiembre de 2021

Flujos de datos	Amazon OpenSearch Service añade soporte para flujos de datos, lo que simplific a el proceso de gestión de datos de series temporales. Su dominio debe ejecutar la OpenSearch versión 1.0 o una versión posterior para utilizar los flujos de datos.	8 de septiembre de 2021
OpenSearch Servicio Amazon	AWS cambia el nombre de Amazon OpenSearch Service para eliminar la antigua marca «Elasticsearch». Amazon OpenSearch Service es compatible con OpenSearch Elasticsearch OSS heredado. Al crear un clúster, tiene la opción de elegir qué motor de búsqueda utilizar. OpenSearc h El servicio ofrece una amplia compatibilidad con Elasticse arch OSS 7.10, la versión final de código abierto del software.	8 de septiembre de 2021

Almacenamiento en frío	El almacenamiento en frío es un nuevo nivel de almacenamiento para datos históricos o a los que se accede con poca frecuencia. Los índices almacenados en frío solo ocupan el almacenamiento de S3 y no tienen informática vinculada a ellos. El almacenamiento en frío requiere un dominio que ejecute Elasticsearch 7.9 o posterior y el software de servicio R20210426 o posterior.	13 de mayo de 2021
Instancias de Graviton basadas en ARM	Amazon OpenSearch Service ahora admite los tipos de instancias de Graviton basados en ARM (M6G, C6G, R6G y R6GD). Los tipos de instancias de Graviton están disponibles en dominios nuevos y existentes que ejecutan Elasticsearch 7.9 o posterior y el software de servicio R20210331 o posterior.	4 de mayo de 2021

Plantillas de ISM	Amazon OpenSearch Service añade compatibilidad con las plantillas de ISM, que te permiten adjuntar automáticamente una política de ISM a un índice si el índice coincide con un patrón definido en la política. Las plantillas de ISM requieren el software de servicio R20210426 o posterior. Esta actualización también deja obsoleta la configuración `policy_id`, que significa que ya no puede utilizar plantillas de índice para aplicar políticas de ISM a los índices recién creados. La actualización introduce un cambio radical en las CloudFormation plantillas existentes que utilizan esta configuración.	27 de abril de 2021
Compatibilidad con Elasticsearch 7.10	Amazon OpenSearch Service ahora es compatible con la versión 7.10 de Elasticsearch. Para obtener más información, consulte las notas de la versión de 7.10.	21 de abril de 2021

Búsqueda asíncrona	Amazon OpenSearch Service ahora admite la búsqueda asíncrona, lo que te permite ejecutar solicitudes de búsqueda en segundo plano. La búsqueda asíncrona requiere un dominio que ejecute Elasticsearch 7.10 o posterior, y el software de servicio R20210331 o posterior.	21 de abril de 2021
Control de acceso basado en etiquetas para la API de configuración	Ahora puede usar AWS etiquetas para controlar el acceso a la API de configuración de Amazon ES.	2 de marzo de 2021
Ajuste automático	Amazon OpenSearch Service añade Auto-Tune, que utiliza las métricas de rendimiento y uso del clúster para sugerir cambios en la configuración de JVM de los nodos. El ajuste automático requiere un dominio que ejecute Elasticsearch 6.7 o posterior y el software de servicio R20201117 o posterior.	24 de febrero de 2021

Trace Analytics	La instalación predeterminada de Kibana for Amazon OpenSearch Service ahora incluye el complemento de análisis de rastreo, que te permite monitorear los datos de rastreo de tus aplicaciones distribuidas. El complemento requiere un dominio que ejecute Elasticsearch 7.9 o posterior y un software de servicio R20210201 o posterior.	17 de febrero de 2021
Métricas de particiones	Amazon OpenSearch Service añade las siguientes CloudWatch métricas para el seguimiento del estado de los fragmentos:`Shards.active` ,`Shards.unassigned` ,, `Shards.delayedUnassigned` `Shards.activePrimary` `Shards.initializing` ,`Shards.relocating` . Las métricas están disponibles en dominios con el software de servicio R20210201 o posterior.	17 de febrero de 2021

Informes de Kibana	La instalación predeterminada de Kibana for Amazon OpenSearch Service ahora admite informes bajo demanda para las páginas Discover, Visualize y Dashboard. Esta característica requiere Elasticsearch 7.9 o posterior y el software de servicio R20210201 o posterior.	17 de febrero de 2021
Compatibilidad con Kibana 5.6.16	Los dominios OpenSearch de Amazon Service que ejecutan Elasticsearch 5.6 ahora admiten la última versión del parche para Kibana 5.6, que añade correcciones de errores y mejora la seguridad. Amazon ES actualizará de forma automática los dominios a esta versión de parche.	17 de febrero de 2021
Cifrado para dominios existentes	Amazon OpenSearch Service ahora permite habilitar el cifrado de los datos en reposo y el node-to-node cifrado en los dominios existentes que ejecutan Elasticsearch 6.7 o una versión posterior. Después de habilitar esta configuración, no podrá deshabilitarla.	27 de enero de 2021

Reindexación remota	Amazon OpenSearch Service ahora admite la reindexación remota, lo que te permite migrar índices desde dominios remotos. Esta característica requiere el software de servicio R20201117 o una versión posterior.	24 de noviembre de 2020
Lenguaje de procesamiento con plecas	Amazon OpenSearch Service ahora es compatible con el lenguaje de procesamiento canalizado (PPL), un lenguaje de consultas que te permite usar la sintaxis pipe (|) para consultar los datos almacenados en Elasticsearch. Esta característica requiere el software de servicio R20201117 o una versión posterior. Para obtener más información, consulte .	24 de noviembre de 2020
Cuadernos de Kibana	Amazon OpenSearch Service añade compatibilidad con las libretas Kibana, lo que te permite combinar visualizaciones en directo y texto narrativo en una única interfaz. Esta característica requiere el software de servicio R20201117 o una versión posterior.	24 de noviembre de 2020

Diagramas de Gantt	La instalación predeterminada de Kibana for Amazon OpenSearch Service ahora admite un nuevo tipo de visualización, los diagramas de Gantt. Esta característica requiere el software de servicio R20201117 o una versión posterior.	24 de noviembre de 2020
Compatibilidad con Elasticsearch 7.9	Amazon OpenSearch Service ahora es compatible con la versión 7.9 de Elasticsearch. Para obtener más información, consulte las notas de la versión de 7.9.	24 de noviembre de 2020
Actualizaciones de detección de anomalías	La detección de anomalías para Amazon OpenSearch Service añade compatibilidad con la alta cardinalidad, lo que te permite clasificar las anomalías con una dimensión como la dirección IP, el identificador del producto, el código de país, etc. Esta característica requiere el software de servicio R20201117 o una versión posterior.	24 de noviembre de 2020

Actualizaciones de diccionario dinámicas	Amazon OpenSearch Service ahora te permite actualizar tus analizadores de búsqueda sin tener que volver a indexarlos. Puede actualizar los archivos de diccionario en algunos o todos sus dominios y Amazon ES realiza un seguimien to de las versiones de los paquetes a lo largo del tiempo para que tenga un historial de lo que cambió y cuándo. Esta caracterí stica requiere el software de servicio R20201019 o una versión posterior.	17 de noviembre de 2020
Puntos de conexión personali zados	Amazon OpenSearch Service ahora admite puntos de enlace personalizados, lo que te permite asignar una nueva URL a tu dominio de Amazon ES. Si alguna vez intercamb ia dominios, puede mantener la misma URL. Esta caracterí stica requiere el software de servicio R20201019 o una versión posterior.	5 de noviembre de 2020

Nuevos complementos de idioma	Amazon OpenSearch Service ahora admite los complementos IK (Chinese) Analysis, Vietnamese Analysis y Thai Analysis en dominios que ejecutan Elasticsearch 7.7 o posterior con el software de servicio R20201019 o posterior.	28 de octubre de 2020
Compatibilidad con Elasticsearch 7.8	Amazon OpenSearch Service ahora es compatible con la versión 7.8 de Elasticsearch. Para obtener más información, consulte las notas de la versión de 7.8.	28 de octubre de 2020
Autenticación SAML para Kibana	Amazon OpenSearch Service ahora admite la autenticación SAML para Kibana, que te permite usar proveedores de identidad de terceros para iniciar sesión en Kibana, administrar un control de acceso detallado, buscar tus datos y crear visualizaciones. Esta característica requiere el software de servicio R20201019 o una versión posterior.	27 de octubre de 2020
Instancias T3	Amazon OpenSearch Service ahora admite los tipos de instancia `t3.medium` y `t3.small`.	23 de septiembre de 2020

Registros de auditoría	Amazon OpenSearch Service ahora admite los registros de auditoría de tus datos, lo que te permite realizar un seguimiento de los intentos fallidos de inicio de sesión, del acceso de los usuarios a índices, documentos y campos, y mucho más. Esta característica requiere el software de servicio R20200910 o una versión posterior.	16 de septiembre de 2020
UltraWarm actualizaciones	UltraWarm for Amazon OpenSearch Service añade nuevas métricas, nuevos ajustes, una cola de migración más grande y una API de cancelación. Estas actualizaciones requieren el software de servicio R20200910 o posterior. Para obtener más información, consulte .	14 de septiembre de 2020
Learning to Rank	Amazon OpenSearch Service ahora es compatible con el complemento de código abierto Learning to Rank, que te permite utilizar tecnologías de aprendizaje automático para mejorar la relevancia de las búsquedas. Esta característica requiere el software de servicio R20200721 o una versión posterior.	27 de julio de 2020

Similitud coseno de k-NN	k-Nearest Neighbor (k-NN) ahora permite buscar “vecinos más cercanos” por similitud coseno además de la distancia euclidiana. Esta caracterí stica requiere el software de servicio R20200721 o una versión posterior.	23 de julio de 2020
Compresión con gzip	Amazon OpenSearch Service ahora admite la compresió n gzip para la mayoría de las solicitudes y respuestas HTTP, lo que puede reducir la latencia y conservar el ancho de banda. Esta caracterí stica requiere el software de servicio R20200721 o una versión posterior.	23 de julio de 2020
Compatibilidad con Elasticse arch 7.7	Amazon OpenSearch Service ahora es compatible con la versión 7.7 de Elasticsearch. Para obtener más informaci ón, consulte las notas de la versión de 7.7.	23 de julio de 2020
Servicio de mapas de Kibana	La instalación predeterm inada de Kibana for Amazon OpenSearch Service ahora incluye un servidor de mapas WMS, excepto para los dominios de las regiones de India y China.	18 de junio de 2020

Mejoras de SQL	El soporte de SQL para Amazon OpenSearch Service ahora admite muchas operaciones nuevas, una interfaz de usuario de Kibana dedicada para la exploración de datos y una CLI interactiva.	3 de junio de 2020
Búsqueda en clústeres	Amazon OpenSearch Service te permite realizar consultas y agregaciones entre clústeres en varios dominios conectados.	3 de junio de 2020
Detección de anomalías	Amazon OpenSearch Service te permite detectar automáticamente las anomalías casi en tiempo real.	3 de junio de 2020
UltraWarm	UltraWarm El almacenamiento de Amazon OpenSearch Service ha salido de la versión preliminar pública y ahora está disponible de forma general. La función ahora es compatible con una gama más amplia de versiones y Regiones de AWS. Para obtener más información, consulte .	5 de mayo de 2020

Diccionarios personalizados	Amazon OpenSearch Service te permite cargar archivos de diccionarios personalizados para usarlos con tu clúster. Estos archivos mejoran los resultados de la búsqueda al indicarle a Elasticsearch que ignore ciertas palabras de alta frecuencia o que trate los términos como equivalentes.	21 de abril de 2020
Compatibilidad con Elasticsearch 7.4	Amazon OpenSearch Service ahora es compatible con la versión 7.4 de Elasticsearch. Para obtener más información, consulte Versiones compatibles.	12 de marzo de 2020
k-NN	Amazon OpenSearch Service añade compatibilidad con la búsqueda del vecino más cercano (k-NN). K-nn requiere el software de servicio R20200302 o posterior.	3 de marzo de 2020
Index State Management	Amazon OpenSearch Service añade Index State Management (ISM), que te permite automatizar tareas rutinarias, como eliminar índices cuando llegan a cierta edad. Esta característica requiere el software de servicio R20200302 o una versión posterior.	3 de marzo de 2020

Compatibilidad con Elasticsearch 5.6.16	Amazon OpenSearch Service ahora es compatible con la última versión del parche de la versión 5.6, que añade correcciones de errores y mejora la seguridad. Amazon ES actualizará de forma automática los dominios 5.6 existentes a esta versión. Tenga en cuenta que esta versión de Elasticsearch informa incorrectamente su versión como 5.6.17.	2 de marzo de 2020
Control de acceso detallado	Amazon OpenSearch Service ahora admite un control de acceso detallado, que ofrece seguridad a nivel de índice, documento y campo, la multitenencia de Kibana y la autenticación básica HTTP opcional para su clúster.	11 de febrero de 2020

UltraWarm almacenamiento (versión preliminar)	Amazon OpenSearch Service añade UltraWarm un nuevo nivel de almacenamiento en caliente que utiliza Amazon S3 y una sofisticada solución de almacenamiento en caché para mejorar el rendimiento. Para los índices en los que no está escribiendo activamente y consultando con menos frecuencia, el UltraWarm almacenamiento ofrece costos por GiB significativamente más bajos.	3 de diciembre de 2019
Características de cifrado para las regiones de China	El cifrado de los datos en reposo y el node-to-node cifrado ya están disponibles en las regiones de `cn-north-1` China (Pekín) y `cn-northwest-1` China (Ningxia).	20 de noviembre de 2019
Exigir HTTPS	Ahora puede exigir que todo el tráfico a sus dominios de Amazon ES llegue a través de HTTPS. Al configurar su dominio, marque la casilla Exigir HTTPS. Esta característica requiere el software de servicio R20190808 o una versión posterior.	3 de octubre de 2019

Compatibilidad con Elasticsearch 7.1 y 6.8	Amazon OpenSearch Service ahora es compatible con las versiones 7.1 y 6.8 de Elasticsearch. Para obtener más información, consulte Versiones compatibles.	13 de agosto de 2019
Instantáneas cada hora	En lugar de instantáneas diarias, Amazon OpenSearch Service ahora toma instantáneas cada hora de los dominios que ejecutan Elasticsearch 5.3 y versiones posteriores para que tengas copias de seguridad más frecuentes desde las que restaurar tus datos.	8 de julio de 2019
Compatibilidad con Elasticsearch 6.7	Amazon OpenSearch Service ahora es compatible con la versión 6.7 de Elasticsearch. Para obtener más información, consulte Versiones compatibles.	29 de mayo de 2019
Compatibilidad con SQL	Amazon OpenSearch Service ahora te permite consultar tus datos mediante SQL. La compatibilidad con SQL requiere el software de servicio R20190418 o una versión posterior.	15 de mayo de 2019

Tipos de instancias de la serie 5	Amazon OpenSearch Service ahora admite los tipos de instancias M5, C5 y R5. En comparación con los tipos de instancias de la generación anterior, estos nuevos tipos ofrecen un mejor rendimien to a precios más bajos. Para obtener más información, consulte Límites.	24 de abril de 2019
Compatibilidad con Elasticse arch 6.5	Amazon OpenSearch Service ahora es compatible con la versión 6.5 de Elasticsearch.	8 de abril de 2019
Alertas	Las alertas para Amazon OpenSearch Service le notifican cuando los datos de uno o más índices de Amazon ES cumplen determinadas condiciones. Las alertas requieren el software de servicio R20190221 o una versión posterior.	25 de marzo de 2019
Compatibilidad con tres zonas de disponibilidad	Amazon OpenSearch Service ahora admite tres zonas de disponibilidad en muchas regiones. Esta versión incluye también una experiencia de consola simplificada. Este multi-AZ requiere el software de servicio R20181023 o una versión posterior.	7 de febrero de 2019

Compatibilidad con Elasticsearch 6.4	Amazon OpenSearch Service ahora es compatible con la versión 6.4 de Elasticsearch.	23 de enero de 2019
Clústeres de 200 nodos	Ahora Amazon ES permite crear clústeres con hasta 200 nodos de datos y un total de 3 PB de almacenamiento.	22 de enero de 2019
Actualizaciones del software del servicio	Amazon ES ahora permite actualizar manualmente el software de servicio de su dominio con el fin de obtener beneficios de nuevas características con mayor rapidez o realizar las actualizaciones en el momento en que hay menos tráfico. Para obtener más información, consulte .	20 de noviembre de 2018
Nuevas métricas CloudWatch	Amazon ES ahora ofrece métricas a nivel de nodo y las pestañas Estado del clúster e Estado de instancia en la consola de Amazon ES.	20 de noviembre de 2018
Compatibilidad en China (Pekín)	Amazon OpenSearch Service ya está disponible en la región cn-north-1, donde admite los tipos de instancias M4, C4 y R4.	17 de octubre de 2018

Node-to-node cifrado	Amazon OpenSearch Service ahora admite el node-to-node cifrado, lo que mantiene los datos cifrados a medida que Amazon ES los distribuye por todo el clúster.	18 de septiembre de 2018
Actualizaciones de versiones in situ	Amazon OpenSearch Service ahora admite actualizaciones de versiones locales.	14 de agosto de 2018
Compatibilidad con Elasticsearch 6.3 y 5.6	Amazon OpenSearch Service ahora es compatible con las versiones 6.3 y 5.6 de Elasticsearch.	14 de agosto de 2018
Registros de errores	Amazon ES ahora le permite publicar los registros de errores de Elasticsearch en Amazon. CloudWatch	31 de julio de 2018
Instancias reservadas en China (Ningxia)	Amazon ES ahora ofrece instancias reservadas en la región China (Ningxia).	29 de mayo de 2018
Instancias reservadas	Ahora Amazon ES ofrece compatibilidad con Instancias reservadas.	7 de mayo de 2018

Actualizaciones anteriores

En la siguiente tabla se describen los cambios importantes en Amazon ES realizados antes de mayo de 2018.

Cambio	Descripción	Fecha
Autenticación de Amazon Cognito para Kibana	Amazon ES ahora ofrece protección de la página de inicio de sesión para Kibana. Para obtener más información, consulte the section called “Autenticación de Amazon Cognito para paneles OpenSearch ”.	2 de abril de 2018
Compatibilidad con Elasticsearch 6.2	Amazon OpenSearch Service ahora es compatible con la versión 6.2 de Elasticsearch.	14 de marzo de 2018
Complemento Korean Analysis	Amazon ES ahora admite una versión optimizada para memoria del complemento del análisis coreano Seunjeon.	13 de marzo de 2018
Actualizaciones de control de acceso instantáneas	Los cambios en las políticas de control de acceso en los dominios de Amazon ES ahora se aplican al instante.	7 de marzo de 2018
Escala de petabytes	Amazon ES admite ahora tipos de instancias I3 y una capacidad total de almacenamiento del dominio de hasta 1,5 PB. Para obtener más información, consulte the section called “Escala de petabytes”.	19 de diciembre de 2017
Cifrado de datos en reposo	Amazon ES ahora admite el cifrado de datos en reposo. Para obtener más información, consulte the section called “Cifrado en reposo”.	7 de diciembre de 2017
Compatibilidad con Elasticsearch 6.0	Amazon ES ahora admite la versión 6.0 de Elasticse arch. Para ver consideraciones e instrucciones sobre las migraciones, consulte the section called “Actualización de dominios”.	6 de diciembre de 2017
Compatibilidad con VPC	Amazon ES ahora permite lanzar dominios dentro de una Amazon Virtual Private Cloud. La compatibilidad con VPC proporciona una capa de seguridad adicional y simplifica las comunicaciones entre Amazon ES y otros servicios dentro de una VPC. Para obtener más informaci ón, consulte the section called “Compatibilidad con VPC”.	17 de octubre de 2017

Cambio	Descripción	Fecha
Publicación de registros lentos	Amazon ES ahora admite la publicación de registros lentos en CloudWatch Logs. Para obtener más información, consulte the section called “Monitoreo de registros”.	16 de octubre de 2017
Compatibilidad con Elasticsearch 5.5	Amazon ES ahora admite la versión 5.5 de Elasticsearch. Ahora puede restablecer las instantáneas automatizadas sin contactar con Soporte y almacenar scripts con la API de `_scripts`.	7 de septiembre de 2017
Compatibilidad con Elasticsearch 5.3	Amazon ES agregó compatibilidad a la versión 5.3 de Elasticsearch.	1 de junio de 2017
Más instancias y capacidad de EBS por clúster	Amazon ES ahora admite hasta 100 nodos y 150 TB de capacidad de EBS por clúster.	5 de abril de 2017
Compatibilidad en Canadá (Central) y la UE (Londres)	Amazon ES agregó compatibilidad en las siguiente s regiones: Canadá (Central), ca-central-1 y Europa (Londres), eu-west-2.	20 de marzo de 2017
Más instancias y mayores volúmenes de EBS	Amazon ES incorpora la compatibilidad con más instancias y mayores volúmenes de EBS.	21 de febrero de 2017
Compatibilidad con Elasticsearch 5.1	Amazon ES agregó compatibilidad a la versión 5.1 de Elasticsearch.	30 de enero de 2017
Compatibilidad con el complemento Phonetic Analysis	Amazon ES lleva incorporada la integración con el complemento Phonetic Analysis, que permite ejecutar consultas con el operador “sounds-like” (sonido similar a) en los datos.	22 de diciembre de 2016
Compatibilidad con la región EE. UU Este (Ohio)	Amazon ES incorpora compatibilidad con la siguiente región: Este de EE. UU. (Ohio), us-east-2.	17 de octubre de 2016

Cambio	Descripción	Fecha
Nueva métrica de rendimiento	Amazon ES agregó una métrica de rendimiento, `ClusterUsedSpace` .	29 de julio de 2016
Compatibilidad con Elasticsearch 2.3	Amazon ES agregó compatibilidad con la versión 2.3 de Elasticsearch.	27 de julio de 2016
Compatibilidad con la región Asia Pacífico (Mumbai)	Amazon ES agregó compatibilidad a la siguiente región: Asia-Pacífico (Mumbai), ap-south-1.	27 de junio de 2016
Más instancias por clúster	Amazon ES aumenta de 10 a 20 el número máximo de instancias (recuento de instancias) por clúster.	18 de mayo de 2016
Compatibilidad con la región Asia Pacífico (Seúl)	Amazon ES agregó compatibilidad a la siguiente región: Asia-Pacífico (Seúl), ap-northeast-2.	28 de enero de 2016
Amazon ES	Versión inicial.	1 de octubre de 2015

AWS Glosario

Para obtener la AWS terminología más reciente, consulte el AWS glosario de la Glosario de AWS Referencia.

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la version original de inglés, prevalecerá la version en inglés.