aws

Ferramentas de alertas e monitoramento e as práticas recomendadas do Amazon RDS para MySQL e MariaDB

AWS Orientação prescritiva

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

AWS Orientação prescritiva: Ferramentas de alertas e monitoramento e as práticas recomendadas do Amazon RDS para MySQL e MariaDB

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

As marcas comerciais e imagens de marcas da Amazon não podem ser usadas no contexto de nenhum produto ou serviço que não seja da Amazon, nem de qualquer maneira que possa gerar confusão entre os clientes ou que deprecie ou desprestigie a Amazon. Todas as outras marcas comerciais que não pertencem à Amazon pertencem a seus respectivos proprietários, que podem ou não ser afiliados, patrocinados pela Amazon ou ter conexão com ela.

Table of Contents

Ferramentas de alertas e monitoramento e as práticas recomendadas do Amazon RDS para MySQL e MariaDB

Igor Obradovic, Amazon Web Services (AWS)

Março de 2025 (histórico do documento)

O monitoramento de banco de dados é o processo de medir, rastrear e avaliar a disponibilidade, a performance e a funcionalidade de um banco de dados. As soluções de monitoramento e alerta ajudam as organizações a garantir que seus serviços de banco de dados e, portanto, suas aplicações e workloads associadas, sejam seguros, resilientes e eficientes. Ativado AWS, você pode coletar e analisar seus registros de carga de trabalho, métricas, eventos e rastreamentos para entender a integridade de sua carga de trabalho e obter insights das operações ao longo do tempo.

Você pode monitorar seus recursos para garantir que estejam funcionando conforme o esperado e para detectar e remediar quaisquer problemas antes que eles afetem seus clientes. Você deve usar as métricas, os logs, os eventos e os rastreamentos que monitora para gerar alarmes quando os limites são violados.

Este guia descreve as ferramentas de monitoramento e observabilidade do banco de dados e as práticas recomendadas para bancos de dados do Amazon Relational Database Service (Amazon RDS). O guia se concentra nos bancos de dados MySQL e MariaDB, embora a maioria das informações também se aplique a outros mecanismos de banco de dados do Amazon RDS.

Este guia é para arquitetos de soluções, arquitetos de banco de dados DBAs, DevOps engenheiros seniores e outros membros da equipe que se dedicam ao projeto, implementação e gerenciamento de soluções de monitoramento e observabilidade para suas cargas de trabalho de banco de dados em execução no. Nuvem AWS

Índice

Visão geral

O monitoramento e o alerta estão incluídos em quatro pilares do AWS Well-Architected Framework.

- O pilar de excelência operacional determina que sua workload deve ser projetada para incluir telemetria e monitoramento. Os serviços da AWS, como o Amazon Relational Database Service (Amazon RDS), fornecem as informações necessárias para que você entenda o estado interno da sua workload (tais como métricas, logs, eventos e rastreamento). Ao operar seus bancos de dados do Amazon RDS, você desejará entender a integridade das instâncias de seus bancos de dados, detectar eventos operacionais e ser capaz de responder a eventos planejados e não planejados. A AWS fornece ferramentas de monitoramento que ajudam a determinar quando os resultados organizacionais e comerciais estão em risco ou podem estar em risco, para que você possa tomar as medidas apropriadas no momento certo.
- O pilar de eficiência de performance determina que você deve monitorar a performance de seus recursos, como instâncias de banco de dados do Amazon RDS, reunindo, agregando e processando métricas relacionadas à performance em tempo real. Você pode identificar a degradação da performance e remediar os fatores, por exemplo, consultas SQL não otimizadas ou parâmetros de configuração inadequados, que a causaram. Você pode acionar alarmes automaticamente quando as medições estão fora dos limites esperados. Recomendamos que você use alarmes não apenas para notificações, mas também para iniciar ações automatizadas em resposta aos eventos detectados. Você pode avaliar as métricas coletadas em relação a limites predefinidos ou usar algoritmos de machine learning para identificar comportamentos anômalos. Por exemplo, para detectar uma tendência de aumento na utilização da CPU, você pode coletar e analisar a métrica `cpuUtilization.total` por um período. Alertar sobre essa anomalia de forma proativa, antes que a utilização da CPU atinja o limite máximo, pode ajudar você a remediar o problema antes que ele afete seus clientes.
- O pilar de confiabilidade define o monitoramento e os alertas como essenciais para garantir que você esteja cumprindo seus requisitos de disponibilidade. Sua solução de monitoramento deve ser capaz de detectar falhas de forma eficaz. Ao detectar problemas ou falhas, seu objetivo principal é alertar sobre essas questões. Implementar práticas contínuas de observabilidade e monitoramento é fundamental para arquiteturas resilientes na nuvem. Para melhorar suas workloads, você deve ser capaz de avaliá-las e entender seu estado e integridade. Os princípios de design para recuperação automática de uma falha, escalabilidade horizontal e provisionamento de capacidade dependem de serviços precisos de monitoramento e alerta.

- O pilar de segurança discute a detecção e prevenção de alterações de configuração inesperadas ou indesejadas e comportamento inesperado. Você pode configurar suas instâncias de banco de dados do Amazon RDS para MySQL e MariaDB com o plug-in de auditoria do MariaDB para registrar atividades do banco de dados, como logins de usuários e operações específicas executadas no banco de dados. O plug-in armazena o registro da atividade do banco de dados em um arquivo de logs, que pode ser integrado e importado para ferramentas de monitoramento e alerta. O arquivo de logs é analisado em tempo real para detectar comportamentos inesperados ou suspeitos em seu banco de dados. Esse comportamento inesperado ou suspeito pode indicar que sua instância de banco de dados do Amazon RDS foi comprometida, o que sinaliza riscos potenciais para sua empresa. Se a ferramenta de monitoramento detectar esse evento, ela ativará um alarme para iniciar uma resposta ao incidente de segurança, o que ajuda a lidar com atividades suspeitas e maliciosas.

Resultados de negócios desejados

A implementação das práticas recomendadas em mecanismos de monitoramento e alerta ajuda você a garantir uma infraestrutura de alta performance, resiliente, eficiente, segura e com otimização de custos para suas aplicações e workloads. Você pode usar ferramentas de observabilidade que coletam, armazenam e visualizam métricas, eventos, rastreamentos e logs em tempo real para observar e analisar o panorama geral da integridade e da performance de seus bancos de dados e, assim, evitar a degradação ou interrupção dos serviços de TI associados. Se a degradação não planejada ou a interrupção do serviço ainda ocorrer, as ferramentas de monitoramento e alerta ajudarão você a detectar o problema em tempo hábil, escalar, reagir e investigar e resolver rapidamente. Uma solução abrangente de monitoramento e alerta para suas workloads de banco de dados em nuvem ajuda você a alcançar os seguintes resultados de negócios:

- Melhorar a experiência do cliente. Um serviço confiável melhora as experiências de seus clientes. Os bancos de dados geralmente são um componente essencial dos serviços digitais, como aplicações da web e móveis, streaming de mídia, pagamentos, APIs business-to-business (B2B) e serviços de integração. Se você puder monitorar e configurar alertas em seus bancos de dados para detectar problemas rapidamente, investigá-los com eficiência e remediá-los o mais rápido possível para minimizar o tempo de inatividade e outras interrupções, poderá melhorar a disponibilidade, a segurança e a performance do serviço digital para seus clientes.
- Conquistar a confiança do cliente.Uma melhor performance e uma experiência de usuário sem atritos ajudam você a conquistar a confiança de seus clientes, o que pode resultar em mais negócios em sua plataforma. Por exemplo, um provedor de serviços de processamento de

pagamentos que oferece um serviço on-line confiável pode esperar uma alta confiança e fidelidade dos clientes, o que resulta em mais clientes e melhor retenção, um aumento nas transações faturáveis e serviços novos e inovadores que geram mais receita.

- Evitar perdas financeiras.Qualquer tempo de inatividade inesperado em sua infraestrutura de banco de dados pode afetar as transações comerciais que seus clientes realizam usando sua aplicação. Isso pode levar a perdas financeiras substanciais em alguns casos. A violação dos acordos de serviço (SLAs) pode resultar na perda da confiança do cliente e, consequentemente, na perda de receita. Também pode se tornar uma base legal para testes caros, em que os clientes podem exigir compensação com base em seus contratos de responsabilidade e garantia. De acordo com um estudo da Atlassian Corporation, uma empresa de software, os custos médios da interrupção do serviço estão na faixa de USD 140 mil a USD 540 mil por hora, dependendo do tipo e porte da empresa. Um ambiente de banco de dados estável é essencial para evitar interrupções prolongadas e perda de negócios.
- Expandir o valor. Mecanismos de monitoramento e alerta podem ajudar você a projetar, desenvolver e operar um serviço digital altamente disponível, resiliente, confiável, de alta performance, econômico e seguro, mas isso é só o começo. Você vai querer que sua organização cresça e prospere com o tempo, aprimore as workloads existentes na nuvem e introduza novos serviços. Novos serviços fornecem valor adicional para seus clientes e mais receita para sua empresa, criando um efeito de aceleração contínua no crescimento da sua empresa.
- Melhorar a produtividade do desenvolvedor.Desenvolvedores que são produtivos e eficientes, e que não encontram problemas e gargalos em suas tarefas de desenvolvimento, podem entregar produtos de alta qualidade em menos tempo. No entanto, a engenharia de software e as operações de TI geralmente têm desafios complexos, e essa complexidade aumenta com a escala das workloads e suas arquiteturas. Para analisar a performance e a consistência em aplicações distribuídas, os desenvolvedores precisam de ferramentas que possam fornecer métricas e rastreamentos correlacionados. Elas ajudam a identificar artefatos de código e componentes de infraestrutura com defeito o mais rápido possível, e ajudam a determinar os impactos nos usuários finais. O conjunto certo de ferramentas de monitoramento e alerta pode ajudar os desenvolvedores a programar e testar de forma melhor e mais rápida.
- Melhorar a eficácia e a eficiência operacionais.Quando você opera workloads na nuvem em grande escala, até mesmo uma pequena porcentagem de melhorias na performance pode resultar em economias de milhões de dólares. Ao monitorar seus bancos de dados e analisar métricas, eventos, logs e rastreamentos, você pode entender e prever suas necessidades futuras de capacidade e aproveitar as economias de custo disponíveis na Nuvem AWS. Compreender as

workloads e a integridade operacional do Amazon RDS pode ajudar você a responder a eventos, corrigir problemas e planejar melhorias.

Práticas recomendadas gerais

As práticas recomendadas a seguir ajudam você a obter visibilidade suficiente da integridade da sua workload do Amazon RDS e a tomar as medidas apropriadas em resposta a eventos operacionais e dados de monitoramento.

- Identificar os KPIs. Identifique indicadores-chave de performance (KPIs) com base nos resultados empresariais desejados. Avalie os KPIs para determinar o sucesso da workload. Por exemplo, se seu negócio principal é comércio eletrônico, um dos resultados comerciais desejados pode ser que sua loja virtual esteja disponível 24 horas por dia, 7 dias por semana, para que seus clientes façam suas compras. Para alcançar esse resultado comercial, você define o KPI de disponibilidade para o banco de dados de backend do Amazon RDS que sua aplicação de loja virtual usa, e define o KPI de linha de base para 99,99% semanalmente. Avaliar o KPI de disponibilidade real em relação ao valor de linha de base ajuda a determinar se você está atingindo a disponibilidade desejada do banco de dados de 99,99% e, assim, alcançando o resultado comercial de oferecer um serviço 24 horas por dia, 7 dias por semana.
- Definir as métricas da workload. Defina as métricas da workload para avaliar as quantidades e as qualidades da sua workload do Amazon RDS. Avalie as métricas para determinar se a workload está alcançando os resultados desejados e para entender a integridade da workload. Por exemplo, para avaliar o KPI de disponibilidade da sua instância de banco de dados Amazon RDS, você deve avaliar métricas como tempo de atividade e tempo de inatividade da instância de banco de dados. Você pode então usar essas métricas para calcular o KPI de disponibilidade da seguinte forma:

  ```
  availability = uptime / (uptime + downtime)
  ```

  As métricas representam conjuntos de pontos de dados ordenados cronologicamente. As métricas também podem incluir dimensões, que são úteis na categorização e análise.
- Coletar e analisar métricas da workload. O Amazon RDS gera diferentes métricas e registros, dependendo da sua configuração. Algumas delas representam eventos, contadores ou estatísticas de instâncias de banco de dados, como `db.Cache.innoDB_buffer_pool_hits`. Outras métricas vêm do sistema operacional, como `memory.Total`, que mede a quantidade total de memória da instância host do Amazon Elastic Compute Cloud (Amazon EC2). A ferramenta de monitoramento deve realizar análises regulares e proativas das métricas coletadas para identificar tendências e determinar se alguma resposta apropriada é necessária.

- Estabelecer linhas de base de métricas da workload. Estabeleça linhas de base para as métricas a fim de definir valores esperados e identificar limites bons ou ruins. Por exemplo, você pode definir a linha de base para `ReadIOPS` para ser de até 1.000 em operações normais de banco de dados. Você pode então usar essa linha de base para comparação e para identificar a utilização excessiva. Se suas novas métricas mostrarem consistentemente que as IOPS de leitura estão na faixa de 2.000 a 3.000, você identificou um desvio que poderia desencadear uma resposta para investigação, intervenção e melhoria.
- Alertar quando os resultados da workload estiverem em risco. Ao determinar que o resultado comercial está em risco, emita um alerta. Você pode então abordar os problemas de forma proativa, antes que eles afetem seus clientes, ou mitigar o impacto do incidente em tempo hábil.
- Identificar os padrões de atividade esperados para sua workload. Com base nas linhas de base de suas métricas, estabeleça padrões de atividade da workload para identificar comportamentos inesperados e responder com ações apropriadas, se necessário. A AWS fornece ferramentas de monitoramento que aplicam algoritmos estatísticos e de machine learning para analisar métricas e detectar anomalias.
- Alertar quando forem detectadas anomalias na workload. Quando forem detectadas anomalias nas operações das workloads do Amazon RDS, emita um alerta para que você possa responder com as ações apropriadas, se necessário.
- Analisar e revisar KPIs e métricas. Confirme se seus bancos de dados do Amazon RDS atendem aos requisitos definidos e identifique áreas de possíveis melhorias para alcançar suas metas comerciais. Valide a eficácia das métricas medidas e dos KPIs avaliados, e revise-os se necessário. Por exemplo, digamos que você defina um KPI para o número ideal de conexões simultâneas de banco de dados e monitore métricas relacionadas a tentativas e falhas de conexões, bem como a threads de usuário que foram criados e estão em execução. Você pode ter mais conexões de banco de dados do que as definidas pela sua linha de base de KPI. Ao analisar suas métricas atuais, você pode detectar o resultado, mas talvez não consiga determinar a causa raiz. Nesse caso, você deve revisar suas métricas e incluir outras medidas de monitoramento, como contadores para bloqueios de tabela. As novas métricas ajudariam a determinar se o aumento do número de conexões de banco de dados é causado por bloqueios inesperados de tabelas.

Ferramentas de monitoramento

Recomendamos que você use ferramentas de observabilidade, monitoramento e alerta para:

- Obter insights sobre a performance do seu ambiente do Amazon RDS
- Detectar comportamentos inesperados e suspeitos
- Planejar a capacidade e tomar decisões informadas sobre a alocação de instâncias do Amazon RDS
- Analisar métricas e logs para prever possíveis problemas de forma proativa
- Gerar alertas quando os limites forem violados para solucionar e resolver problemas antes que seus usuários sejam afetados

Você tem diferentes opções e soluções para escolher, incluindo ferramentas e serviços de observabilidade e monitoramento nativos da nuvem fornecidos pela AWS; soluções de software de código aberto gratuitas; e soluções comerciais de terceiros para monitorar instâncias de banco de dados do Amazon RDS. Algumas dessas ferramentas são analisadas nas seções a seguir.

Para determinar qual ferramenta atende melhor às suas necessidades, compare os recursos e as capacidades de cada ferramenta com os requisitos da sua organização. Também recomendamos que você avalie as ferramentas quanto à facilidade de implantação, configuração e integração, atualizações e manutenção de software, método de implantação (por exemplo, hardware ou sem servidor), licenciamento, preço e quaisquer outros fatores específicos de sua organização.

Seções

Ferramentas incluídas no Amazon RDS

O Amazon Relational Database Service (Amazon RDS) é um serviço de banco de dados gerenciado na Nuvem AWS. Como o Amazon RDS é um serviço gerenciado, ele libera você da maioria das tarefas de gerenciamento, como backups de banco de dados, instalações de sistema operacional (SO) e software de banco de dados, aplicação de patches de sistemas operacionais e software, configuração de alta disponibilidade, ciclo de vida de hardware e operações de data centers. A AWS também fornece um conjunto abrangente de ferramentas que permitem criar uma solução completa de observabilidade para suas instâncias de banco de dados Amazon RDS.

Algumas das ferramentas de monitoramento estão incluídas, pré-configuradas e habilitadas automaticamente no serviço do Amazon RDS. Duas ferramentas automatizadas estão disponíveis para você assim que você inicia sua nova instância do Amazon RDS:

- O status da instância do Amazon RDS fornece detalhes sobre a integridade atual da sua instância de banco de dados. Por exemplo, os códigos de status incluem Available, Stopped, Creating, Backing-up e Failed. Você pode usar o console do Amazon RDS, a AWS Command Line Interface (AWS CLI) ou a API do Amazon RDS para ver o status das instâncias. Para obter as informações, consulte Visualizar o status de uma instância de banco de dados do Amazon RDS na documentação do Amazon RDS.
- As recomendações do Amazon RDS fornecem recomendações automatizadas para instâncias de banco de dados, réplicas de leitura e grupos de parâmetros de banco de dados. Essas recomendações são fornecidas pela análise do uso de instâncias de bancos de dados, dos dados de performance e da configuração, e são entregues como orientação. Por exemplo, a recomendação da versão desatualizada do Engine sugere que suas instâncias de banco de dados não estão executando a versão mais recente do software de banco de dados e que você deve atualizar sua instância de banco de dados para se beneficiar das correções de segurança e outras melhorias mais recentes. Para obter as informações, consulte Visualizar as recomendações Amazon RDS na documentação do Amazon RDS.

Namespaces do CloudWatch

O Amazon RDS se integra ao Amazon CloudWatch, que é um serviço de monitoramento e alerta para recursos e aplicações em nuvem executados na AWS. O Amazon RDS coleta automaticamente métricas, arquivos de logs, rastreamentos e eventos sobre a operação, a utilização, a performance

e a integridade das instâncias de banco de dados e os envia ao CloudWatch para armazenamento, análise e alertas de longo prazo.

O Amazon RDS para MySQL e o Amazon RDS para MariaDB publicam automaticamente um conjunto padrão de métricas no CloudWatch em intervalos de um minuto, sem custo adicional. Essas métricas são coletadas em dois namespaces, que são contêineres para métricas:

- O namespace AWS/RDS inclui métricas em nível de instância de banco de dados. Os exemplos incluem `BinLogDiskUsage` (a quantidade de espaço em disco ocupado pelos logs binários), `CPUUtilization` (a porcentagem de utilização da CPU), `DatabaseConnections` (o número de conexões de rede do cliente com a instância de banco de dados) e muito mais.
- O namespace AWS/Usage inclui métricas de uso em nível de conta, que são usadas para determinar se você está operando dentro das cotas de serviço do Amazon RDS. Os exemplos incluem `DBInstances` (o número de instâncias de banco de dados em sua conta ou região da AWS), `DBSubnetGroups` (o número de grupos de sub-redes de banco de dados em sua conta ou região da AWS) e `ManualSnapshots` (o número de snapshots de banco de dados criados manualmente em sua conta ou região da AWS).

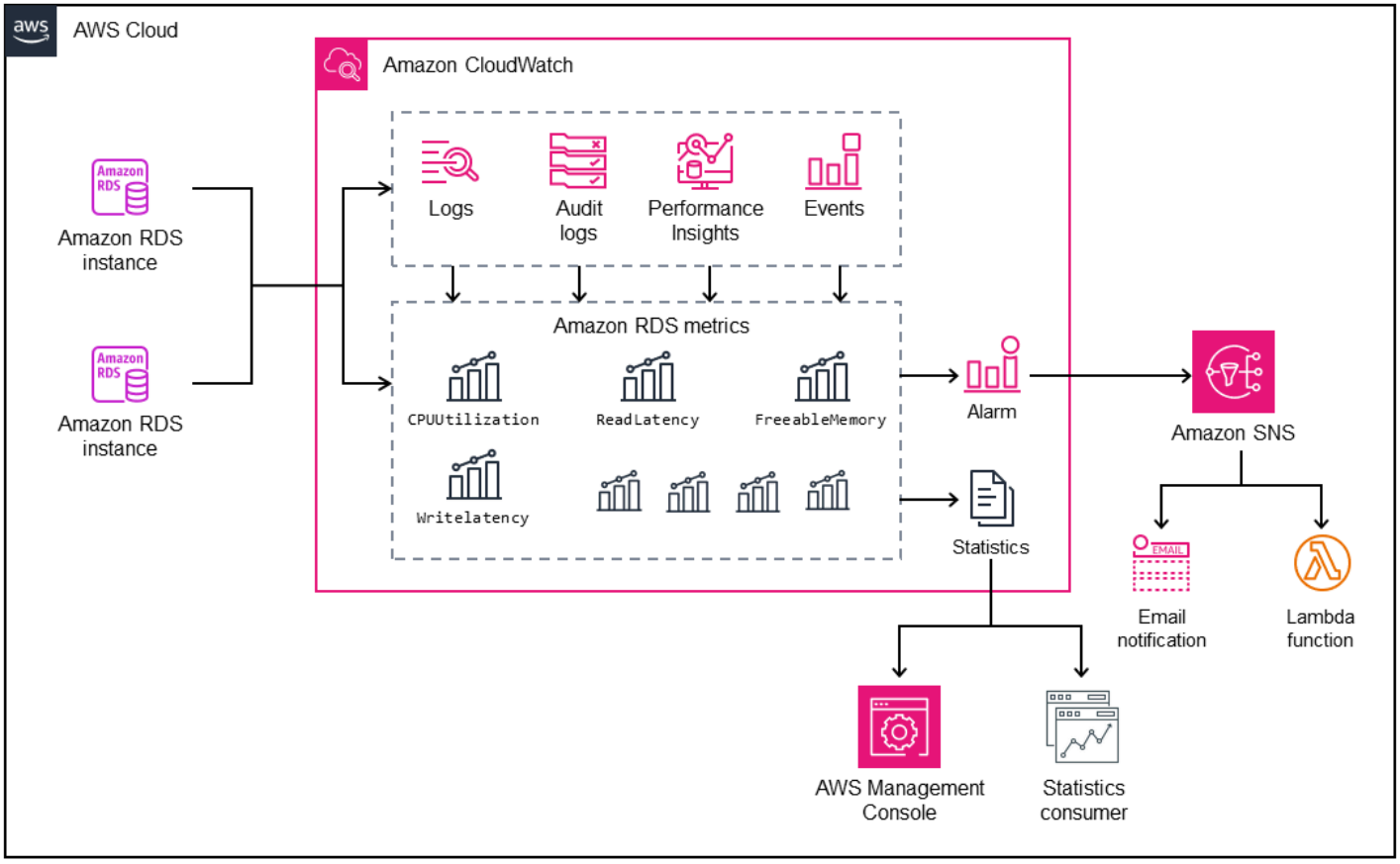
O CloudWatch mantém os dados de métrica da seguinte forma:

- 3 horas: métricas personalizadas de alta resolução com um período inferior a 60 segundos são mantidas por 3 horas. Depois de 3 horas, os pontos de dados são agregados em métricas de período de 1 minuto e mantidos por 15 dias.
- 15 dias: pontos de dados com um período de 60 segundos (1 minuto) ficam retidos por 15 dias. Depois de 15 dias, os pontos de dados são agregados em métricas de período de 5 minutos e são mantidos por 63 dias.
- 63 dias: pontos de dados com um período de 300 segundos (5 minutos) ficam retidos por 63 dias. Depois de 63 dias, os pontos de dados são agregados em métricas de período de 1 minuto e são mantidos por 15 meses.
- 15 meses: pontos de dados com um período de 3.600 segundos (1 hora) ficam disponíveis por 15 meses (455 dias).

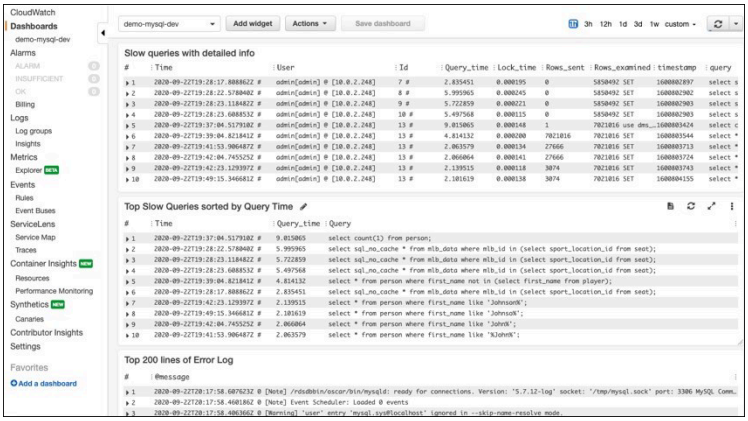
Para obter mais informações, consulte Métricas na documentação do CloudWatch.

Alarmes e painéis do CloudWatch

Você pode usar alarmes do Amazon CloudWatch para monitorar uma métrica específica do Amazon RDS por um período. Por exemplo, você pode monitorar `FreeStorageSpace` e, em seguida, realizar uma ou mais ações se o valor da métrica ultrapassar o limite definido. Se você definir o limite para 250 MB e o espaço de armazenamento gratuito for 200 MB (menos do que o limite), o alarme será ativado e poderá acionar uma ação para provisionar automaticamente armazenamento adicional para a instância de banco de dados do Amazon RDS. O alarme também pode enviar uma notificação por SMS para o DBA usando o Amazon Simple Notification Service (Amazon SNS). O diagrama a seguir ilustra esse processo.

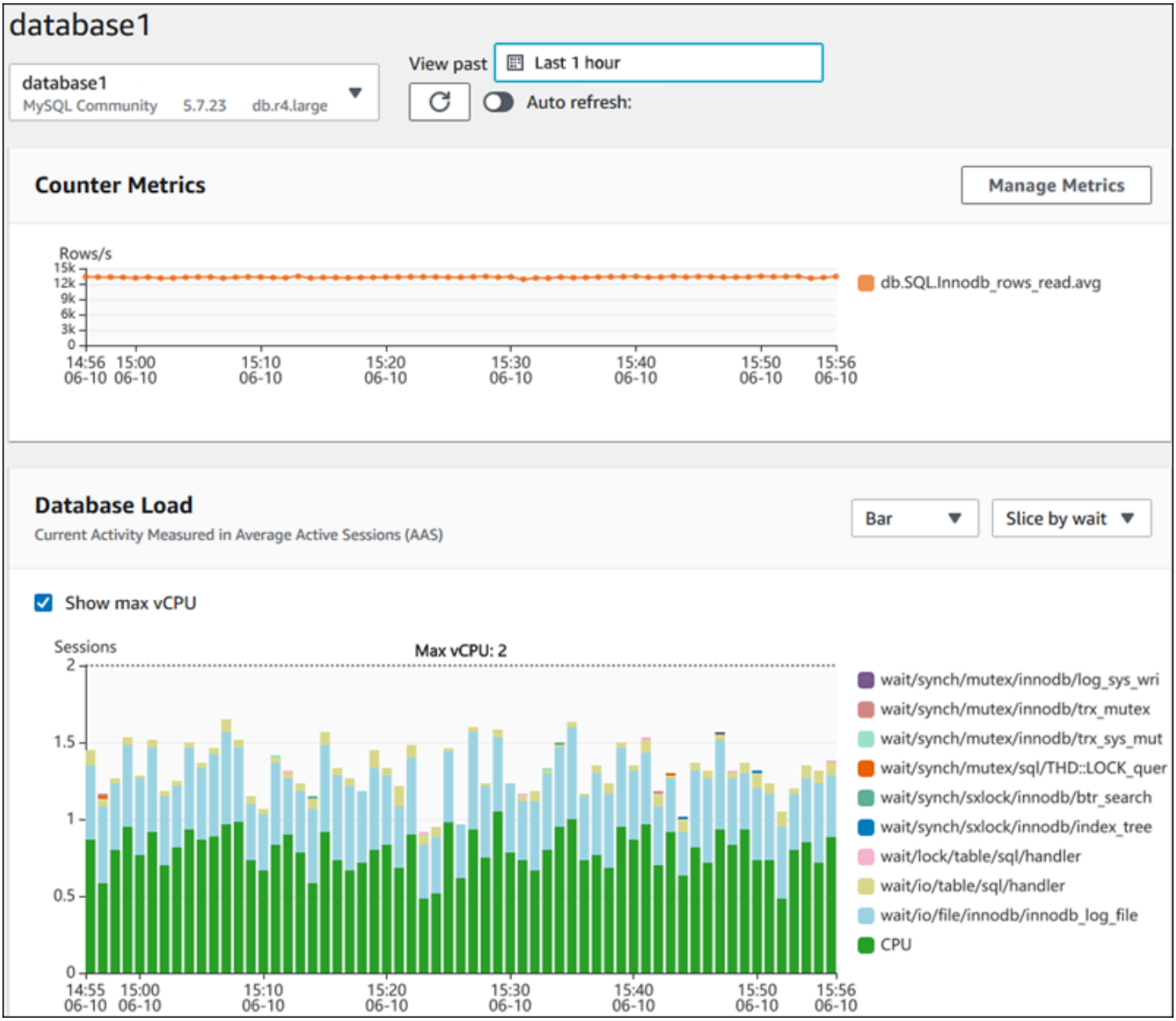


O CloudWatch também fornece painéis, que você pode usar para criar, personalizar, interagir e salvar visualizações personalizadas (grafos) das métricas. Você também pode usar o CloudWatch Logs Insights para criar um painel para monitorar o log de consultas lentas e o log de erros, além de receber alertas se um padrão específico for detectado nesses logs. A tela a seguir mostra um exemplo de painel do CloudWatch.



Amazon RDS Performance Insights

O Insights de Performance do Amazon RDS é uma ferramenta de ajuste e monitoramento de performance do banco de dados que expande os recursos de monitoramento do Amazon RDS. Ele ajuda você a analisar a performance do seu banco de dados, visualizando a carga das instâncias de bancos de dados e filtrando-a por esperas, instruções SQL, hosts ou usuários. A ferramenta combina várias métricas em um único grafo interativo que ajuda a identificar o tipo de gargalo que sua instância de banco de dados pode ter, como esperas de bloqueio, alto consumo de CPU ou latência de E/S, e determinar quais instruções SQL estão criando o gargalo. A tela a seguir mostra um exemplo de visualização.



Você precisa habilitar o Insights de Performance durante o processo de criação da instância de banco de dados para coletar métricas para as instâncias de banco de dados do Amazon RDS em sua conta. O nível gratuito inclui sete dias de histórico de dados de performance e um milhão de solicitações de API por mês. Opcionalmente, você pode comprar períodos de retenção mais longos. Para obter informações completas sobre custos, consulte Definição de preço do Performance Insights.

Para obter informações sobre como você pode usar o Insights de Performance para monitorar suas instâncias de banco de dados, consulte a seção de Monitoramento de instâncias de bancos de dados mais adiante neste guia.

O Insights de Performance publica automaticamente as métricas no CloudWatch. Além de usar a ferramenta Insights de Performance, você pode aproveitar os recursos adicionais que o CloudWatch fornece. Você pode examinar as métricas do Insights de Performance usando o console do CloudWatch, a AWS CLI ou a API do CloudWatch. Você também pode adicionar alarmes do CloudWatch, como acontece com qualquer outra métrica. Por exemplo, você talvez queira acionar uma notificação por SMS para DBAs ou executar uma ação corretiva se a métrica `DBLoad` ultrapassar o valor limite definido. Você também pode adicionar as métricas do Insights de Performance aos seus painéis existentes do CloudWatch.

Monitoramento avançado

O Monitoramento Aprimorado é uma ferramenta que captura métricas em tempo real para o sistema operacional (SO) em que a instância de banco de dados do Amazon RDS é executada. Essas métricas fornecem granularidade de até um segundo para CPU, memória, processos do Amazon RDS e do sistema operacional, sistema de arquivos e dados de E/S de disco, entre outros. Você pode acessar e analisar essas métricas no console do Amazon RDS. Assim como no Insights de Performance, as métricas do Monitoramento Aprimorado são fornecidas do Amazon RDS para o CloudWatch, onde você pode se beneficiar de recursos adicionais, como a preservação de longo prazo de métricas para análise, a criação de filtros de métricas, a exibição de grafos no painel do CloudWatch e a configuração de alarmes. Por padrão, o Monitoramento Aprimorado é desabilitado quando você cria uma nova instância de banco de dados do Amazon RDS. Você pode habilitar o recurso ao criar ou modificar uma instância de banco de dados. Os preços são baseados na quantidade de dados transferidos do Amazon RDS para o CloudWatch Logs e nas taxas de armazenamento. Dependendo da granularidade e do número de instâncias de banco de dados em que o Monitoramento Aprimorado está habilitado, parte dos dados de monitoramento pode ser incluída no nível gratuito do CloudWatch Logs. Para obter mais detalhes sobre os preços, consulte Definição de preço do Amazon CloudWatch. Para obter mais informações sobre a ferramenta, consulte a documentação do Amazon RDS e as perguntas frequentes sobre Monitoramento Aprimorado.

Serviços adicionais da AWS

A AWS fornece vários serviços de apoio, que também se integram ao Amazon RDS e ao CloudWatch, para aprimorar ainda mais a observabilidade de seus bancos de dados. Inclui o Amazon EventBridge, o Amazon CloudWatch Logs e o AWS CloudTrail.

- O Amazon EventBridge é um barramento de eventos sem servidor que pode receber, filtrar, transformar, rotear e entregar eventos de suas aplicações e recursos da AWS, incluindo suas instâncias de banco de dados do Amazon RDS. Um evento do Amazon RDS indica uma alteração no ambiente do Amazon RDS. Por exemplo, quando uma instância de banco de dados muda seu status de Avaiable para Stopped, o Amazon RDS gera o evento `RDS-EVENT-0087 / The DB instance has been stopped`. O Amazon RDS entrega eventos ao CloudWatch Events e ao EventBridge quase em tempo real. Usando o EventBridge e o CloudWatch Events, você pode definir regras para enviar alertas sobre eventos de interesse específicos do Amazon RDS e automatizar ações a serem executadas quando um evento corresponder à regra. Várias destinos estão disponíveis em resposta a um evento, como uma função do AWS Lambda que pode executar uma ação corretiva ou um tópico do Amazon SNS que pode enviar um e-mail ou SMS para notificar os DBAs ou engenheiros de DevOps sobre o evento.
- O Amazon CloudWatch Logs é um serviço que centraliza o armazenamento de arquivos de logs de todas as suas aplicações, sistemas e serviços da AWS, incluindo instâncias do Amazon RDS para banco de dados MySQL e MariaDB e o AWS CloudTrail. Se você habilitar o recurso para suas instâncias de banco de dados, o Amazon RDS publicará automaticamente os seguintes logs no CloudWatch Logs:
  - Log de erros
  - Log de consultas lentas
  - Log geral
  - Log de auditoria

  Você pode usar o CloudWatch Logs Insights para consultar e analisar os dados de logs. O recurso inclui uma linguagem de consulta específica que ajuda você a pesquisar eventos de logs que correspondam aos padrões definidos por você. Por exemplo, você pode rastrear a corrupção de tabelas em sua instância de banco de dados MySQL monitorando o arquivo de logs de erros em busca do seguinte padrão: `"ERROR 1034 (HY000): Incorrect key file for table '*'; try to repair it OR Table * is marked as crashed"`. Os dados de logs filtrados podem ser convertidos em métricas do CloudWatch. Você pode então usar as métricas para criar painéis com grafos ou dados tabulares ou definir um alarme se o valor limite definido for violado. Isso é particularmente útil ao usar o log de auditoria, pois você pode monitorar, enviar alertas e executar ações corretivas automaticamente se algum comportamento inesperado ou suspeito for detectado. Você pode acessar e gerenciar os logs do banco de dados usando o Console de Gerenciamento da AWS, a AWS CLI, a API do Amazon RDS ou o AWS SDK do CloudWatch Logs.

- O AWS CloudTrail registra em log e monitora continuamente a atividade do usuário e da API em sua Conta da AWS. Ele ajuda você com a auditoria, o monitoramento de segurança e a solução de problemas operacionais de suas instâncias do Amazon RDS para banco de dados MySQL ou MariaDB. O CloudTrail está integrado ao Amazon RDS. Todas as ações podem ser registradas em log, e o CloudTrail fornece um registro das ações executadas por um usuário, perfil ou serviço da AWS no Amazon RDS. Por exemplo, quando um usuário cria uma nova instância de banco de dados do Amazon RDS, um evento é detectado, e o log inclui informações sobre a ação solicitada (`"eventName": "CreateDBInstance"`), a data e hora da ação (`"eventTime": "2022-07-30T22:14:06Z"`), os parâmetros de solicitação (`"requestParameters": {"dBInstanceIdentifier": "test-instance", "engine": "mysql", "dBInstanceClass": "db.m6g.large"}`) e assim por diante. Os eventos registrados em log pelo CloudTrail incluem as chamadas do console do Amazon RDS e as chamadas de código que usa as APIs do Amazon RDS.

Ferramentas de monitoramento de terceiros

Em alguns cenários, além do conjunto completo de ferramentas de monitoramento e observabilidade nativas da nuvem que a AWS fornece para o Amazon RDS, você talvez queira usar ferramentas de monitoramento de outros provedores de software. Esses cenários incluem implantações híbridas, em que você pode ter vários bancos de dados em execução em seu data center on-premises e outro conjunto de bancos de dados em execução na Nuvem AWS. Se você já estabeleceu sua solução corporativa de observabilidade, talvez queira continuar usando suas ferramentas existentes e estendê-las às suas implantações da Nuvem AWS. O desafio de configurar uma solução de monitoramento de terceiros geralmente está nas proteções impostas pelo Amazon RDS como um serviço gerenciado na nuvem. Por exemplo, você não pode instalar o software do agente no sistema operacional host que executa a instância de banco de dados, porque o acesso à máquina host do banco de dados é negado. No entanto, você pode integrar várias soluções de monitoramento de terceiros com o Amazon RDS usando o CloudWatch e outros serviços da Nuvem AWS. Por exemplo, métricas, logs, eventos e rastreamentos do Amazon RDS podem ser exportados e depois importados para a ferramenta de monitoramento de terceiros para análise, visualização e alertas adicionais. Algumas dessas soluções de terceiros incluem o Prometheus, Grafana e Percona.

Prometheus e Grafana

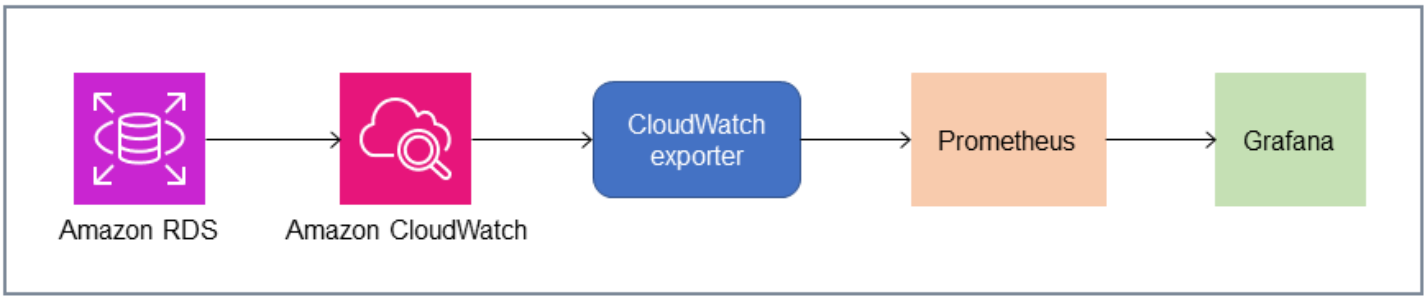
O Prometheus é uma solução de monitoramento de código aberto que coleta métricas de destinos configurados em determinados intervalos. É uma solução de monitoramento de uso geral que pode

monitorar qualquer aplicação ou serviço. Quando você monitora instâncias de banco de dados do Amazon RDS, o CloudWatch coleta as métricas do Amazon RDS. As métricas são então exportadas para o servidor Prometheus usando um exportador de código aberto, como o YACE Exporter ou CloudWatch Exporter.

- O YACE Exporter otimiza as tarefas de exportação de dados recuperando várias métricas em uma única solicitação para a API do CloudWatch. Depois que as métricas são armazenadas no servidor Prometheus, ele avalia as expressões de regras e pode gerar alertas quando condições especificadas são observadas.
- O CloudWatch Exporter é mantido oficialmente pelo Prometheus. Ele recupera as métricas do CloudWatch por meio da API do CloudWatch e as armazena no servidor Prometheus em um formato compatível com o Prometheus, usando solicitações da API REST para o endpoint HTTP.

Ao escolher um exportador, projetar seu modelo de implantação e configurar instâncias exportadoras, considere as cotas de serviços e APIs do CloudWatch e do CloudWatch Logs, pois a exportação das métricas do CloudWatch para um servidor Prometheus é implementada na API do CloudWatch. Por exemplo, implantar várias instâncias do CloudWatch Exporter em uma única região e Conta da AWS para monitorar centenas de instâncias de banco de dados do Amazon RDS pode resultar em um erro de controle de utilização (ThrottlingException) e erros de código 400. Para superar essas limitações, considere usar o YACE Exporter, que é otimizado para coletar até 500 métricas diferentes em uma única solicitação. Além disso, para implantar um grande número de instâncias de banco de dados do Amazon RDS, você deve considerar usar várias Contas da AWS, em vez de centralizar a workload em uma única Conta da AWS e limitar o número de instâncias do exportador em cada Conta da AWS.

Os alertas são gerados pelo servidor Prometheus e gerenciados pelo Alertmanager. Essa ferramenta se encarrega de desduplicar, agrupar e rotear alertas para o destinatário correto, como e-mail, SMS ou Slack, ou iniciar uma ação de resposta automática. Outra ferramenta de código aberto chamada Grafana exibe as visualizações dessas métricas. O Grafana fornece widgets de visualização avançados, como grafos avançados, painéis dinâmicos e recursos de analytics, como consultas ad-hoc e detalhamento dinâmico. Ele também pode pesquisar e analisar logs e inclui recursos de alerta para avaliar continuamente métricas e logs e enviar notificações quando os dados corresponderem às regras de alerta.



Percona

O Percona Monitoring and Management (PMM) é uma solução gratuita de monitoramento, gerenciamento e observabilidade de banco de dados de código aberto para MySQL e MariaDB. O PMM coleta milhares de métricas de performance de instâncias de banco de dados e seus hosts. Ele fornece uma interface de usuário da web para visualizar dados em painéis e recursos adicionais, como consultores automáticos para avaliações de integridade do banco de dados. Você pode usar o PMM para monitorar o Amazon RDS. No entanto, o cliente PMM (agente) não está instalado nos hosts subjacentes das instâncias de banco de dados do Amazon RDS porque não tem acesso aos hosts. Em vez disso, a ferramenta se conecta às instâncias de banco de dados do Amazon RDS, consulta as estatísticas do servidor, o `INFORMATION_SCHEMA`, o esquema do sistema e o Esquema de Performance, e usa a API do CloudWatch para adquirir métricas, logs, eventos e rastreamentos. O PMM exige uma chave de acesso do usuário do AWS Identity and Access Management (IAM) (perfil do IAM) e descobre automaticamente as instâncias de banco de dados do Amazon RDS que estão disponíveis para monitoramento. A ferramenta PMM tem um perfil para monitoramento de banco de dados e coleta mais métricas específicas do banco de dados do que o Prometheus. Para usar o painel Query Analytics do PMM, você deve configurar o Esquema de Performance como a fonte da consulta, pois o agente do Query Analytics não está instalado para o Amazon RDS e não consegue ler o log de consultas lentas. Em vez disso, ele consulta diretamente o `performance_schema` das instâncias dos bancos de dados MySQL e MariaDB para obter métricas. Uma das características proeminentes do PMM é sua capacidade de alertar e aconselhar os DBAs sobre problemas que a ferramenta identifica em seus bancos de dados. O PMM oferece conjuntos de verificações que podem detectar ameaças comuns à segurança, degradação da performance e perda e corrupção de dados.

Além dessas ferramentas, há várias soluções comerciais de observabilidade e monitoramento disponíveis no mercado que podem ser integradas ao Amazon RDS. Alguns exemplos incluem Datadog Database Monitoring, Dynatrace Amazon RDS Monitoring e AppDynamics Database Monitoring.

Monitoramento de instâncias de bancos de dados

Um instância de banco de dados é o elemento básico do Amazon RDS. É um ambiente de banco de dados isolado executado na nuvem. Para bancos de dados MySQL e MariaDB, a instância de banco de dados é o programa mysqld, também conhecido como servidor MySQL, que inclui vários threads e componentes, como o analisador SQL, o otimizador de consultas, o thread/connection manipulador, as variáveis de sistema e de status e um ou mais mecanismos de armazenamento conectáveis. Cada mecanismo de armazenamento foi projetado para oferecer suporte a um caso de uso especializado. O mecanismo de armazenamento padrão e recomendado é o InnoDB, que é um mecanismo de banco de dados relacional transacional, de uso geral e compatível com o modelo de atomicidade, consistência, isolamento e durabilidade (ACID). O InnoDB apresenta estruturas na memória (grupo de buffers, buffer de alterações, índice de hash adaptativo, buffer de logs), bem como estruturas em disco (espaços para tabelas, tabelas, índices, log de undo, log de redo, arquivos de buffer de doublewrite). Para garantir que seu banco de dados siga rigorosamente o modelo ACID, o mecanismo de armazenamento InnoDB implementa vários recursos para proteger seus dados, incluindo transações, commit, reversão, recuperação de falhas, bloqueio em nível de linha e controle de simultaneidade multiversão (MVCC).

Todos esses componentes internos de uma instância de banco de dados trabalham em conjunto para ajudar a manter a disponibilidade, a integridade e a segurança de seus dados no nível de performance esperado e satisfatório. Dependendo da sua workload, cada componente e recurso pode impor demandas de recursos aos subsistemas de CPU, memória, rede e armazenamento. Quando um aumento na demanda por um recurso específico excede a capacidade provisionada ou os limites de software desse recurso (impostos pelos parâmetros de configuração ou pelo design do software), a instância de banco de dados pode sofrer degradação da performance ou ficar totalmente indisponível e corrompida. Portanto, é fundamental medir e monitorar esses componentes internos, compará-los com os valores de linha de base definidos e gerar alertas se os valores monitorados se desviarem dos valores esperados.

Conforme descrito anteriormente, você pode usar ferramentas diferentes para monitorar suas instâncias do MySQL e do MariaDB. Recomendamos que você use o Amazon RDS Performance Insights e as CloudWatch ferramentas para monitoramento e alertas, porque essas ferramentas são integradas ao Amazon RDS, reúnem métricas de alta resolução, apresentam as informações de desempenho mais recentes quase em tempo real e geram alarmes.

Independentemente da sua ferramenta de monitoramento preferencial, recomendamos que você ative o Esquema de Performance em suas instâncias dos bancos de dados MySQL e MariaDB. O

Esquema de Performance é um recurso opcional para monitorar a operação do servidor MySQL (a instância de banco de dados) em um nível baixo, e foi projetado para ter um impacto mínimo na performance geral do banco de dados. Você pode gerenciar esse recurso usando o parâmetro `performance_schema`. Embora esse parâmetro seja opcional, você deve usá-lo para coletar métricas de alta resolução (um segundo) por SQL, métricas de sessão ativa, eventos de espera e outras informações detalhadas de monitoramento de baixo nível, que são coletadas pelo Insights de Performance do Amazon RDS.

Seções

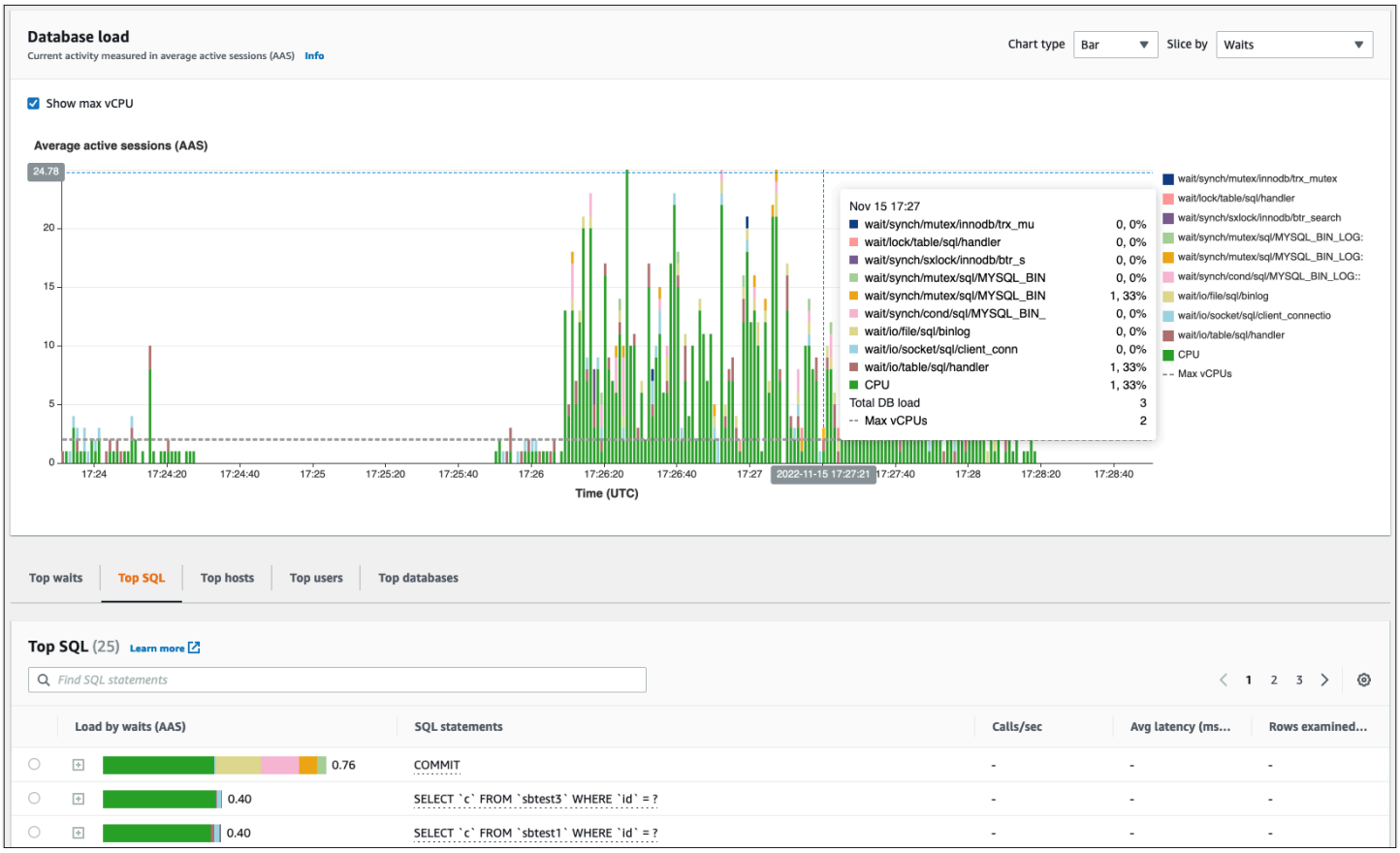
Métricas do Insights de Performance para instâncias de banco de dados

O Insights de Performance monitora diferentes tipos de métricas, conforme discutido nas seções a seguir.

Carga de banco de dados

A carga do banco de dados (`DBLoad`) é uma métrica essencial no Insights de Performance que mede o nível de atividade no seu banco de dados. Ele é coletado a cada segundo e publicado automaticamente na Amazon CloudWatch. Ela representa a atividade da instância do banco de dados em média de sessões ativas (AAS), que é o número de sessões que estão executando consultas SQL simultaneamente. A métrica `DBLoad` é diferente de outras métricas de séries temporais, pois pode ser interpretada usando qualquer uma destas cinco dimensões: esperas, SQL, hosts, usuários e bancos de dados. Essas dimensões são subcategorias da métrica `DBLoad`. Você pode usá-las segmentadas por categorias para representar diferentes características da carga do banco de dados. Para obter uma descrição detalhada de como calculamos a carga do banco de dados, consulte Carga de banco de dados na documentação do Amazon RDS.

A ilustração de tela a seguir mostra a ferramenta Insights de Performance.



Dimensões

- Eventos de espera são condições em que uma sessão de banco de dados espera pela conclusão de um recurso ou outra operação para continuar seu processamento. Se você executar uma instrução SQL como `SELECT * FROM big_table` e se essa tabela for muito maior do que o buffer pool alocado do InnoDB, sua sessão provavelmente aguardará eventos de espera`wait/io/file/innodb/innodb_data_file`, que são causados por operações I/O físicas no arquivo de dados. Eventos de espera são uma dimensão importante para o monitoramento de banco de dados, pois indicam possíveis gargalos de performance. Os eventos de espera indicam os recursos e as operações pelos quais as instruções SQL que você está executando nas sessões passam a maior parte do tempo esperando. Por exemplo, o evento `wait/synch/mutex/innodb/trx_sys_mutex` ocorre quando há alta atividade do banco de dados com um grande número de transações, e o evento `wait/synch/mutex/innodb/buf_pool_mutex` ocorre quando um thread adquiriu um bloqueio no grupo de buffers do InnoDB para acessar uma página na memória. Para obter informações sobre todos os eventos de espera do MariaDB e do MySQL, consulte Wait Event Summary Tables na documentação do MySQL. Para entender como

interpretar nomes de instrumentos, consulte Performance Schema Instrument Naming Conventions na documentação do MySQL.

- O SQL mostra quais instruções SQL estão contribuindo mais para a carga total do banco de dados. A tabela Principais dimensões, localizada abaixo do gráfico Carga do banco de dados no Insights de Performance do Amazon RDS, é interativa. Você pode obter uma lista detalhada dos eventos de espera associados à instrução SQL clicando na barra na coluna Carregar por esperas (AAS). Quando você seleciona uma instrução SQL na lista, o Insights de Performance exibe os eventos de espera associados no gráfico Carga do banco de dados e o texto da instrução SQL na seção Texto SQL. As estatísticas do SQL são exibidas no lado direito da tabela de Principais dimensões.
- Hosts mostra os nomes dos hosts dos clientes conectados. Essa dimensão ajuda a identificar quais hosts clientes estão enviando a maior parte da carga para o banco de dados.
- Usuários agrupa a carga do banco de dados por usuários que estão conectados ao banco de dados.
- Bancos de dados agrupa a carga do banco de dados pelo nome do banco de dados ao qual o cliente está conectado.

Métricas de contador

As métricas de contador são métricas cumulativas cujos valores só podem aumentar ou ser redefinidos para zero quando a instância de banco de dados é reiniciada. O valor de uma métrica de contador não pode ser reduzido ao valor anterior. Essas métricas representam um contador único e monotonicamente crescente.

- Contadores nativos são métricas definidas pelo mecanismo de banco de dados e não pelo Amazon RDS. Por exemplo:
  - `SQL.Innodb_rows_inserted` representa o número de linhas inseridas nas tabelas do InnoDB.
  - `SQL.Select_scan` representa o número de junções que concluíram uma varredura completa da primeira tabela.
  - `Cache.Innodb_buffer_pool_reads` representa o número de leituras lógicas que o mecanismo do InnoDB não conseguiu recuperar do grupo de buffers e precisou ler diretamente do disco.
  - `Cache.Innodb_buffer_pool_read_requests` representa o número de solicitações de leitura lógica.

Para obter definições para essas métricas nativas, consulte Server Status Variables na documentação do MySQL.

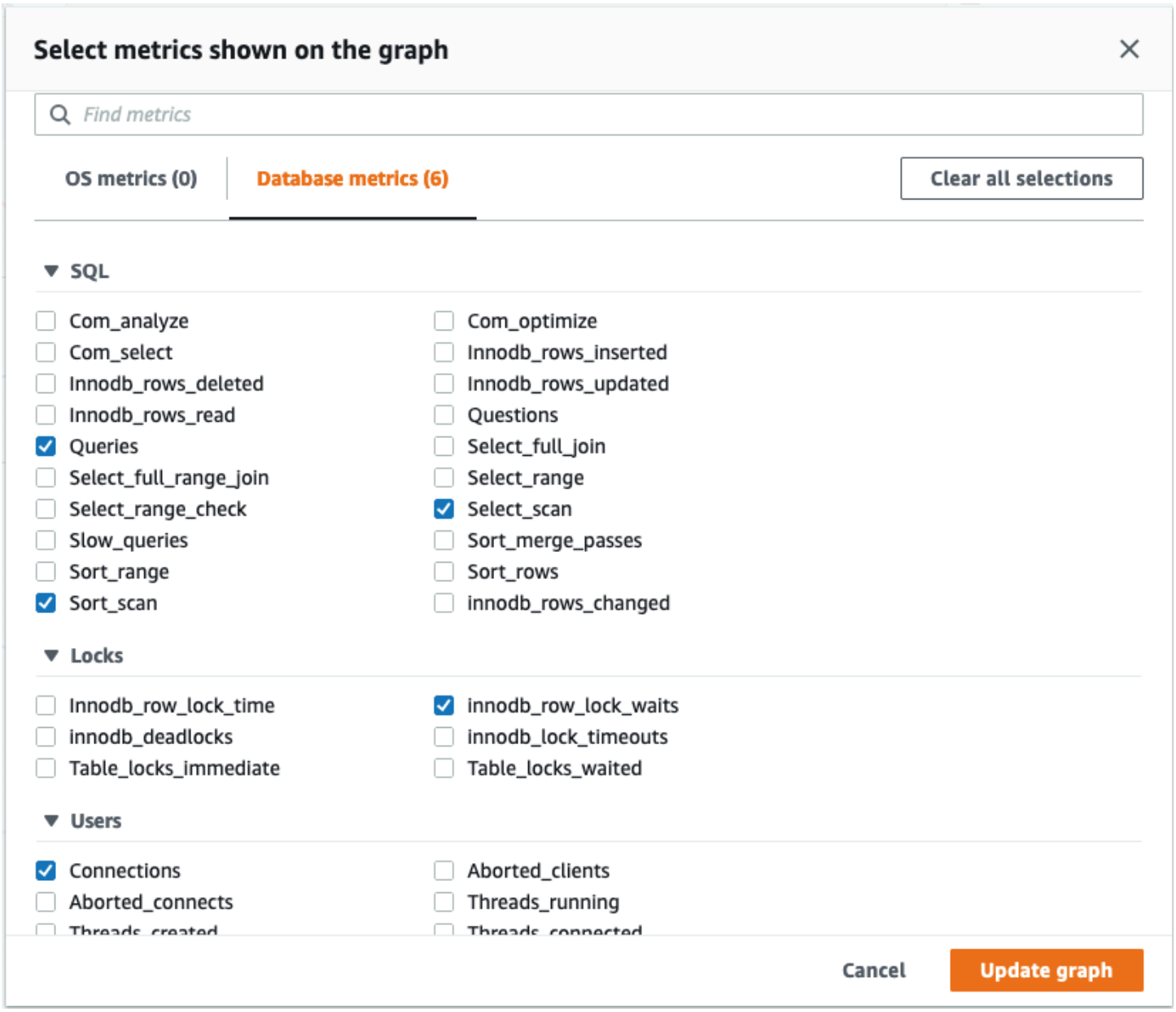
- Contadores não nativos são definidos pelo Amazon RDS. Você pode obter essas métricas usando uma consulta específica ou derivá-las usando duas ou mais métricas nativas nos cálculos. As métricas de contador não nativas podem representar latências, proporções ou taxas de acerto. Por exemplo:
  - `Cache.innoDB_buffer_pool_hits` representa o número de operações de leitura que o InnoDB pode recuperar do grupo de buffers sem utilizar o disco. É calculado com base nas métricas do contador nativo da seguinte forma:

    ```
    db.Cache.Innodb_buffer_pool_read_requests - db.Cache.Innodb_buffer_pool_reads
    ```

  - `IO.innoDB_datafile_writes_to_disk` representa o número de operações de gravação do arquivo de dados do InnoDB no disco. Ele captura somente operações em arquivos de dados, não operações de gravação de doublewrite ou redo logging. É calculado da seguinte forma:

    ```
    db.IO.Innodb_data_writes - db.IO.Innodb_log_writes - db.IO.Innodb_dblwr_writes
    ```

Você pode visualizar métricas de instâncias de banco de dados diretamente no painel do Insights de Performance. Escolha Gerenciar métricas, depois a guia Métricas do banco de dados, e selecione as métricas de interesse, conforme mostrado na ilustração a seguir.



Escolha o botão Atualizar grafo para exibir as métricas selecionadas, conforme mostrado na ilustração a seguir.

Estatísticas SQL

O Insights de Performance coleta métricas relacionadas a performance sobre as consultas SQL para cada segundo de execução de uma consulta e para cada chamada SQL. Em geral, o Insights de Performance coleta estatísticas do SQL nos níveis de instrução e resumo. No entanto, para instâncias de banco de dados MariaDB e MySQL, as estatísticas são coletadas apenas no nível de resumo.

- As estatísticas de resumo são uma métrica composta de todas as consultas que têm o mesmo padrão, mas que acabam tendo valores literais diferentes. O resumo substitui valores literais específicos por uma variável, por exemplo:

```
SELECT department_id, department_name FROM departments WHERE location_id = ?
```

- Há métricas que representam estatísticas por segundo para cada instrução SQL resumida. Por exemplo, `sql_tokenized.stats.count_star_per_sec` representa chamadas por segundo (ou seja, quantas vezes por segundo a instrução SQL foi executada).

- O Insights de Performance também inclui métricas que fornecem estatísticas por chamada para uma instrução SQL. Por exemplo, `sql_tokenized.stats.sum_timer_wait_per_call` mostra a latência média da instrução SQL por chamada, em milissegundos.

As estatísticas do SQL estão disponíveis no painel do Insights de Performance na guia SQL principal da tabela Principais dimensões.

CloudWatch métricas para instâncias de banco de dados

A Amazon CloudWatch também contém métricas que o Amazon RDS publica automaticamente. As métricas que residem no namespace `AWS/RDS` são métricas em nível de instância que se referem à instância (serviço) do Amazon RDS (ou seja, o ambiente de banco de dados isolado executado na nuvem) em vez de à instância de banco de dados no sentido estrito do processo mysqld. Portanto, a maioria dessas métricas padrão se enquadra na categoria de métricas de sistema operacional, na definição exata do termo. Alguns exemplos incluem: `CPUUtilization`, `WriteIOPS`, `SwapUsage`, entre outras. No entanto, existem algumas métricas de instâncias de banco de dados que são aplicáveis ao MariaDB e ao MySQL:

- `BinLogDiskUsage`: o espaço em disco ocupado por logs binários.
- `DatabaseConnections`: o número de conexões de rede cliente com a instância de banco de dados.
- `ReplicaLag`: o tempo que uma instância de banco de dados de réplica de leitura atrasa em relação à instância de banco de dados de origem.

Publicando métricas do Performance Insights em CloudWatch

O Amazon RDS Performance Insights monitora a maioria das métricas e dimensões da instância de banco de dados e as disponibiliza por meio do painel Performance Insights no AWS Management Console. Esse painel é adequado para a solução de problemas de banco de dados e a análise da causa raiz. No entanto, não é possível criar alarmes no Insights de Performance para métricas relacionadas à performance. Se você quiser criar alarmes com base nas métricas do Performance Insights, essas métricas devem estar presentes CloudWatch.

O Performance Insights publica automaticamente métricas em CloudWatch. Você pode consultar os mesmos dados do Performance Insights, mas ter as métricas inseridas CloudWatch facilita a adição de CloudWatch alarmes e a adição das métricas aos CloudWatch painéis existentes. Os contadores são métricas de performance do sistema operacional e do banco de dados, como `os.memory.free` ou `db.Locks.Innodb_row_lock_time`. A coleta de métricas do sistema operacional depende da configuração do Monitoramento Aprimorado. Se esse recurso estiver desativado, as métricas do sistema operacional serão coletadas uma vez por minuto. Se estiver ativado, as métricas do sistema operacional serão coletadas para o período selecionado. Para obter mais informações, consulte Ativar e desativar o Monitoramento Aprimorado na documentação do Amazon RDS.

O Performance Insights permite que você exporte o painel de métricas pré-configurado ou personalizado da sua instância de banco de dados para CloudWatch. Você pode exportar o painel de métricas como um novo painel ou adicioná-lo a um CloudWatch painel existente. A exportação do painel de métricas do Performance Insights para o CloudWatch painel oferece uma visão unificada e holística da integridade do seu sistema, fornecendo uma visão geral das métricas associadas a vários recursos em seu sistema, como instâncias do EC2, recursos do Amazon Elastic File System (Amazon EFS) e recursos do Elastic Load Balancing (ELB), junto com as métricas da sua instância de banco de dados.

Você pode usar a função matemática CloudWatch `DB_PERF_INSIGHTS` métrica para consultar e criar alarmes e gráficos com base nas métricas do Performance Insights de. CloudWatch Para criar um alarme em uma métrica do Performance Insights, siga as instruções na CloudWatch documentação. Por exemplo, se você quiser acionar um alarme quando o total de transações ativas em sua instância de banco de dados atingir um limite específico, siga as instruções na página, use a expressão matemática `DB_PERF_INSIGHTS` e escolha Aplicar:

```
DB_PERF_INSIGHTS('RDS', 'db-BQ2TPYY7HG2GDFC7APMB3BVB3M',
 'db.Transactions.active_transactions.avg')
```

em que `db-BQ2TPYY7HG2GDFC7APMB3BVB3M` é o ID do recurso da sua instância de banco de dados. Especifique o período (por exemplo, 1 minuto) e as condições (por exemplo, maior que 1000). Para finalizar a criação do alarme, configure as ações do alarme, adicione um nome e uma descrição e pré-visualize e crie o alarme.

Monitoramento do sistema operacional

Uma instância de banco de dados no Amazon RDS para MySQL ou MariaDB é executada no sistema operacional Linux, que usa recursos subjacentes do sistema: CPU, memória, rede e armazenamento.

```
MySQL [(none)]> SHOW variables LIKE 'version%';
+-------------------------+---------------------+
| Variable_name           | Value               |
+-------------------------+---------------------+
| version                 | 8.0.28              |
| version_comment         | Source distribution |
| version_compile_machine | aarch64             |
| version_compile_os      | Linux               |
| version_compile_zlib    | 1.2.11              |
+-------------------------+---------------------+
5 rows in set (0.00 sec)
```

A performance geral do banco de dados e do sistema operacional subjacente depende muito da utilização dos recursos do sistema. Por exemplo, a CPU é o componente chave para a performance do sistema, pois executa as instruções do software do banco de dados e gerencia outros recursos do sistema. Se a CPU for superutilizada (ou seja, se a carga exigir mais potência de CPU do que a provisionada para sua instância de banco de dados), esse problema afetará a performance e a estabilidade do seu banco de dados e, consequentemente, da sua aplicação.

O mecanismo de banco de dados aloca e libera memória dinamicamente. Quando não há memória suficiente na RAM para fazer o trabalho atual, o sistema grava páginas de memória na memória swap, que reside no disco. Como o disco é muito mais lento que a memória, mesmo que seja baseado na tecnologia SSD NVMe, a alocação excessiva de memória leva à degradação da performance. A alta utilização da memória causa maior latência das respostas do banco de dados, porque o tamanho de um arquivo de paginação aumenta para suportar memória adicional. Se a alocação de memória for tão alta que esgote tanto a RAM quanto os espaços de memória swap, o serviço de banco de dados poderá ficar indisponível e os usuários poderão observar erros como `[ERROR] mysqld: Out of memory (Needed xyz bytes)`.

Os sistemas de gerenciamento de banco de dados MySQL e MariaDB utilizam o subsistema de armazenamento, que consiste em discos que armazenam estruturas em disco, como tabelas, índices, logs binários, logs de redo, logs de undo e arquivos de buffer de doublewrite. Portanto, o

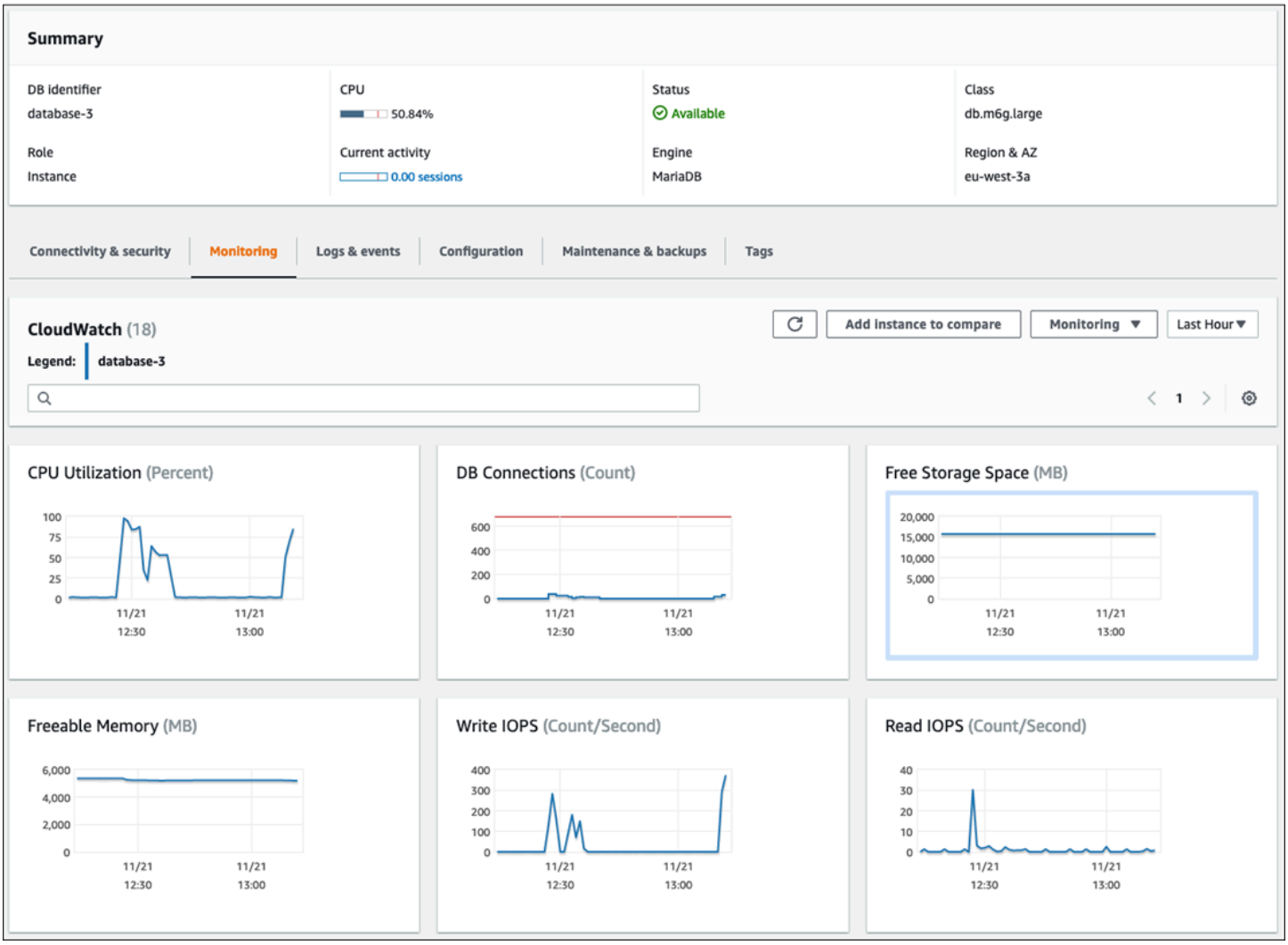
banco de dados, ao contrário de outros tipos de software, deve realizar muita atividade no disco. Para a operação ideal do seu banco de dados, é importante monitorar e ajustar a utilização de E/S do disco e a alocação de espaço em disco. A performance do banco de dados pode ser afetada quando o banco de dados atinge os limites de IOPS ou o throughput máximo suportado pelo disco. Por exemplo, expansões de acesso aleatório causados por uma verificação de índice podem causar um grande número de operações de E/S por segundo, o que eventualmente pode atingir as limitações do armazenamento subjacente. As verificações completas das tabelas podem não atingir o limite de IOPS, mas podem causar um alto throughput mensurado em megabytes por segundo. É fundamental monitorar e gerar alertas sobre a alocação de espaço em disco, pois erros como `OS error code 28: No space left on device` podem causar indisponibilidade e corrupção do banco de dados.

O Amazon RDS fornece métricas em tempo real para o sistema operacional em que sua instância de banco de dados é executada. O Amazon RDS publica automaticamente um conjunto de métricas do sistema operacional no CloudWatch. Essas métricas estão disponíveis para exibição e análise no console do Amazon RDS e nos painéis do CloudWatch, e você pode definir alarmes nas métricas selecionadas no CloudWatch. Os exemplos incluem:

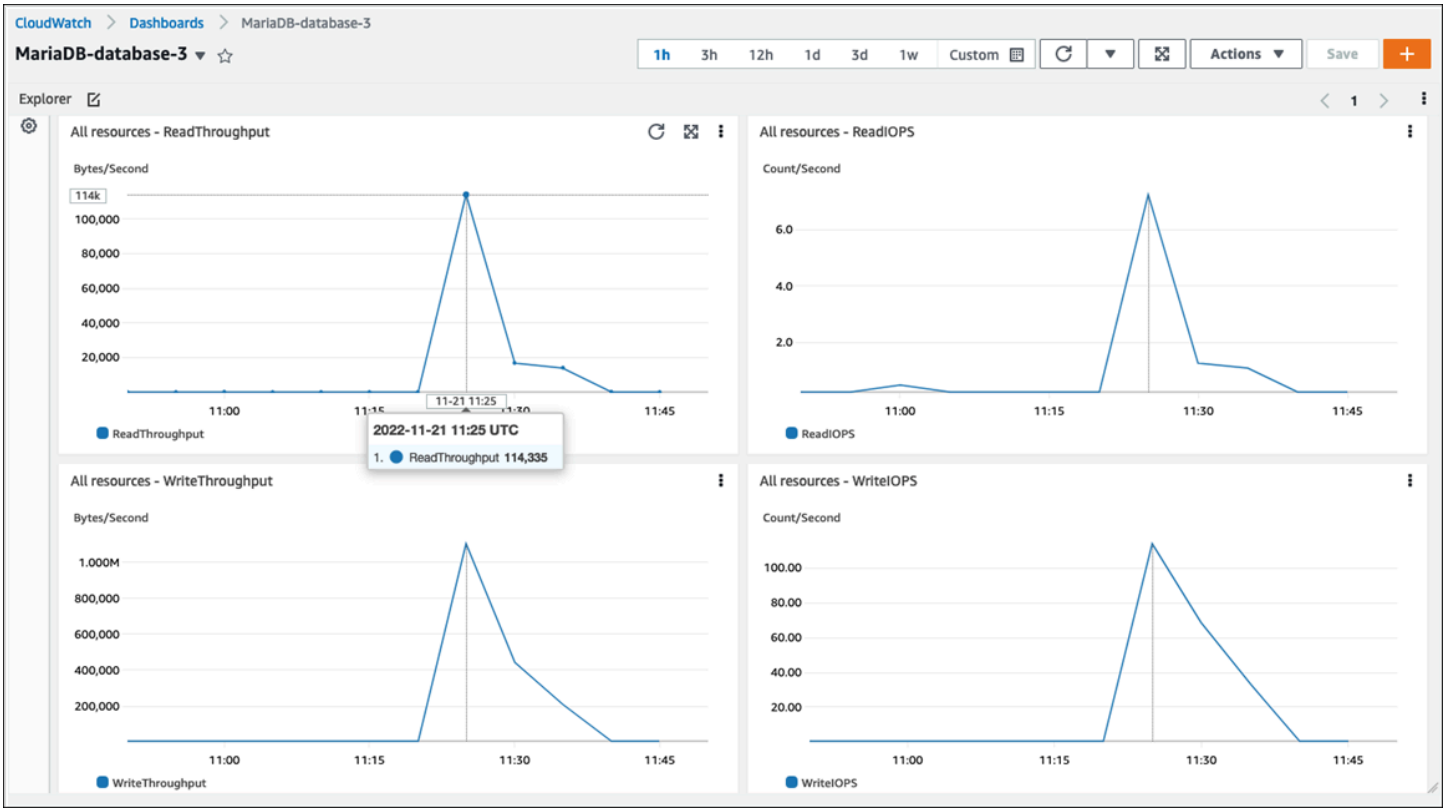
- `CPUUtilization`: o percentual de utilização da CPU.
- `BinLogDiskUsage`: o volume do espaço em disco ocupado por logs binários.
- `FreeableMemory`: a quantidade de memória de acesso aleatório disponível. Isso representa o valor do campo `MemAvailable` de `/proc/meminfo`.
- `ReadIOPS`: o número médio de operações E/S de leitura de disco por segundo.
- `WriteThroughput`: o número médio de bytes gravados no disco por segundo para o armazenamento local.
- `NetworkTransmitThroughput`: o tráfego de rede de saída no nó do banco de dados, que combina o tráfego de banco de dados e o tráfego do Amazon RDS usado para monitoramento e replicação.

Para obter uma referência completa de todas as métricas publicadas pelo Amazon RDS no CloudWatch, consulte as métricas do Amazon CloudWatch para o Amazon RDS na documentação do Amazon RDS.

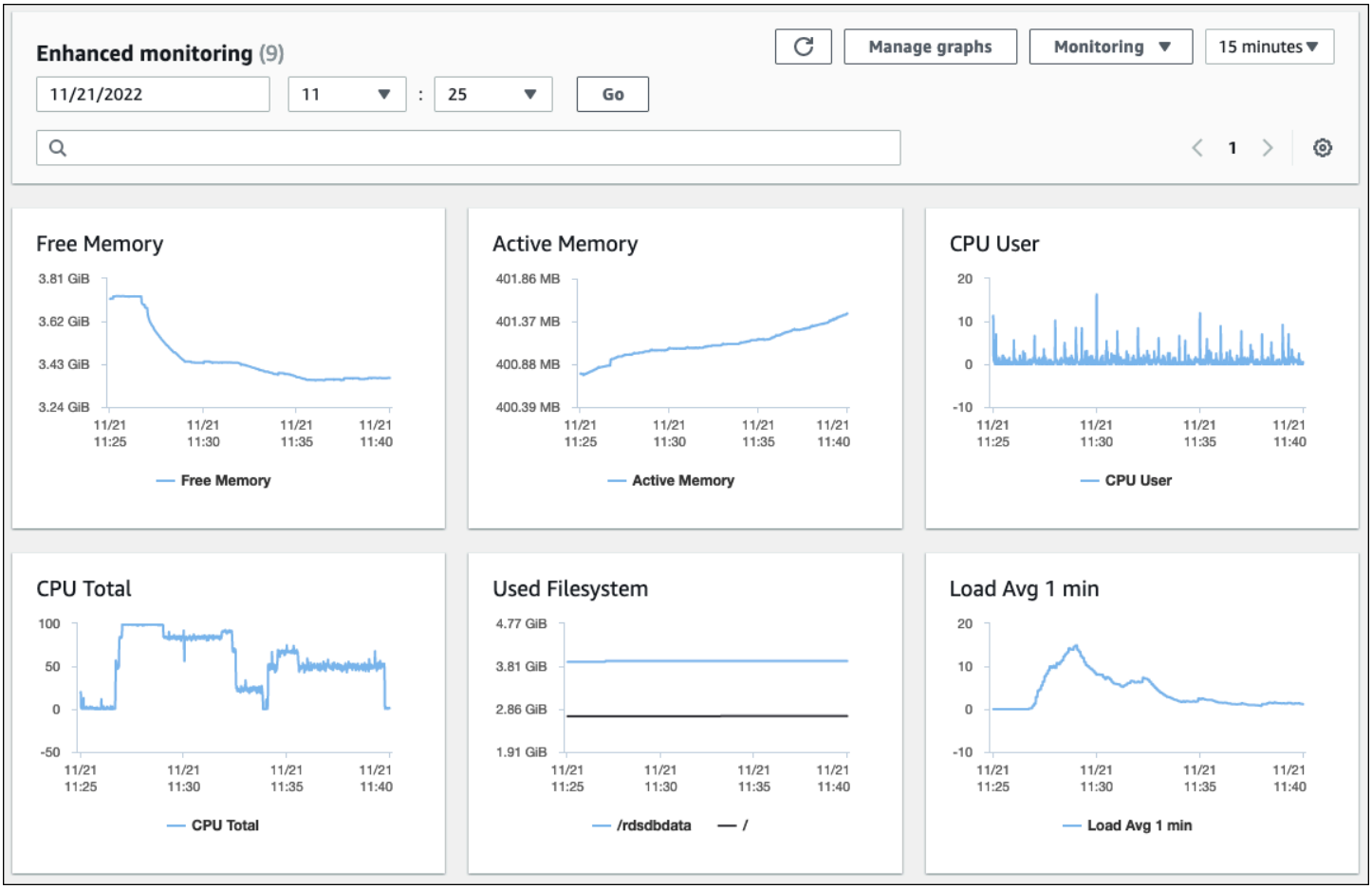
O gráfico a seguir mostra exemplos de métricas do CloudWatch para o Amazon RDS que são exibidas no console do Amazon RDS.



O gráfico a seguir mostra métricas semelhantes exibidas no painel do CloudWatch.



O outro conjunto de métricas do sistema operacional é coletado pelo Monitoramento Aprimorado para Amazon RDS. Essa ferramenta oferece uma visibilidade mais profunda da integridade de suas instâncias de banco de dados do Amazon RDS para MariaDB e Amazon RDS para MySQL, fornecendo métricas do sistema em tempo real e informações sobre o processo do sistema operacional. Quando você habilita o Monitoramento Aprimorado em sua instância de banco de dados e define a granularidade desejada, a ferramenta coleta as métricas do sistema operacional e as informações do processo, que você pode exibir e analisar no console do Amazon RDS, conforme mostrado na tela a seguir.



Algumas das principais métricas fornecidas pelo Monitoramento Aprimorado são:

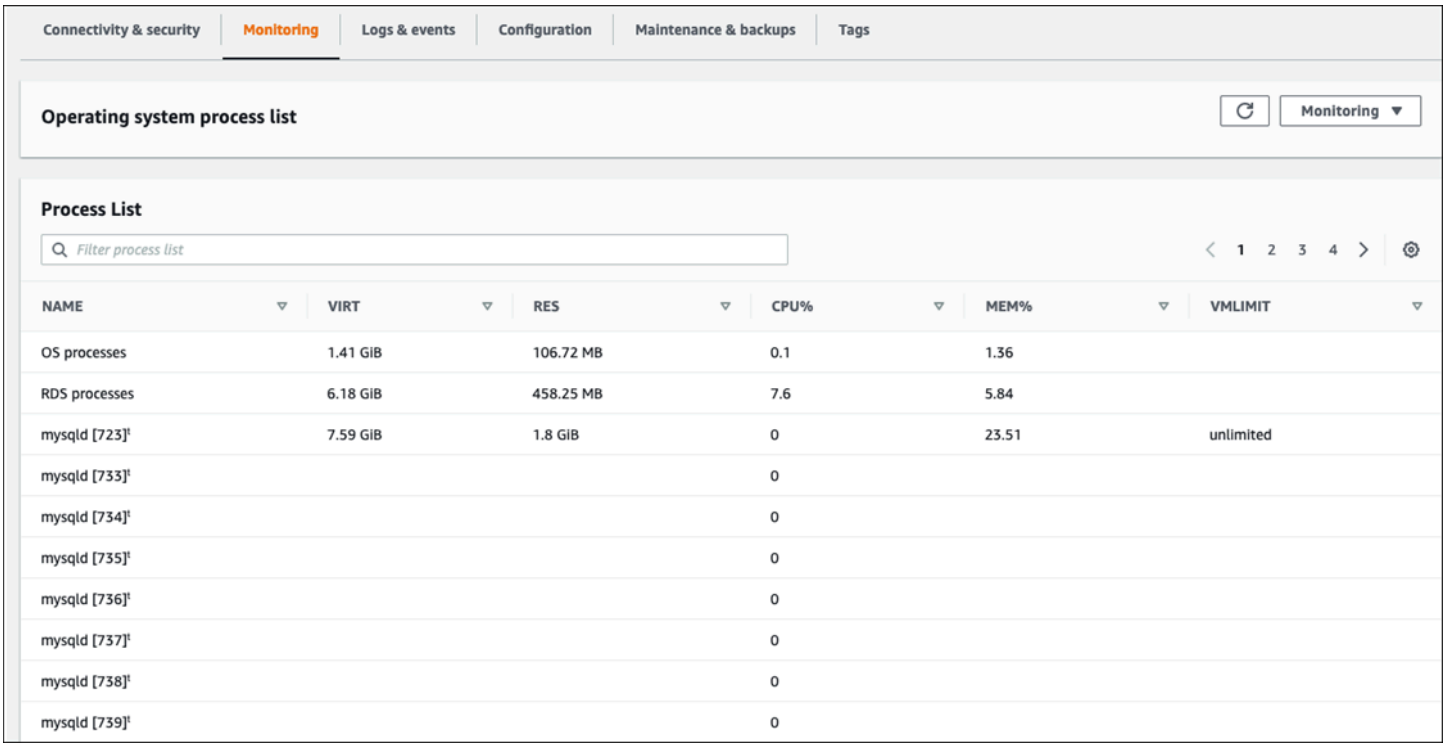
- `cpuUtilization.total`: a porcentagem total da CPU em uso.
- `cpuUtilization.user`: a porcentagem de CPU em uso por programas do usuário.
- `memory.active`: a quantidade de memória atribuída, em kilobytes.
- `memory.cached`: a quantidade de memória utilizada para o armazenamento em cache da E/S baseada no sistema de arquivos.
- `loadAverageMinute.one`: o número de processos que solicitaram tempo de CPU no último minuto.

Para obter uma lista completa de métricas, consulte as Métricas do sistema operacional no Monitoramento Aprimorado na documentação do Amazon RDS.

No console do Amazon RDS, a lista de processos do sistema operacional fornece detalhes de cada processo que está sendo executado na sua instância de banco de dados. A lista está organizada em três seções:

- Processos do sistema operacional: esta seção representa um resumo agregado de todos os processos do kernel e do sistema. Esses processos geralmente têm impacto mínimo na performance do banco de dados.
- Processos do RDS: esta seção representa um resumo dos processos necessários da AWS para oferecer suporte a uma instância de banco de dados do Amazon RDS. Por exemplo, inclui o agente de gerenciamento do Amazon RDS, processos de monitoramento e diagnóstico e processos similares.
- Processos secundários do RDS: esta seção representa um resumo dos processos do Amazon RDS que oferecem suporte à instância de banco de dados, nesse caso, o processo `mysqld` e seus threads. Os threads do `mysqld` aparecem aninhados abaixo do processo principal do `mysqld`.

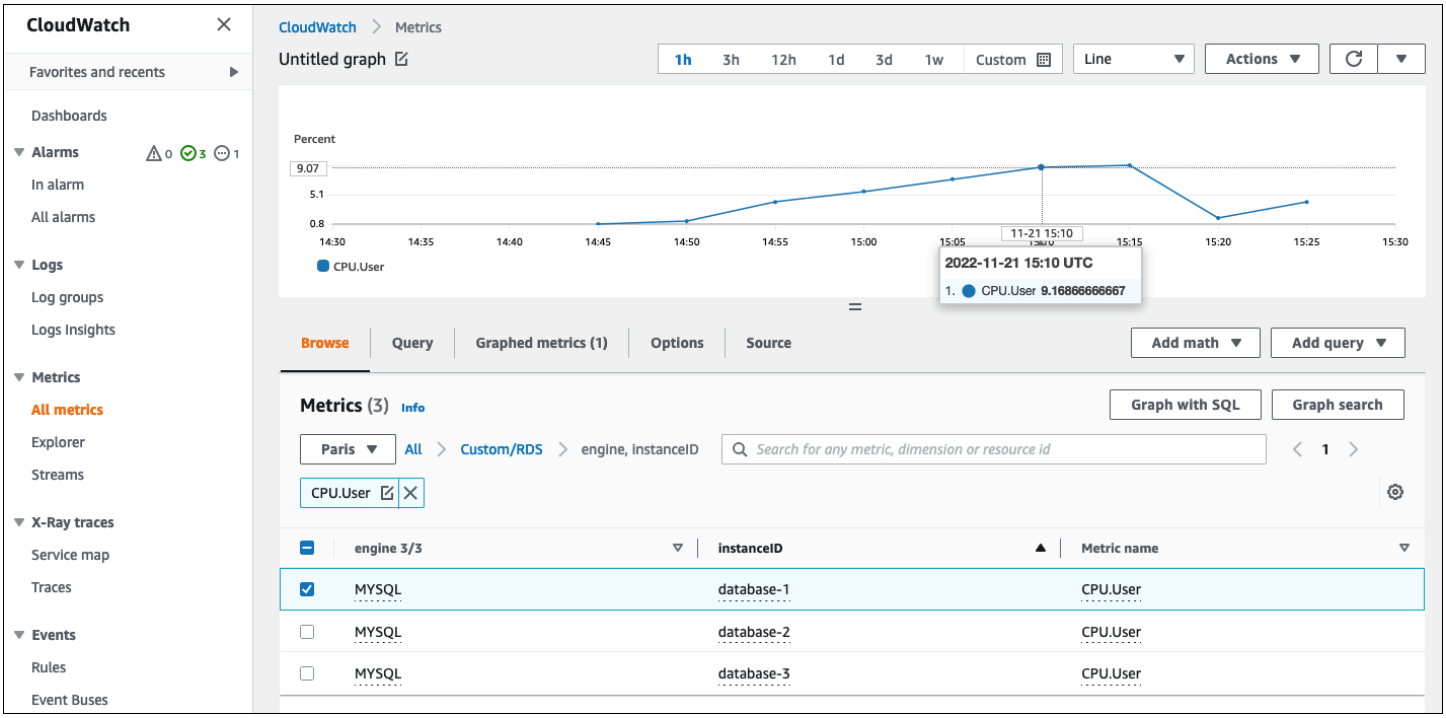
A ilustração de tela a seguir mostra a lista de processos do sistema operacional no console do Amazon RDS.



O Amazon RDS entrega as métricas do Monitoramento Aprimorado à sua conta do Amazon CloudWatch Logs. Os dados de monitoramento que são mostrados no console do Amazon RDS são recuperados do CloudWatch Logs. Você também pode recuperar as métricas para uma instância de banco de dados como um fluxo de logs do CloudWatch Logs. Essas métricas são armazenadas no formato JSON. É possível consumir o resultado do JSON de monitoramento avançado do CloudWatch Logs em um sistema de monitoramento de sua escolha.

Para exibir grafos no painel do CloudWatch e criar alarmes que iniciarão uma ação se uma métrica ultrapassar o limite definido, você deve criar filtros de métricas no CloudWatch do CloudWatch Logs. Para obter instruções detalhadas, consulte o artigo do AWS re:Post sobre como filtrar os logs de Monitoramento Aprimorado do CloudWatch para gerar métricas personalizadas automatizadas para o Amazon RDS.

O exemplo a seguir ilustra a métrica personalizada `CPU.User` no namespace `Custom/RDS`. Essa métrica personalizada é criada filtrando a métrica `cpuUtilization.user` do Monitoramento Aprimorado do CloudWatch Logs.



Quando a métrica está disponível no repositório do CloudWatch, você pode exibi-la e analisá-la nos painéis do CloudWatch, aplicar mais operações matemáticas e de consulta e definir um alarme para monitorar essa métrica específica e gerar alertas se os valores observados não estiverem de acordo com as condições de alarme definidas.

Eventos, logs e trilhas de auditoria

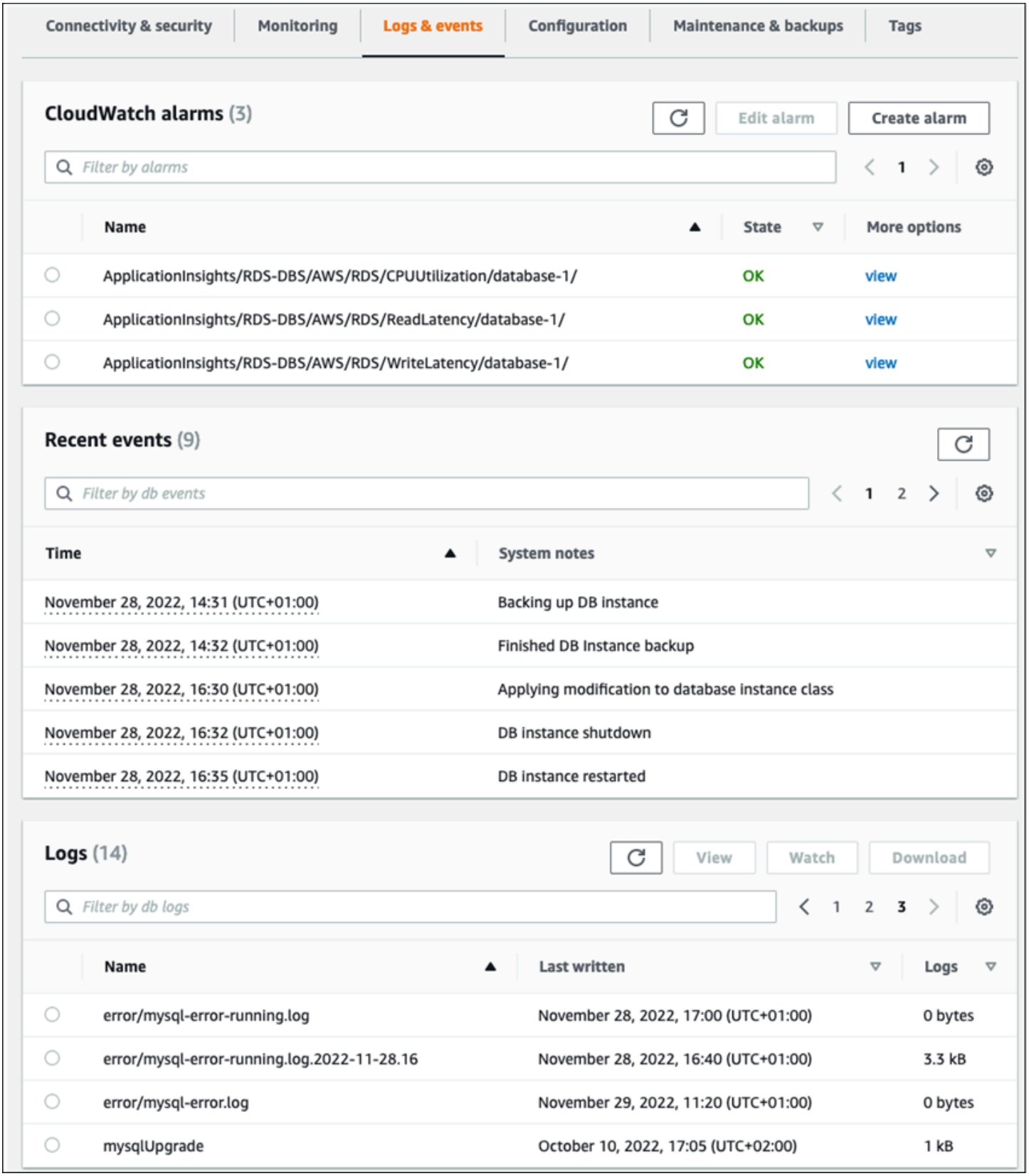
Monitorar métricas de instâncias de banco de dados e métricas do sistema operacional, analisar as tendências e comparar as métricas com valores de linha de base e gerar alertas quando os valores ultrapassam os limites definidos são práticas recomendadas e necessárias que ajudam você a alcançar e manter a confiabilidade, a disponibilidade, a performance e a segurança de suas instâncias de banco de dados do Amazon RDS. No entanto, uma solução completa também deve monitorar eventos de banco de dados, arquivos de logs e trilhas de auditoria dos bancos de dados MySQL e MariaDB.

Seções

- eventos do Amazon RDS
- Logs de banco de dados
- Trilhas de auditoria

eventos do Amazon RDS

Um evento do Amazon RDS indica uma alteração no ambiente do Amazon RDS. Por exemplo, quando o status da instância de banco de dados muda de Starting para Available, o Amazon RDS gera o evento `RDS-EVENT-0088 The DB instance has been started`. O Amazon RDS entrega eventos ao Amazon EventBridge quase em tempo real. Você pode acessar eventos por meio do console do Amazon RDS, do comando describe-events da AWS CLI ou da operação da API DescribeEvents do Amazon RDS. A ilustração de tela a seguir mostra eventos e logs exibidos no console do Amazon RDS.


Connectivity & security
Monitoring
Logs & events
Configuration
Maintenance & backups
Tags
CloudWatch alarms (3)
Edit alarm
Create alarm
Filter by alarms
1
Name
State
More options
ApplicationInsights/RDS-DBS/AWS/RDS/CPUUtilization/database-1/
OK
view
ApplicationInsights/RDS-DBS/AWS/RDS/ReadLatency/database-1/
OK
view
ApplicationInsights/RDS-DBS/AWS/RDS/WriteLatency/database-1/
OK
view
Recent events (9)
Filter by db events
1
2
Time
System notes
November 28, 2022, 14:31 (UTC+01:00)
Backing up DB instance
November 28, 2022, 14:32 (UTC+01:00)
Finished DB Instance backup
November 28, 2022, 16:30 (UTC+01:00)
Applying modification to database instance class
November 28, 2022, 16:32 (UTC+01:00)
DB instance shutdown
November 28, 2022, 16:35 (UTC+01:00)
DB instance restarted
Logs (14)
View
Watch
Download
Filter by db logs
1
2
3
Name
Last written
Logs
error/mysql-error-running.log
November 28, 2022, 17:00 (UTC+01:00)
0 bytes
error/mysql-error-running.log.2022-11-28.16
November 28, 2022, 16:40 (UTC+01:00)
3.3 kB
error/mysql-error.log
November 29, 2022, 11:20 (UTC+01:00)
0 bytes
mysqlUpgrade
October 10, 2022, 17:05 (UTC+02:00)
1 kB

O Amazon RDS emite diferentes tipos de eventos, incluindo eventos de instância de banco de dados, eventos de grupos de parâmetros de banco de dados, eventos de grupos de segurança de banco de dados, eventos de snapshots de banco de dados, eventos de proxy do RDS e eventos de implantação azul/verde. As informações incluem:

- Nome da origem e tipo de origem; por exemplo: `"SourceIdentifier": "database-1", "SourceType": "db-instance"`
- Data e hora do evento; por exemplo: `"Date": "2022-12-01T09:20:28.595000+00:00"`
- Mensagem associada ao evento; por exemplo: `"Message": "Finished updating DB parameter group"`
- Categoria do evento; por exemplo: `"EventCategories": ["configuration change"]`

Para uma referência completa, consulte Categorias de eventos e mensagens de eventos do Amazon RDS na documentação do Amazon RDS.

Recomendamos que você monitore os eventos do Amazon RDS, pois eles indicam mudanças de status na disponibilidade de instâncias de banco de dados, alterações de configuração, mudanças de status de réplica de leitura, eventos de backup e recuperação, ações de failover, eventos de falha, modificações em grupos de segurança e muitas outras notificações. Por exemplo, se você configurou uma instância de banco de dados de réplica de leitura para fornecer performance e durabilidade aprimorados para seu banco de dados, recomendamos que você monitore os eventos do Amazon RDS da categoria de eventos de réplica de leitura associada às instâncias de banco de dados. Isso ocorre porque eventos como `RDS-EVENT-0057 Replication on the read replica was terminated` indicam que sua réplica de leitura não está mais sincronizada com a instância de banco de dados primária. Uma notificação à equipe responsável de que tal evento ocorreu pode ajudar a mitigar o problema em tempo hábil. O Amazon EventBridge e os Serviços da AWS adicionais, como o AWS Lambda, o Amazon Simple Queue Service (Amazon SQS) e o Amazon Simple Notification Service (Amazon SNS), podem ajudar você a automatizar respostas a eventos do sistema, como problemas de disponibilidade de banco de dados ou alterações de recursos.

No console do Amazon RDS, você pode recuperar eventos das últimas 24 horas. Se você usar a AWS CLI ou a API do Amazon RDS para visualizar eventos, poderá recuperar eventos dos últimos 14 dias usando o comando describe-events conforme a seguir.

```
$ aws rds describe-events --source-identifier database-1 --source-type db-instance
```

```
{
    "Events": [
        {
            "SourceIdentifier": "database-1",
            "SourceType": "db-instance",
            "Message": "CloudWatch Logs Export enabled for logs [audit, error, general,
 slowquery]",
            "EventCategories": [],
            "Date": "2022-12-01T09:20:28.595000+00:00",
            "SourceArn": "arn:aws:rds:eu-west-3:111122223333:db:database-1"
        },
        {
            "SourceIdentifier": "database-1",
            "SourceType": "db-instance",
            "Message": "Finished updating DB parameter group",
            "EventCategories": [
                "configuration change"
            ],
            "Date": "2022-12-01T09:22:40.413000+00:00",
            "SourceArn": "arn:aws:rds:eu-west-3:111122223333:db:database-1"
        }
    ]
}
```

Se você quiser armazenar eventos no longo prazo, até o período de expiração especificado ou permanentemente, é possível usar o CloudWatch Logs para registrar as informações sobre os eventos que foram gerados pelo Amazon RDS. Para implementar essa solução, você pode usar um tópico do Amazon SNS para receber notificações de eventos do Amazon RDS e, em seguida, chamar uma função do Lambda para registrar o evento no CloudWatch Logs.

1. Crie uma função do Lambda que será chamada no evento e registre as informações do evento no CloudWatch Logs. O CloudWatch Logs é integrado ao Lambda e fornece uma maneira conveniente de registrar informações de eventos de logs, usando a função imprimir para `stdout`.
2. Crie um tópico do SNS com uma assinatura para uma função do Lambda (defina o Protocolo como Lambda) e defina o Endpoint como o nome do recurso da Amazon (ARN) da função do Lambda que você criou na etapa anterior.
3. Configure seu tópico do SNS para receber notificações de eventos do Amazon RDS. Para obter instruções detalhadas, consulte o artigo do AWS re:Post sobre como fazer com que seu tópico do Amazon SNS receba notificações do Amazon RDS.

4. No console do Amazon RDS, crie uma nova assinatura de evento. Defina Destino como ARN e selecione o tópico do SNS que você criou anteriormente. Defina o Tipo de origem e as Categorias de eventos a serem incluídas de acordo com seus requisitos. Para obter mais informações, consulte Inscrever-se em notificações de eventos do Amazon RDS na documentação do Amazon RDS.

Logs de banco de dados

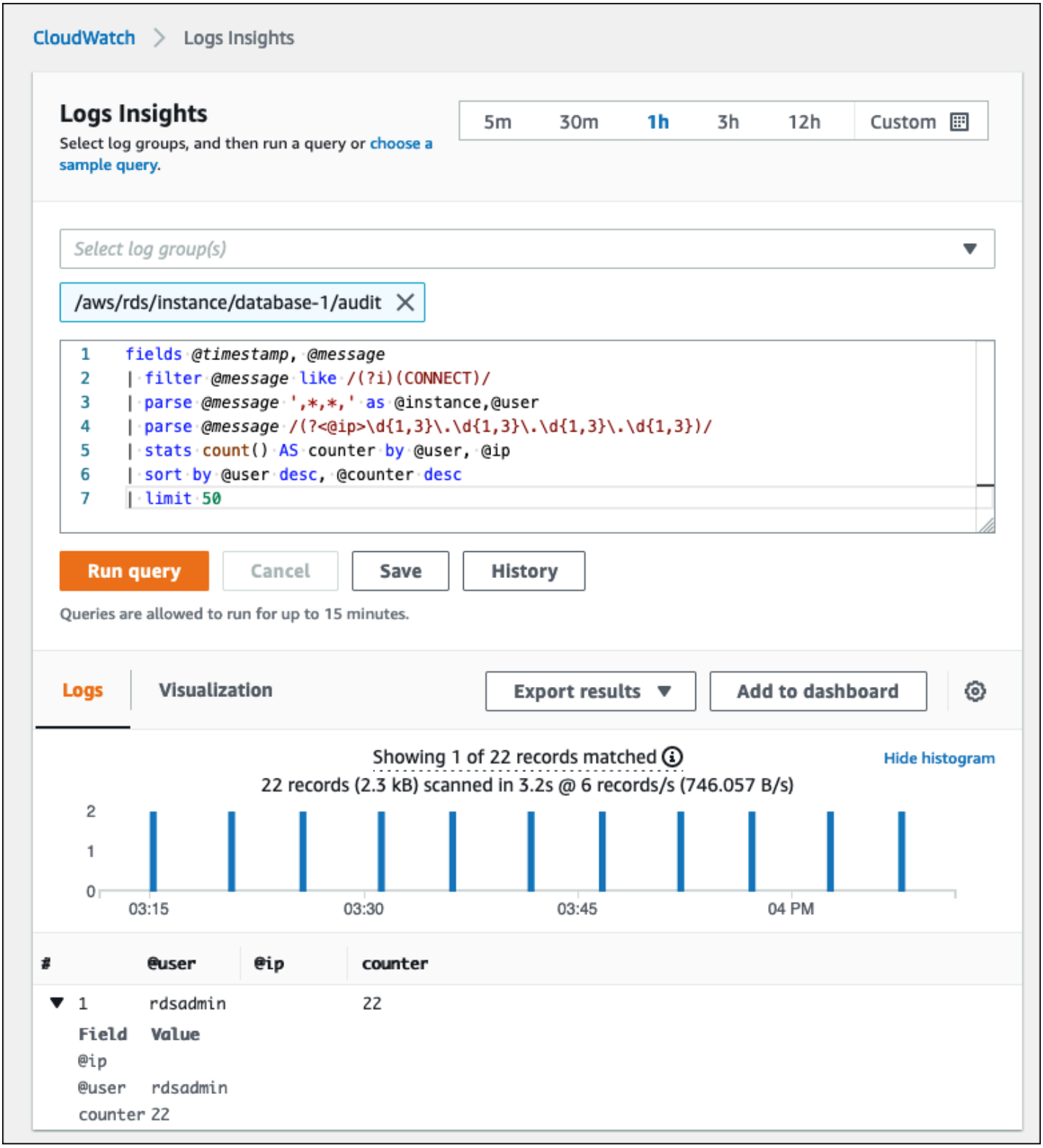
Os bancos de dados MySQL e MariaDB geram logs que você pode acessar para auditoria e solução de problemas. Esses logs são:

- Auditoria: a trilha de auditoria é um conjunto de registros que documentam a atividade do servidor. Para cada sessão do cliente, ele registra quem se conectou ao servidor (nome de usuário e host), quais consultas foram executadas, quais tabelas foram acessadas e quais variáveis do servidor foram alteradas.
- Erro: este log contém os horários de inicialização e desligamento do servidor (`mysqld`) e mensagens de diagnóstico, como erros, avisos e observações que ocorrem durante a inicialização e o desligamento do servidor e enquanto o servidor está em execução.
- Geral: este log registra a atividade de `mysqld`, incluindo a atividade de conexão e desconexão de cada cliente e as consultas SQL recebidas dos clientes. O log geral de consultas pode ser muito útil quando você suspeita de um erro e quer saber exatamente o que o cliente enviou para o `mysqld`.
- Consulta lenta: este log fornece um registro das consultas SQL que demoraram muito para serem executadas.

Como prática recomendada, você deve publicar logs de banco de dados do Amazon RDS no Amazon CloudWatch Logs. Com o CloudWatch Logs, é possível executar uma análise em tempo real dos dados em log, armazenar os dados em um armazenamento de alta durabilidade e gerenciar os dados com o agente do CloudWatch Logs. Você pode acessar e monitorar os logs do seu banco de dados no console do Amazon RDS. Você também pode usar o CloudWatch Logs Insights para pesquisar e analisar dados de log de forma interativa no CloudWatch Logs. O exemplo a seguir ilustra uma consulta no log de auditoria que verifica quantas vezes os eventos `CONNECT` aparecem no log, quem se conectou e de qual cliente (endereço IP) eles se conectaram. O trecho do log de auditoria seria o seguinte:

```
20221201 14:07:05,ip-10-22-1-51,rdsadmin,localhost,821,0,CONNECT,,,0,SOCKET
20221201 14:07:05,ip-10-22-1-51,rdsadmin,localhost,821,0,DISCONNECT,,,0,SOCKET
20221201 14:12:20,ip-10-22-1-51,rdsadmin,localhost,822,0,CONNECT,,,0,SOCKET
20221201 14:12:20,ip-10-22-1-51,rdsadmin,localhost,822,0,DISCONNECT,,,0,SOCKET
20221201 14:17:35,ip-10-22-1-51,rdsadmin,localhost,823,0,CONNECT,,,0,SOCKET
20221201 14:17:35,ip-10-22-1-51,rdsadmin,localhost,823,0,DISCONNECT,,,0,SOCKET
20221201 14:22:50,ip-10-22-1-51,rdsadmin,localhost,824,0,CONNECT,,,0,SOCKET
20221201 14:22:50,ip-10-22-1-51,rdsadmin,localhost,824,0,DISCONNECT,,,0,SOCKET
```

O exemplo de consulta do Log Insights mostra que `rdsadmin` se conectou ao banco de dados no `localhost` a cada cinco minutos, totalizando 22 vezes, conforme mostrado na ilustração a seguir. Esses resultados indicam que a atividade se originou de processos internos do Amazon RDS, como o próprio sistema de monitoramento.



Os eventos de logs frequentemente incluem mensagens importantes que você deseja contabilizar, como avisos ou erros sobre operações associadas às instâncias de banco de dados MySQL e

MariaDB. Por exemplo, se uma operação falhar, um erro poderá ocorrer e ser registrado no arquivo de logs de erros da seguinte forma: `ERROR 1114 (HY000): The table zip_codes is full`. É possível monitorar essas entradas para entender a tendência dos erros. Você pode criar métricas personalizadas do CloudWatch dos logs do Amazon RDS usando filtros para permitir o monitoramento automático dos logs do banco de dados do Amazon RDS para monitorar um log específico de padrões específicos e gerar um alarme se houver violações do comportamento esperado. Por exemplo, crie um filtro de métrica para o grupo de logs `/aws/rds/instance/database-1/error` que monitore o log de erros e busque o padrão específico, como `ERROR`. Defina o Padrão de filtro como `ERROR` e o Valor da métrica como `1`. O filtro detectará cada registro de log que tenha a palavra-chave `ERROR` e incrementará a contagem em 1 para cada evento de log que contenha “ERROR”. Depois de criar o filtro, você pode definir um alarme para notificar você caso sejam detectados erros no log de erros do MySQL ou do MariaDB.

Para saber mais sobre como monitorar o log de consultas lentas e o log de erros criando um painel do CloudWatch e usando o CloudWatch Logs Insights, consulte a publicação no blog Creating an Amazon CloudWatch dashboard to monitor Amazon RDS and Amazon Aurora MySQL.

Trilhas de auditoria

A trilha de auditoria (ou log de auditoria) fornece um registro cronológico relevante para a segurança dos eventos em sua Conta da AWS. Inclui eventos do Amazon RDS, que fornecem evidências documentais da sequência de atividades que afetaram seu banco de dados ou seu ambiente de nuvem. No Amazon RDS para MySQL ou MariaDB, o uso da trilha de auditoria envolve:

- Monitoramento do log de auditoria da instância de banco de dados
- Monitoramento de chamadas da API do Amazon RDS no AWS CloudTrail

Para uma instância de banco de dados Amazon RDS, os objetivos da auditoria geralmente incluem:

- Possibilitar a responsabilização pelo seguinte:
  - Modificações realizadas no parâmetro ou na configuração de segurança
  - Ações executadas em um esquema, tabela ou linha de banco de dados, ou ações que afetam um conteúdo específico
- Detecção e investigação de intrusões
- Detecção e investigação de atividades suspeitas

- Detecção de problemas de autorização, por exemplo, para identificar abusos de direitos de acesso por usuários regulares ou privilegiados

A trilha de auditoria do banco de dados tenta responder a estas perguntas frequentes: Quem visualizou ou modificou dados sensíveis em seu banco de dados? Quando isso ocorreu? De onde um usuário específico acessou os dados? Os usuários privilegiados abusaram de seus direitos de acesso ilimitado?

Tanto o MySQL quanto o MariaDB implementam o recurso de trilha de auditoria de instâncias de banco de dados usando o plug-in de auditoria do MariaDB. Esse plug-in registra a atividade do banco de dados, como usuários que fazem login no banco de dados, as consultas que são executadas no banco de dados e muito mais. O registro da atividade do banco de dados é armazenado em um arquivo de log. Para acessar o log de auditoria, a instância de banco de dados deve usar um grupo de opções personalizado com a opção `MARIADB_AUDIT_PLUGIN`. Para obter mais informações, consulte Suporte ao plug-in de auditoria do MariaDB para MySQL na documentação do Amazon RDS. Os registros no log de auditoria são armazenados em um formato específico, conforme definido pelo plug-in. Você pode encontrar mais detalhes sobre o formato do log de auditoria na documentação do MariaDB Server.

A trilha de auditoria da Nuvem AWS para sua conta da AWS é fornecida pelo serviço do AWS CloudTrail. O CloudTrail captura as chamadas de API para o Amazon RDS como eventos. Todas as ações do Amazon RDS são registradas em log. O CloudTrail fornece um registro de ações executadas no Amazon RDS por um usuário, um perfil ou outro serviço da AWS. Os eventos incluem ações executadas no Console de Gerenciamento da AWS, na AWS CLI e nos SDKs e APIs da AWS.

Exemplo

Em um cenário de auditoria comum, talvez seja necessário combinar trilhas do AWS CloudTrail com o log de auditoria do banco de dados e o monitoramento de eventos do Amazon RDS. Por exemplo, você pode ter um cenário em que os parâmetros do banco de dados da sua instância de banco de dados do Amazon RDS (por exemplo,`database-1`) tenham sido modificados e sua tarefa seja identificar quem fez a modificação, o que foi alterado e quando ocorreu a alteração.

Para realizar a tarefa, siga estas etapas:

1. Liste os eventos do Amazon RDS que ocorreram com a instância `database-1` do banco de dados e determine se há um evento na categoria `configuration change` que contém a mensagem `Finished updating DB parameter group`.

```
$ aws rds describe-events --source-identifier database-1 --source-type db-instance
{
    "Events": [
        {
            "SourceIdentifier": "database-1",
            "SourceType": "db-instance",
            "Message": "Finished updating DB parameter group",
            "EventCategories": [
                "configuration change"
            ],
            "Date": "2022-12-01T09:22:40.413000+00:00",
            "SourceArn": "arn:aws:rds:eu-west-3:111122223333:db:database-1"
        }
    ]
}
```

2. Identifique qual grupo de parâmetros de banco de dados a instância de banco de dados está usando:

```
$ aws rds describe-db-instances --db-instance-identifier database-1 --query
 'DBInstances[*].[DBInstanceIdentifier,Engine,DBParameterGroups]'
[
    [
        "database-1",
        "mariadb",
        [
            {
                "DBParameterGroupName": "mariadb10-6-test",
                "ParameterApplyStatus": "pending-reboot"
            }
        ]
    ]
]
```

3. Use a AWS CLI para pesquisar eventos do CloudTrail na região em que `database-1` está implantado, no período próximo ao evento do Amazon RDS descoberto na etapa 1, e onde `EventName=ModifyDBParameterGroup`.

```
$ aws cloudtrail --region eu-west-3 lookup-events --lookup-attributes
 AttributeKey=EventName,AttributeValue=ModifyDBParameterGroup --start-time
 "2022-12-01, 09:00 AM" --end-time "2022-12-01, 09:30 AM"
```

```
{
    "eventVersion": "1.08",
    "userIdentity": {
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AIDACKCEVSQ6C2EXAMPLE",
                "arn": "arn:aws:iam::111122223333:role/Role1",
                "accountId": "111122223333",
                "userName": "User1"
            }
        }
    },
    "eventTime": "2022-12-01T09:18:19Z",
    "eventSource": "rds.amazonaws.com",
    "eventName": "ModifyDBParameterGroup",
    "awsRegion": "eu-west-3",
    "sourceIPAddress": "AWS Internal",
    "userAgent": "AWS Internal",
    "requestParameters": {
        "parameters": [
            {
                "isModifiable": false,
                "applyMethod": "pending-reboot",
                "parameterName": "innodb_log_buffer_size",
                "parameterValue": "8388612"
            },
            {
                "isModifiable": false,
                "applyMethod": "pending-reboot",
                "parameterName": "innodb_write_io_threads",
                "parameterValue": "8"
            }
        ],
        "dBParameterGroupName": "mariadb10-6-test"
    },
    "responseElements": {
        "dBParameterGroupName": "mariadb10-6-test"
    },
    "requestID": "fdf19353-de72-4d3d-bf29-751f375b6378",
```

```
    "eventID": "0bba7484-0e46-4e71-93a8-bd01ca8386fe",
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "eventCategory": "Management",
    "sessionCredentialFromConsole": "true"
}
```

O evento do CloudTrail revela que o `User1`, com o perfil `Role1` da conta 111122223333 da AWS, modificou o grupo de parâmetros `mariadb10-6-test` do banco de dados, que foi usado pela instância `database-1` do banco de dados em `2022-12-01 at 09:18:19 h`. Dois parâmetros foram modificados e configurados com os seguintes valores:

- `innodb_log_buffer_size = 8388612`
- `innodb_write_io_threads = 8`

Recursos adicionais do CloudTrail e do CloudWatch Logs

É possível resolver problemas operacionais e incidentes de segurança dos últimos 90 dias no console do CloudTrail visualizando o Histórico de eventos. Para estender o período de retenção e aproveitar os recursos adicionais de consulta, você pode usar o AWS CloudTrail Lake. Com o AWS CloudTrail Lake, você pode manter os dados de eventos em um armazenamento de dados de eventos por até sete anos. Além disso, o serviço é compatível com consultas SQL complexas que oferecem uma visão mais profunda e personalizável dos eventos do que as visualizações fornecidas por pesquisas simples de chave/valor no Histórico de eventos.

Para monitorar suas trilhas de auditoria, definir alarmes e receber notificações quando uma atividade específica ocorrer, você precisa configurar o CloudTrail para enviar seus registros de trilha para o CloudWatch Logs. Depois que os registros da trilha forem armazenados como CloudWatch Logs, você poderá definir filtros de métricas para avaliar os eventos de logs de acordo com termos, frases ou valores e atribuir métricas aos filtros de métricas. Além disso, você pode criar alarmes do CloudWatch que são gerados de acordo com os limites e os períodos que você especificar. Por exemplo, você pode configurar alarmes que enviam notificações às equipes responsáveis para que elas possam tomar as medidas apropriadas. Você também pode configurar o CloudWatch para realizar uma ação automaticamente em resposta a um alarme.

Geração de alertas

Os alertas são uma das fontes de informação mais importantes quando se trata de segurança, disponibilidade, performance e confiabilidade de sua infraestrutura e serviços de TI. Eles notificam e informam suas equipes de TI sobre ameaças de segurança contínuas, interrupções, problemas de performance ou falhas no sistema.

A Biblioteca de Infraestrutura de Tecnologia da Informação (ITIL), especificamente as práticas de gerenciamento de serviços de TI (ITSM), define alertas automatizados no ponto focal das práticas recomendadas de monitoramento e gerenciamento de eventos e incidentes.

Os alertas de incidentes ocorrem quando as ferramentas de monitoramento geram alertas para notificar sua equipe e as ferramentas automatizadas (para itens que podem ser acionados automaticamente) sobre mudanças, ações de alto risco ou falhas no ambiente de TI. Os alertas de TI são a primeira linha de defesa contra mudanças ou interrupções do sistema que podem se transformar em incidentes graves. Ao monitorar automaticamente os sistemas e gerar alertas para interrupções e mudanças arriscadas, as equipes de TI podem minimizar o tempo de inatividade e reduzir o alto custo que o acompanha.

Como melhores práticas, o AWS Well-Architected Framework prescreve que você use o monitoramento para gerar notificações baseadas em alarmes e monitorar e alarmar proativamente. Use CloudWatch um serviço de monitoramento terceirizado para definir alarmes que indicam quando as métricas estão fora dos limites esperados.

O objetivo do gerenciamento de alertas é estabelecer procedimentos eficientes e padronizados para lidar com eventos e incidentes relacionados à TI por meio de registro em log, classificação, definição e implementação de ações, encerramento e atividades de análise pós-incidentes.

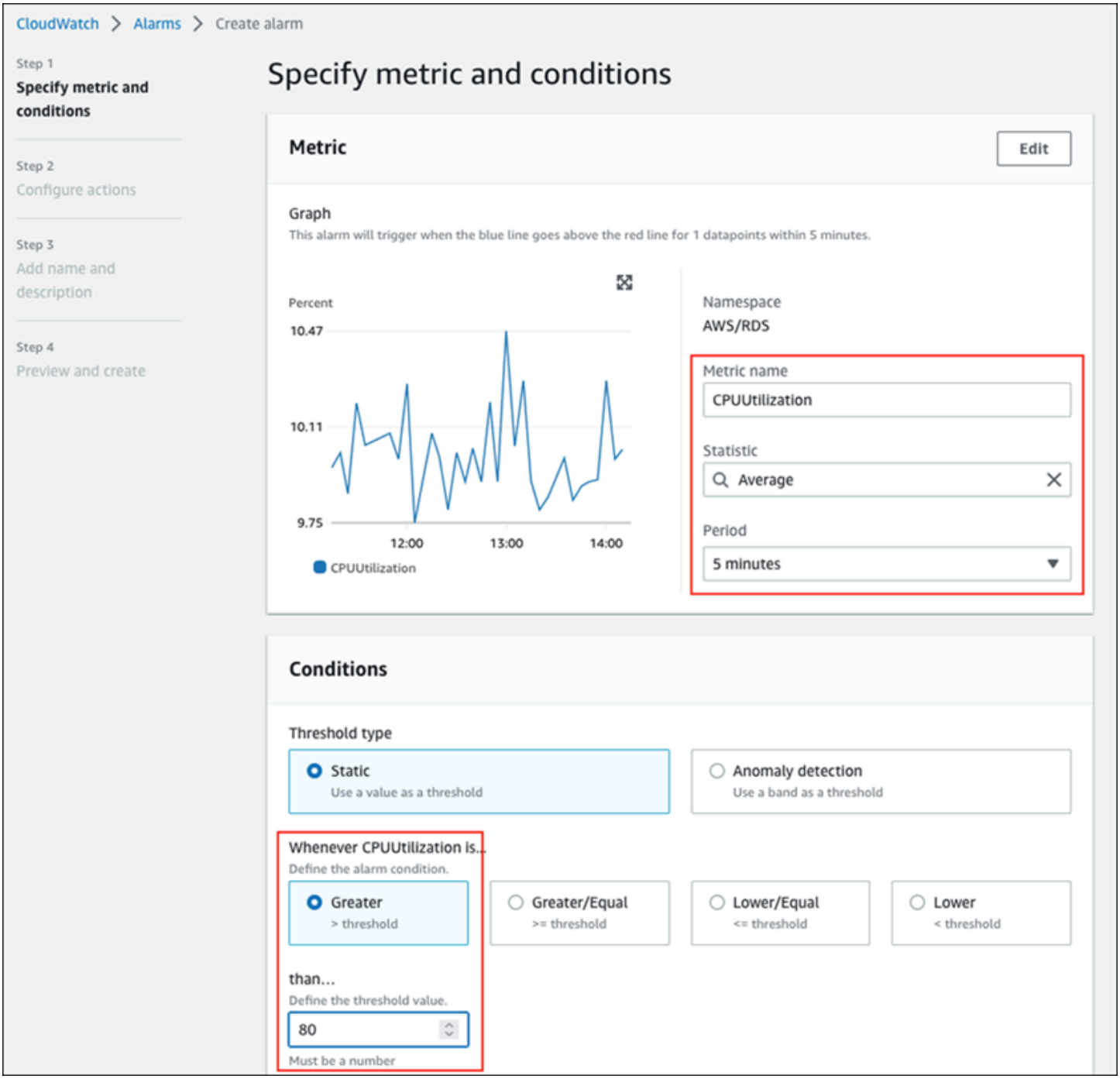
Seções

- CloudWatch alarmes
- EventBridge regras
- Especificação de ações e habilitação e desabilitação de alarmes

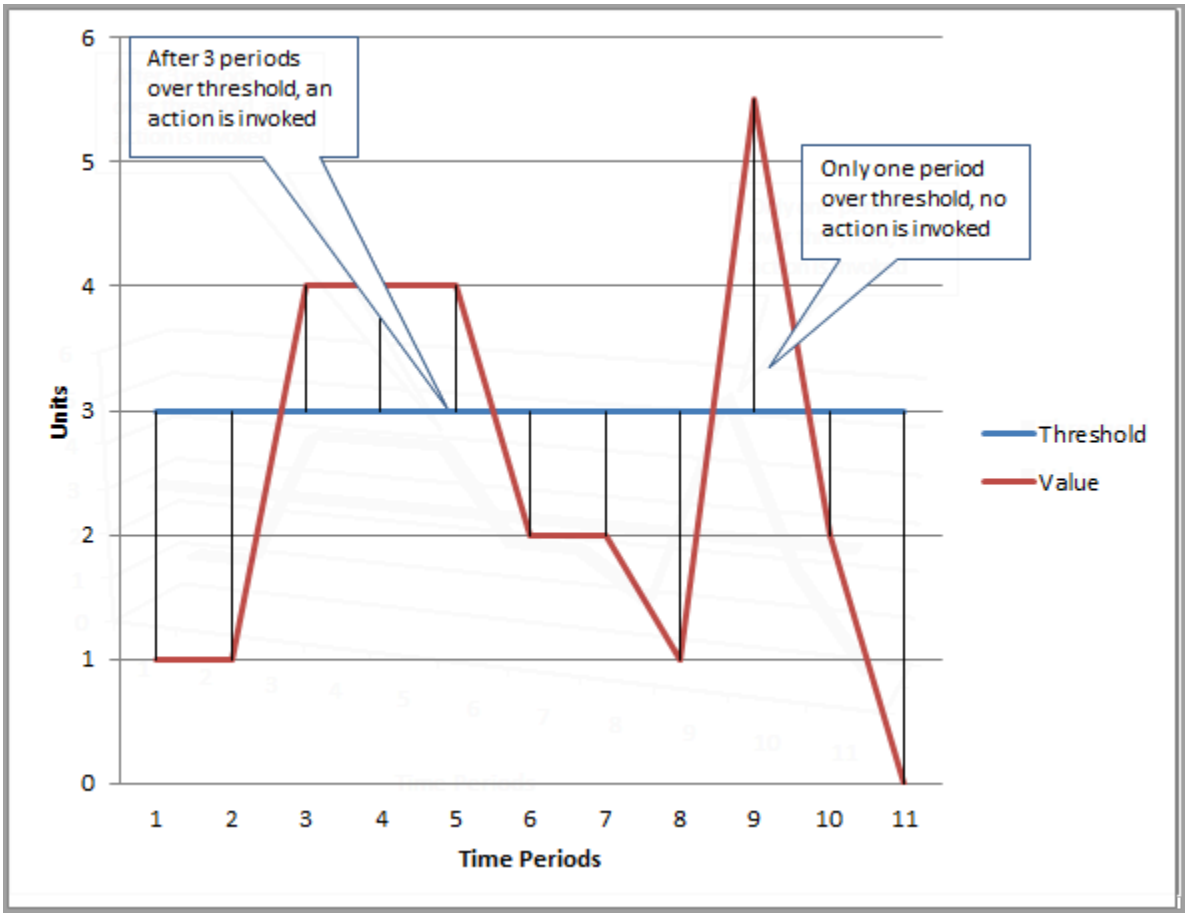
CloudWatch alarmes

Ao operar suas instâncias de banco de dados do Amazon RDS, você deseja monitorar e gerar alertas sobre diferentes tipos de métricas, eventos e rastreamentos. Para bancos de dados MySQL e MariaDB, as fontes críticas de informação são métricas das instâncias do banco de dados, métricas do sistema operacional, eventos, logs e trilhas de auditoria. Recomendamos que você use CloudWatch alarmes para observar uma única métrica durante um período especificado por você.

O exemplo a seguir ilustra como você pode definir um alarme que monitora a métrica `CPUUtilization` (porcentagem de utilização da CPU) em todas as suas instâncias de banco de dados Amazon RDS. Você configura o alarme para ser acionado se a utilização da CPU em qualquer instância de banco de dados for maior que 80% durante o período de avaliação de cinco minutos.



Isso significa que o alarme entrará no estado `ALARM` se algum de seus bancos de dados apresentar uma alta utilização da CPU (mais de 80%) por cinco minutos ou mais. O alarme permanecerá no estado `OK` se a CPU ocasionalmente atingir mais de 80% de utilização por um curto período de tempo e, em seguida, cair novamente para abaixo do limite. O grafo a seguir ilustra essa lógica.



CloudWatch os alarmes suportam alarmes métricos e compostos.

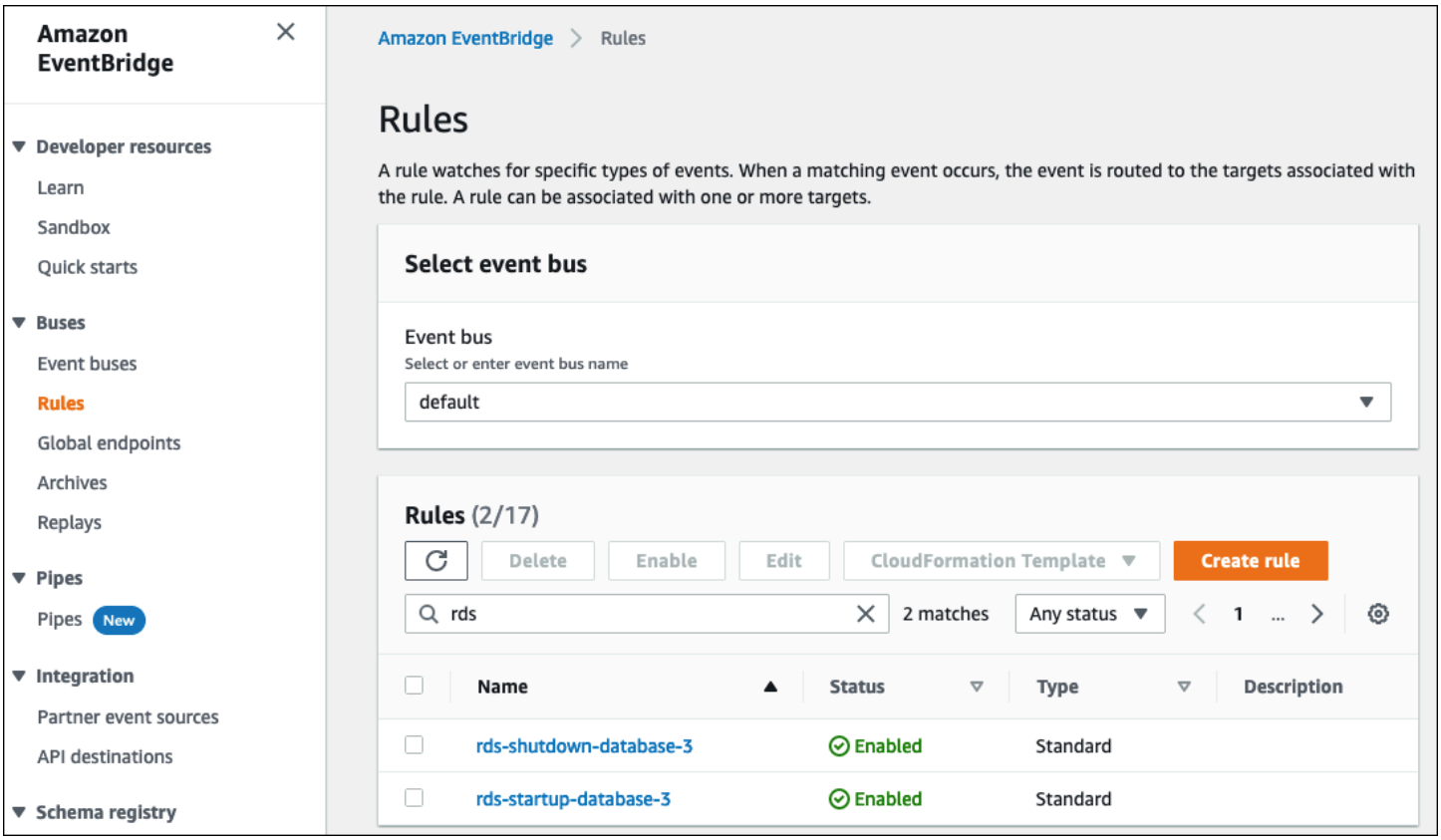
- Um alarme métrico observa uma única CloudWatch métrica e pode executar expressões matemáticas na métrica. Um alarme de métrica pode enviar mensagens para o Amazon SNS, que, por sua vez, pode executar uma ou mais ações com base no valor da métrica relativo a um determinado limite ao longo de vários períodos.
- Um alarme composto é baseado em uma expressão de regra, que avalia os estados de vários alarmes e entra no estado `ALARM` somente se todas as condições da regra são atendidas. Os alarmes compostos são normalmente usados para reduzir o número de alertas desnecessários. Por exemplo, você pode ter um alarme composto que contém vários alarmes de métricas configurados para nunca executar ações. O alarme composto enviará um alerta quando todos os alarmes individuais de métricas no composto já estivessem no estado `ALARM`

CloudWatch os alarmes só podem observar CloudWatch métricas. Se você quiser criar um alarme com base no erro, na consulta lenta ou nos registros gerais, deverá criar CloudWatch métricas a partir dos registros. Você pode realizar isso, conforme discutido anteriormente nas seções

Monitoramento do sistema operacional e Eventos, logs e trilhas de auditoria, usando filtros para criar métricas de eventos de logs. Da mesma forma, para alertar sobre métricas de monitoramento aprimorado, você deve criar filtros de métricas CloudWatch a partir dos CloudWatch registros.

EventBridge regras

Os eventos do Amazon RDS são entregues à Amazon EventBridge, e você pode usar EventBridge regras para reagir a esses eventos. Por exemplo, você pode criar EventBridge regras que o notificariam e tomariam uma ação se uma instância de banco de dados específica parasse ou fosse inicializada, conforme mostra a tela a seguir.

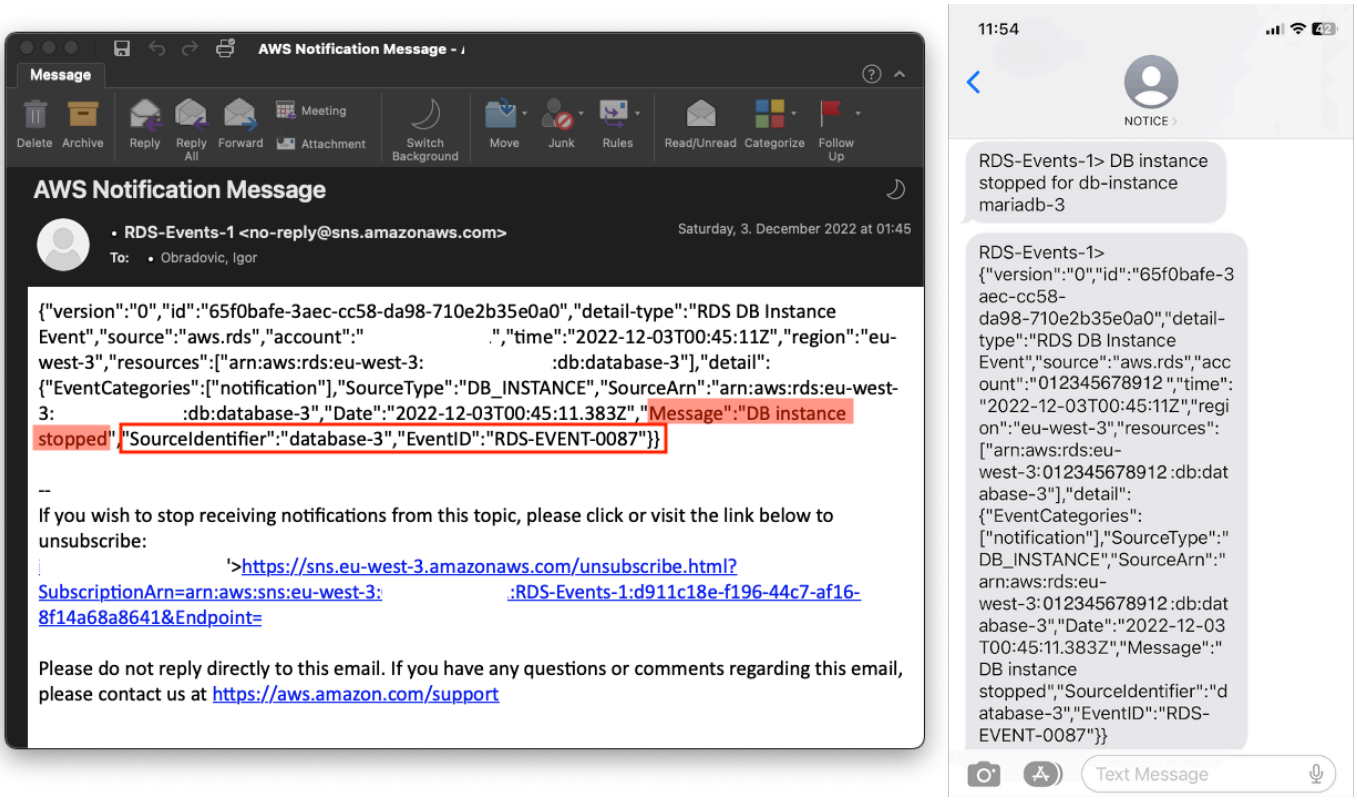


A regra que detecta o evento `The DB instance has been stopped` tem o ID `RDS-EVENT-0087` do evento do Amazon RDS, então você define a propriedade `Event Pattern` da regra como:

```
{
  "source": ["aws.rds"],
  "detail-type": ["RDS DB Instance Event"],
  "detail": {
```

```
    "SourceArn": ["arn:aws:rds:eu-west-3:111122223333:db:database-3"],
    "EventID": ["RDS-EVENT-0087"]
  }
}
```

Essa regra monitora somente a instância `database-3` de banco de dados e o evento `RDS-EVENT-0087`. Quando EventBridge detecta o evento, ele envia o evento para um recurso ou endpoint, conhecido como alvo. É aqui que você pode especificar a ação que deseja realizar se a instância do Amazon RDS for encerrada. Você pode enviar o evento para vários destinos possíveis, incluindo um tópico do SNS, uma fila do Amazon Simple Queue Service (Amazon SQS) AWS Lambda , uma AWS Systems Manager função, automação, um trabalho AWS Batch , Amazon API Gateway e muitos outros. Por exemplo, você pode criar um tópico do SNS que enviará um e-mail de notificação e SMS e atribuirá esse tópico do SNS como o destino da EventBridge regra. Se a instância de banco de dados do Amazon RDS `database-3` tiver sido interrompida, o Amazon RDS entrega o evento `RDS-EVENT-0087` para EventBridge, onde ele será detectado. EventBridge em seguida, chama o alvo, que é o tópico do SNS. O tópico do SNS está configurado para enviar um e-mail (conforme mostrado na ilustração a seguir) e um SMS.



Especificação de ações e habilitação e desabilitação de alarmes

Você pode usar um CloudWatch alarme para especificar quais ações o alarme deve tomar quando muda entre os `INSUFFICIENT_DATA` estados `OK``ALARM`, e. CloudWatch tem integração integrada com tópicos do SNS e várias categorias de ações adicionais que não são aplicáveis às métricas do Amazon RDS, como ações do Amazon Elastic Compute Cloud (Amazon EC2) ou ações de grupo do Amazon EC2 Auto Scaling. EventBridge geralmente é usado para escrever regras e definir metas que realizam ações quando o alarme é acionado para as métricas do Amazon RDS. CloudWatch envia eventos para EventBridge toda vez que um CloudWatch alarme muda de estado. Você pode usar esses eventos de mudança de estado de alarme para acionar um destino de evento EventBridge. Para obter mais informações, consulte Eventos de alarme e EventBridge na CloudWatch documentação.

Você talvez também precise gerenciar alarmes, por exemplo, desabilitar automaticamente um alarme durante testes ou alterações planejadas na configuração e, em seguida, reabilitar o alarme quando a ação planejada terminar. Por exemplo, se você tiver uma atualização planejada e programada do software do banco de dados que exija tempo de inatividade e tiver alarmes que serão ativados se o banco de dados ficar indisponível, você poderá desativar e ativar os alarmes usando as ações da API DisableAlarmActionse EnableAlarmActionsou os comandos disable-alarm-actionse enable-alarm-actionsno. AWS CLI Você também pode ver o histórico do alarme no CloudWatch console ou usando a ação da DescribeAlarmHistoryAPI ou o describe-alarm-historycomando no AWS CLI. CloudWatch preserva o histórico de alarmes por duas semanas. No CloudWatch console, você pode escolher o menu Favoritos e recentes no painel de navegação para definir e acessar seus alarmes favoritos e os mais visitados recentemente.

Próximas etapas e recursos

Para obter mais informações sobre como migrar seus bancos de dados relacionais para a Nuvem AWS, consulte a estratégia a seguir no site das Recomendações da AWS:

- Estratégia de migração para bancos de dados relacionais

Você pode explorar os padrões de migração de banco de dados nas Recomendações da AWS para obter instruções passo a passo sobre seus bancos de dados relacionais específicos executados na Nuvem AWS, incluindo tarefas relacionadas a monitoramento, migração e gerenciamento de dados.

Para recursos adicionais, consulte o seguinte:

- Guia do usuário do Amazon Relational Database Service
- Guia do usuário do Amazon CloudWatch
- Perguntas frequentes do Amazon RDS
- Perguntas frequentes do Insights de Performance
- Deliver Amazon RDS Performance Insights counter metrics to a third-party Application Performance Monitoring service provider using Amazon CloudWatch Metrics Stream (publicação do Blog da AWS)
- Creating an Amazon CloudWatch dashboard to monitor Amazon RDS and Amazon Aurora MySQL (publicação do Blog da AWS)
- Tuning Amazon RDS for MySQL with Performance Insights (publicação do Blog da AWS)

Histórico do documento

A tabela a seguir descreve alterações significativas feitas neste guia. Se desejar receber notificações sobre futuras atualizações, inscreva-se em um feed RSS.

Alteração	Descrição	Data
Atualizadas as informações sobre o Insights de Performance	Atualizada a seção sobre a publicação de métricas do Insights de Performance no CloudWatch com as informações mais recentes.	11 de março de 2025
Atualizadas as informações sobre exportadores	Atualizadas as informações sobre exportadores e adicionadas as orientações para escolher um exportador.	13 de junho de 2024
Publicação inicial	—	30 de junho de 2023

AWS Glossário de orientação prescritiva

A seguir estão os termos comumente usados em estratégias, guias e padrões fornecidos pela Orientação AWS Prescritiva. Para sugerir entradas, use o link Fornecer feedback no final do glossário.

Números

7 Rs

Sete estratégias comuns de migração para mover aplicações para a nuvem. Essas estratégias baseiam-se nos 5 Rs identificados pela Gartner em 2011 e consistem em:

- Refatorar/rearquitetar: mova uma aplicação e modifique sua arquitetura aproveitando ao máximo os recursos nativos de nuvem para melhorar a agilidade, a performance e a escalabilidade. Isso normalmente envolve a portabilidade do sistema operacional e do banco de dados. Exemplo: migrar seu banco de dados Oracle on-premises para o Amazon Aurora Edição Compatível com PostgreSQL.
- Redefinir a plataforma (mover e redefinir [mover e redefinir (lift-and-reshape]): mova uma aplicação para a nuvem e introduza algum nível de otimização a fim de aproveitar os recursos da nuvem. Exemplo: migrar seu banco de dados Oracle on-premises para o Amazon Relational Database Service (Amazon RDS) para Oracle na Nuvem AWS.
- Recomprar (drop and shop): mude para um produto diferente, normalmente migrando de uma licença tradicional para um modelo SaaS. Exemplo: migrar seu sistema de gerenciamento de relacionamento com o cliente (CRM) para o Salesforce.com.
- Redefinir a hospedagem (mover sem alterações [lift-and-shift])mover uma aplicação para a nuvem sem fazer nenhuma alteração a fim de aproveitar os recursos da nuvem. Exemplo: migrar seu banco de dados Oracle on-premises para o Oracle em uma instância do EC2 na Nuvem AWS.
- Realocar (mover o hipervisor sem alterações [hypervisor-level lift-and-shift]): mover a infraestrutura para a nuvem sem comprar novo hardware, reescrever aplicações ou modificar suas operações existentes. Você migra servidores de uma plataforma on-premises para um serviço de nuvem para a mesma plataforma. Exemplo: Migrar um Microsoft Hyper-V aplicativo para o. AWS
- Reter (revisitar): mantenha as aplicações em seu ambiente de origem. Isso pode incluir aplicações que exigem grande refatoração, e você deseja adiar esse trabalho para um

momento posterior, e aplicações antigas que você deseja manter porque não há justificativa comercial para migrá-las.

- Retirar: desative ou remova aplicações que não são mais necessárias em seu ambiente de origem.

A

ABAC

Consulte controle de acesso baseado em atributo.

serviços abstraídos

Veja serviços gerenciados.

ACID

Veja atomicidade, consistência, isolamento, durabilidade.

migração ativa-ativa

Um método de migração de banco de dados no qual os bancos de dados de origem e de destino são mantidos em sincronia (por meio de uma ferramenta de replicação bidirecional ou operações de gravação dupla), e ambos os bancos de dados lidam com transações de aplicações conectadas durante a migração. Esse método oferece suporte à migração em lotes pequenos e controlados, em vez de exigir uma substituição única. É mais flexível, mas exige mais trabalho do que a migração ativa-passiva.

migração ativa-passiva

Um método de migração de banco de dados em que os bancos de dados de origem e de destino são mantidos em sincronia, mas somente o banco de dados de origem manipula as transações das aplicações conectadas, enquanto os dados são replicados no banco de dados de destino. O banco de dados de destino não aceita nenhuma transação durante a migração.

AGGREGATE FUNCTION

Uma função SQL que opera em um grupo de linhas e calcula um único valor de retorno para o grupo. Exemplos de funções agregadas incluem `SUM` e `MAX`.

AI

Veja inteligência artificial.

AIOps

Veja operações de inteligência artificial.

anonimização

O processo de excluir permanentemente informações pessoais em um conjunto de dados. A anonimização pode ajudar a proteger a privacidade pessoal. Dados anônimos não são mais considerados dados pessoais.

antipadrões

Uma solução frequentemente usada para um problema recorrente em que a solução é contraproducente, ineficaz ou menos eficaz do que uma alternativa.

controle de aplicações

Uma abordagem de segurança que permite o uso somente de aplicações aprovadas para ajudar a proteger um sistema contra malware.

portfólio de aplicações

Uma coleção de informações detalhadas sobre cada aplicação usada por uma organização, incluindo o custo para criar e manter a aplicação e seu valor comercial. Essas informações são fundamentais para o processo de descoberta e análise de portfólio e ajudam a identificar e priorizar as aplicações a serem migradas, modernizadas e otimizadas.

inteligência artificial (IA)

O campo da ciência da computação que se dedica ao uso de tecnologias de computação para desempenhar funções cognitivas normalmente associadas aos humanos, como aprender, resolver problemas e reconhecer padrões. Para obter mais informações, consulte O que é inteligência artificial?

operações de inteligência artificial (AIOps)

O processo de usar técnicas de machine learning para resolver problemas operacionais, reduzir incidentes operacionais e intervenção humana e aumentar a qualidade do serviço. Para obter mais informações sobre como AIOps é usado na estratégia de AWS migração, consulte o guia de integração de operações.

criptografia assimétrica

Um algoritmo de criptografia que usa um par de chaves, uma chave pública para criptografia e uma chave privada para descriptografia. É possível compartilhar a chave pública porque ela não é usada na descriptografia, mas o acesso à chave privada deve ser altamente restrito.

atomicidade, consistência, isolamento, durabilidade (ACID)

Um conjunto de propriedades de software que garantem a validade dos dados e a confiabilidade operacional de um banco de dados, mesmo no caso de erros, falhas de energia ou outros problemas.

controle de acesso por atributo (ABAC)

A prática de criar permissões minuciosas com base nos atributos do usuário, como departamento, cargo e nome da equipe. Para obter mais informações, consulte ABAC AWS na documentação AWS Identity and Access Management (IAM).

fonte de dados autorizada

Um local onde você armazena a versão principal dos dados, que é considerada a fonte de informações mais confiável. Você pode copiar dados da fonte de dados autorizada para outros locais com o objetivo de processar ou modificar os dados, como anonimizá-los, redigi-los ou pseudonimizá-los.

Zona de disponibilidade

Um local distinto dentro de um Região da AWS que está isolado de falhas em outras zonas de disponibilidade e fornece conectividade de rede barata e de baixa latência a outras zonas de disponibilidade na mesma região.

AWS Estrutura de adoção da nuvem (AWS CAF)

Uma estrutura de diretrizes e melhores práticas AWS para ajudar as organizações a desenvolver um plano eficiente e eficaz para migrar com sucesso para a nuvem. AWS O CAF organiza a orientação em seis áreas de foco chamadas perspectivas: negócios, pessoas, governança, plataforma, segurança e operações. As perspectivas de negócios, pessoas e governança têm como foco habilidades e processos de negócios; as perspectivas de plataforma, segurança e operações concentram-se em habilidades e processos técnicos. Por exemplo, a perspectiva das pessoas tem como alvo as partes interessadas que lidam com recursos humanos (RH), funções de pessoal e gerenciamento de pessoal. Nessa perspectiva, o AWS CAF fornece orientação para desenvolvimento, treinamento e comunicação de pessoas para ajudar a preparar a organização

para a adoção bem-sucedida da nuvem. Para obter mais informações, consulte o site da AWS CAF e o whitepaper da AWS CAF.

AWS Estrutura de qualificação da carga de trabalho (AWS WQF)

Uma ferramenta que avalia as cargas de trabalho de migração do banco de dados, recomenda estratégias de migração e fornece estimativas de trabalho. AWS O WQF está incluído com AWS Schema Conversion Tool ()AWS SCT. Ela analisa esquemas de banco de dados e objetos de código, código de aplicações, dependências e características de performance, além de fornecer relatórios de avaliação.

B

bot malicioso

Um bot destinado a causar disrupção ou danos a indivíduos ou organizações.

BCP

Veja planejamento de continuidade de negócios

gráfico de comportamento

Uma visualização unificada e interativa do comportamento e das interações de recursos ao longo do tempo. É possível usar um gráfico de comportamento com o Amazon Detective para examinar tentativas de login malsucedidas, chamadas de API suspeitas e ações similares. Para obter mais informações, consulte Dados em um gráfico de comportamento na documentação do Detective.

sistema big-endian

Um sistema que armazena o byte mais significativo antes. Veja também endianness.

classificação binária

Um processo que prevê um resultado binário (uma de duas classes possíveis). Por exemplo, seu modelo de ML pode precisar prever problemas como “Este e-mail é ou não é spam?” ou “Este produto é um livro ou um carro?”

filtro de bloom

Uma estrutura de dados probabilística e eficiente em termos de memória que é usada para testar se um elemento é membro de um conjunto.

blue/green deployment (implantação azul/verde)

Uma estratégia de implantação em que você cria dois ambientes separados, mas idênticos. Você executa a versão atual da aplicação em um ambiente (azul) e a nova versão da aplicação no outro ambiente (verde). Essa estratégia ajuda você a reverter rapidamente com o mínimo de impacto.

bot

Uma aplicação de software que executa tarefas automatizadas na internet e simula a atividade ou interação humana. Alguns bots são úteis ou benéficos, como crawlers da web que indexam informações na internet. Outros bots, conhecidos como bots maliciosos, têm como objetivo causar disrupção ou danos a indivíduos ou organizações.

botnet

Redes de bots infectadas por malware e sob o controle de uma única parte, conhecidas como bot herder ou operador de bots. Os botnets são o mecanismo mais conhecido para escalar bots e seu impacto.

ramo

Uma área contida de um repositório de código. A primeira ramificação criada em um repositório é a ramificação principal. Você pode criar uma nova ramificação a partir de uma ramificação existente e, em seguida, desenvolver recursos ou corrigir bugs na nova ramificação. Uma ramificação que você cria para gerar um recurso é comumente chamada de ramificação de recurso. Quando o recurso estiver pronto para lançamento, você mesclará a ramificação do recurso de volta com a ramificação principal. Para obter mais informações, consulte Sobre filiais (GitHub documentação).

Acesso de emergência

Em circunstâncias excepcionais e por meio de um processo aprovado, um meio rápido para um usuário obter acesso a um Conta da AWS que ele normalmente não tem permissão para acessar. Para obter mais informações, consulte o indicador Implement break-glass procedures nas orientações do AWS Well-Architected.

estratégia brownfield

A infraestrutura existente em seu ambiente. Ao adotar uma estratégia brownfield para uma arquitetura de sistema, você desenvolve a arquitetura de acordo com as restrições dos sistemas e da infraestrutura atuais. Se estiver expandindo a infraestrutura existente, poderá combinar as estratégias brownfield e greenfield.

cache do buffer

A área da memória em que os dados acessados com mais frequência são armazenados.

capacidade de negócios

O que uma empresa faz para gerar valor (por exemplo, vendas, atendimento ao cliente ou marketing). As arquiteturas de microsserviços e as decisões de desenvolvimento podem ser orientadas por recursos de negócios. Para obter mais informações, consulte a seção Organizados de acordo com as capacidades de negócios do whitepaper Executar microsserviços conteinerizados na AWS.

planejamento de continuidade de negócios (BCP)

Um plano que aborda o impacto potencial de um evento disruptivo, como uma migração em grande escala, nas operações e permite que uma empresa retome as operações rapidamente.

C

CAF

Veja AWS Cloud Adoption Framework.

implantação canário

O lançamento lento e incremental de uma versão para usuários finais. Quando estiver confiante, você implanta a nova versão e substitui a versão atual por completo.

CCoE

Veja Centro de Excelência da Nuvem.

CDC

Veja captura de dados de alteração.

captura de dados de alterações (CDC)

O processo de rastrear alterações em uma fonte de dados, como uma tabela de banco de dados, e registrar metadados sobre a alteração. É possível usar o CDC para várias finalidades, como auditar ou replicar alterações em um sistema de destino para manter a sincronização.

engenharia do caos

Introduzir intencionalmente falhas ou eventos disruptivos para testar a resiliência de um sistema. Você pode usar AWS Fault Injection Service (AWS FIS) para realizar experimentos que estressam suas AWS cargas de trabalho e avaliar sua resposta.

CI/CD

Veja integração e entrega contínuas.

classificação

Um processo de categorização que ajuda a gerar previsões. Os modelos de ML para problemas de classificação predizem um valor discreto. Os valores discretos são sempre diferentes uns dos outros. Por exemplo, um modelo pode precisar avaliar se há ou não um carro em uma imagem.

criptografia no lado do cliente

Criptografia de dados localmente, antes que o alvo os AWS service (Serviço da AWS) receba.

Centro de excelência em nuvem (CCoE)

Uma equipe multidisciplinar que impulsiona os esforços de adoção da nuvem em toda a organização, incluindo o desenvolvimento de práticas recomendas de nuvem, a mobilização de recursos, o estabelecimento de cronogramas de migração e a liderança da organização em transformações em grande escala. Para obter mais informações, consulte as publicações CCo E no blog de estratégia Nuvem AWS corporativa.

computação em nuvem

A tecnologia de nuvem normalmente usada para armazenamento de dados remoto e gerenciamento de dispositivos de IoT. A computação em nuvem é normalmente conectada à tecnologia de computação de borda.

modelo operacional em nuvem

Em uma organização de TI, o modelo operacional usado para criar, amadurecer e otimizar um ou mais ambientes de nuvem. Para obter mais informações, consulte Criar seu modelo operacional de nuvem.

estágios de adoção da nuvem

As quatro fases pelas quais as organizações normalmente passam ao migrar para a Nuvem AWS:

- Projeto: executar alguns projetos relacionados à nuvem para fins de prova de conceito e aprendizado
- Fundação — Fazer investimentos fundamentais para escalar sua adoção da nuvem (por exemplo, criar uma landing zone, definir um CCo E, estabelecer um modelo de operações)
- Migração: migrar aplicações individuais
- Reinvenção: otimizar produtos e serviços e inovar na nuvem

Esses estágios foram definidos por Stephen Orban na postagem do blog The Journey Toward Cloud-First & the Stages of Adoption no blog de estratégia Nuvem AWS empresarial. Para obter informações sobre como eles se relacionam com a estratégia de AWS migração, consulte o guia de preparação para migração.

CMDB

Veja banco de dados de gerenciamento de configuração.

repositório de código

Um local onde o código-fonte e outros ativos, como documentação, amostras e scripts, são armazenados e atualizados por meio de processos de controle de versão. Os repositórios de nuvem comuns incluem o GitHub ou o Bitbucket Cloud. Cada versão do código é chamada de ramificação. Em uma estrutura de microsserviços, cada repositório é dedicado a uma única peça de funcionalidade. Um único pipeline de CI/CD pode usar vários repositórios.

cache frio

Um cache de buffer que está vazio, não está bem preenchido ou contém dados obsoletos ou irrelevantes. Isso afeta a performance porque a instância do banco de dados deve ler da memória principal ou do disco, um processo que é mais lento do que a leitura do cache do buffer.

dados frios

Dados que raramente são acessados e geralmente são históricos. Ao consultar esse tipo de dados, consultas lentas geralmente são aceitáveis. Mover esses dados para níveis ou classes de armazenamento de baixo desempenho e menos caros pode reduzir os custos.

visão computacional (CV)

Um campo de IA que usa machine learning para analisar e extrair informações de formatos visuais, como vídeos e imagens digitais. Por exemplo, a Amazon SageMaker AI fornece algoritmos de processamento de imagem para CV.

desvio de configuração

Em uma workload, uma alteração de configuração em relação ao estado esperado. Isso pode fazer com que a workload se torne incompatível e, normalmente, é gradual e não intencional.

banco de dados de gerenciamento de configuração (CMDB)

Um repositório que armazena e gerencia informações sobre um banco de dados e seu ambiente de TI, incluindo componentes de hardware e software e suas configurações. Normalmente, os dados de um CMDB são usados no estágio de descoberta e análise do portfólio da migração.

pacote de conformidade

Um conjunto de AWS Config regras e ações de remediação que você pode montar para personalizar suas verificações de conformidade e segurança. Você pode implantar um pacote de conformidade como uma entidade única em uma Conta da AWS região ou em uma organização usando um modelo YAML. Para obter mais informações, consulte Pacotes de conformidade na documentação. AWS Config

integração contínua e entrega contínua (CI/CD)

O processo de automatizar os estágios de origem, criação, teste, preparação e produção do processo de lançamento do software. CI/CD é comumente descrito como um pipeline. CI/CD pode ajudá-lo a automatizar processos, melhorar a produtividade, melhorar a qualidade do código e entregar com mais rapidez. Para obter mais informações, consulte Benefícios da entrega contínua. CD também pode significar implantação contínua. Para obter mais informações, consulte Entrega contínua versus implantação contínua.

CV

Veja visão computacional.

D

dados em repouso

Dados estacionários em sua rede, por exemplo, dados que estão em um armazenamento.

classificação de dados

Um processo para identificar e categorizar os dados em sua rede com base em criticalidade e confidencialidade. É um componente crítico de qualquer estratégia de gerenciamento de riscos de

segurança cibernética, pois ajuda a determinar os controles adequados de proteção e retenção para os dados. A classificação de dados é um componente do pilar de segurança no AWS Well-Architected Framework. Para obter mais informações, consulte Classificação de dados.

desvio de dados

Uma variação significativa entre os dados de produção e os dados usados para treinar um modelo de ML ou uma alteração significativa nos dados de entrada ao longo do tempo. O desvio de dados pode reduzir a qualidade geral, a precisão e a imparcialidade das previsões do modelo de ML.

dados em trânsito

Dados que estão se movendo ativamente pela sua rede, como entre os recursos da rede.

data mesh

Um framework de arquitetura que fornece propriedade de dados distribuída e descentralizada com gerenciamento e governança centralizados.

minimização de dados

O princípio de coletar e processar apenas os dados estritamente necessários. Praticar a minimização de dados no Nuvem AWS pode reduzir os riscos de privacidade, os custos e a pegada de carbono de sua análise.

perímetro de dados

Um conjunto de proteções preventivas em seu AWS ambiente que ajudam a garantir que somente identidades confiáveis acessem recursos confiáveis das redes esperadas. Para obter mais informações, consulte Construindo um perímetro de dados em. AWS

pré-processamento de dados

A transformação de dados brutos em um formato que seja facilmente analisado por seu modelo de ML. O pré-processamento de dados pode significar a remoção de determinadas colunas ou linhas e o tratamento de valores ausentes, inconsistentes ou duplicados.

proveniência dos dados

O processo de rastrear a origem e o histórico dos dados ao longo de seu ciclo de vida, por exemplo, como os dados foram gerados, transmitidos e armazenados.

titular dos dados

Um indivíduo cujos dados estão sendo coletados e processados.

data warehouse

Um sistema de gerenciamento de dados compatível com business intelligence, como analytics. Os data warehouses geralmente contêm grandes quantidades de dados históricos e geralmente são usados para consultas e análises.

linguagem de definição de dados (DDL)

Instruções ou comandos para criar ou modificar a estrutura de tabelas e objetos em um banco de dados.

linguagem de manipulação de dados (DML)

Instruções ou comandos para modificar (inserir, atualizar e excluir) informações em um banco de dados.

DDL

Veja linguagem de definição de banco de dados.

deep ensemble

A combinação de vários modelos de aprendizado profundo para gerar previsões. Os deep ensembles podem ser usados para produzir uma previsão mais precisa ou para estimar a incerteza nas previsões.

Aprendizado profundo

Um subcampo do ML que usa várias camadas de redes neurais artificiais para identificar o mapeamento entre os dados de entrada e as variáveis-alvo de interesse.

defense-in-depth

Uma abordagem de segurança da informação na qual uma série de mecanismos e controles de segurança são cuidadosamente distribuídos por toda a rede de computadores para proteger a confidencialidade, a integridade e a disponibilidade da rede e dos dados nela contidos. Ao adotar essa estratégia AWS, você adiciona vários controles em diferentes camadas da AWS Organizations estrutura para ajudar a proteger os recursos. Por exemplo, uma defense-in-depth abordagem pode combinar autenticação multifatorial, segmentação de rede e criptografia.

administrador delegado

Em AWS Organizations, um serviço compatível pode registrar uma conta de AWS membro para administrar as contas da organização e gerenciar as permissões desse serviço. Essa conta

é chamada de administrador delegado para esse serviço Para obter mais informações e uma lista de serviços compatíveis, consulte Serviços que funcionam com o AWS Organizations na documentação do AWS Organizations .

implantação

O processo de criar uma aplicação, novos recursos ou correções de código disponíveis no ambiente de destino. A implantação envolve a implementação de mudanças em uma base de código e, em seguida, a criação e execução dessa base de código nos ambientes da aplicação

ambiente de desenvolvimento

Veja ambiente.

controle detectivo

Um controle de segurança projetado para detectar, registrar e alertar após a ocorrência de um evento. Esses controles são uma segunda linha de defesa, alertando você sobre eventos de segurança que contornaram os controles preventivos em vigor. Para obter mais informações, consulte Controles detectivos em Como implementar controles de segurança na AWS.

mapeamento do fluxo de valor de desenvolvimento (DVSM)

Um processo usado para identificar e priorizar restrições que afetam negativamente a velocidade e a qualidade em um ciclo de vida de desenvolvimento de software. O DVSM estende o processo de mapeamento do fluxo de valor originalmente projetado para práticas de manufatura enxuta. Ele se concentra nas etapas e equipes necessárias para criar e movimentar valor por meio do processo de desenvolvimento de software.

gêmeo digital

Uma representação virtual de um sistema real, como um prédio, fábrica, equipamento industrial ou linha de produção. Os gêmeos digitais oferecem suporte à manutenção preditiva, ao monitoramento remoto e à otimização da produção.

tabela de dimensões

Em um esquema em estrela, uma tabela menor que contém atributos de dados sobre dados quantitativos em uma tabela de fatos. Os atributos da tabela de dimensões geralmente são campos de texto ou números discretos que se comportam como texto. Esses atributos normalmente são usados para restringir consultas, filtrar e rotular conjuntos de resultados.

desastre

Um evento que impede que uma workload ou sistema cumpra seus objetivos de negócios em seu local principal de implantação. Esses eventos podem ser desastres naturais, falhas técnicas ou o resultado de ações humanas, como configuração incorreta não intencional ou ataque de malware.

Recuperação de desastres (RD)

A estratégia e o processo que você usa para minimizar o tempo de inatividade e a perda de dados causados por um disastre. Para obter mais informações, consulte Recuperação de desastres de cargas de trabalho em AWS: Recuperação na nuvem no AWS Well-Architected Framework.

DML

Veja linguagem de manipulação de banco de dados.

design orientado por domínio

Uma abordagem ao desenvolvimento de um sistema de software complexo conectando seus componentes aos domínios em evolução, ou principais metas de negócios, atendidos por cada componente. Esse conceito foi introduzido por Eric Evans em seu livro, Design orientado por domínio: lidando com a complexidade no coração do software (Boston: Addison-Wesley Professional, 2003). Para obter informações sobre como usar o design orientado por domínio com o padrão strangler fig, consulte Modernizar incrementalmente os serviços web herdados do Microsoft ASP.NET (ASMX) usando contêineres e o Amazon API Gateway.

DR

Veja recuperação de desastres.

Detecção da oscilação

Rastreamento de desvios de uma configuração de linha de base. Por exemplo, você pode usar AWS CloudFormation para detectar desvios nos recursos do sistema ou AWS Control Tower para detectar mudanças em seu landing zone que possam afetar a conformidade com os requisitos de governança.

DVSM

Veja mapeamento do fluxo de valor de desenvolvimento.

E

EDA

Veja análise exploratória de dados.

EDI

Veja intercâmbio eletrônico de dados.

computação de borda

A tecnologia que aumenta o poder computacional de dispositivos inteligentes nas bordas de uma rede de IoT. Quando comparada com a computação em nuvem, a computação de borda pode reduzir a latência da comunicação e melhorar o tempo de resposta.

intercâmbio eletrônico de dados (EDI)

A troca automatizada de documentos comerciais entre organizações. Para obter mais informações, consulte O que é EDI (Intercâmbio eletrônico de dados)?.

criptografia

Um processo de computação que transforma dados de texto simples, legíveis por humanos, em texto cifrado.

chave de criptografia

Uma sequência criptográfica de bits aleatórios que é gerada por um algoritmo de criptografia. As chaves podem variar em tamanho, e cada chave foi projetada para ser imprevisível e exclusiva.

endianismo

A ordem na qual os bytes são armazenados na memória do computador. Os sistemas big-endian armazenam o byte mais significativo antes. Os sistemas little-endian armazenam o byte menos significativo antes.

endpoint

Veja endpoint de serviço.

serviço de endpoint

Um serviço que pode ser hospedado em uma nuvem privada virtual (VPC) para ser compartilhado com outros usuários. Você pode criar um serviço de endpoint com AWS PrivateLink e conceder permissões a outros diretores Contas da AWS ou a AWS Identity and Access Management (IAM).

Essas contas ou entidades principais podem se conectar ao serviço de endpoint de maneira privada criando endpoints da VPC de interface. Para obter mais informações, consulte Criar um serviço de endpoint na documentação do Amazon Virtual Private Cloud (Amazon VPC).

planejamento de recursos empresariais (ERP)

Um sistema que automatiza e gerencia os principais processos de negócios (como contabilidade, MES e gerenciamento de projetos) para uma empresa.

criptografia envelopada

O processo de criptografar uma chave de criptografia com outra chave de criptografia. Para obter mais informações, consulte Criptografia de envelope na documentação AWS Key Management Service (AWS KMS).

ambiente

Uma instância de uma aplicação em execução. Estes são tipos comuns de ambientes na computação em nuvem:

- ambiente de desenvolvimento: uma instância de uma aplicação em execução que está disponível somente para a equipe principal responsável pela manutenção da aplicação. Ambientes de desenvolvimento são usados para testar mudanças antes de promovê-las para ambientes superiores. Esse tipo de ambiente às vezes é chamado de ambiente de teste.
- ambientes inferiores: todos os ambientes de desenvolvimento para uma aplicação, como aqueles usados para compilações e testes iniciais.
- ambiente de produção: uma instância de uma aplicação em execução que os usuários finais podem acessar. Em um CI/CD pipeline, o ambiente de produção é o último ambiente de implantação.
- ambientes superiores: todos os ambientes que podem ser acessados por usuários que não sejam a equipe principal de desenvolvimento. Isso pode incluir um ambiente de produção, ambientes de pré-produção e ambientes para testes de aceitação do usuário.

epic

Em metodologias ágeis, categorias funcionais que ajudam a organizar e priorizar seu trabalho. Os epics fornecem uma descrição de alto nível dos requisitos e das tarefas de implementação. Por exemplo, os épicos de segurança AWS da CAF incluem gerenciamento de identidade e acesso, controles de detetive, segurança de infraestrutura, proteção de dados e resposta a incidentes. Para obter mais informações sobre epics na estratégia de migração da AWS , consulte o guia de implementação do programa.

ERP

Veja planejamento de recursos empresariais.

análise exploratória de dados (EDA)

O processo de analisar um conjunto de dados para entender suas principais características. Você coleta ou agrega dados e, em seguida, realiza investigações iniciais para encontrar padrões, detectar anomalias e verificar suposições. O EDA é realizado por meio do cálculo de estatísticas resumidas e da criação de visualizações de dados.

F

tabela de fatos

A tabela central em um esquema em estrela. Ela armazena dados quantitativos sobre as operações comerciais. Normalmente, uma tabela de fatos contém dois tipos de colunas: as que contêm medidas e as que contêm uma chave externa para uma tabela de dimensões.

Antecipar-se à falha

Uma filosofia que usa testes frequentes e incrementais para reduzir o ciclo de vida do desenvolvimento. É uma parte essencial de uma abordagem ágil.

delimitação de isolamento contra falhas

No Nuvem AWS, um limite, como uma zona de disponibilidade, Região da AWS um plano de controle ou um plano de dados, que limita o efeito de uma falha e ajuda a melhorar a resiliência das cargas de trabalho. Para obter mais informações, consulte AWS Fault Isolation Boundaries.

ramificação de recursos

Veja ramificação.

recursos

Os dados de entrada usados para fazer uma previsão. Por exemplo, em um contexto de manufatura, os recursos podem ser imagens capturadas periodicamente na linha de fabricação.

importância do recurso

O quanto um recurso é importante para as previsões de um modelo. Isso geralmente é expresso como uma pontuação numérica que pode ser calculada por meio de várias técnicas, como

Shapley Additive Explanations (SHAP) e gradientes integrados. Para obter mais informações, consulte Interpretabilidade do modelo de aprendizado de máquina com AWS.

transformação de recursos

O processo de otimizar dados para o processo de ML, incluindo enriquecer dados com fontes adicionais, escalar valores ou extrair vários conjuntos de informações de um único campo de dados. Isso permite que o modelo de ML se beneficie dos dados. Por exemplo, se a data “2021-05-27 00:15:37” for dividida em “2021”, “maio”, “quinta” e “15”, isso poderá ajudar o algoritmo de aprendizado a aprender padrões diferenciados associados a diferentes componentes de dados.

prompt few shot

Fornecer a um LLM um pequeno número de exemplos que demonstram a tarefa e o resultado desejado antes de solicitar que ele execute uma tarefa semelhante. Essa técnica é uma aplicação do aprendizado em contexto, em que os modelos aprendem com exemplos (shots) incorporados aos prompts. Prompts few-shot podem ser eficazes para tarefas que exigem formatação, raciocínio ou conhecimento de domínio específicos. Veja também prompts zero-shot.

FGAC

Veja controle de acesso refinado.

Controle de acesso refinado (FGAC)

O uso de várias condições para permitir ou negar uma solicitação de acesso.

migração flash-cut

Um método de migração de banco de dados que usa replicação contínua de dados via captura de dados de alteração para migrar os dados no menor tempo possível, em vez de usar uma abordagem em fases. O objetivo é reduzir ao mínimo o tempo de inatividade.

FM

Veja modelo de base.

modelo de base (FM)

Uma grande rede neural de aprendizado profundo que vem treinando em grandes conjuntos de dados generalizados e não rotulados. FMs são capazes de realizar uma ampla variedade de tarefas gerais, como entender a linguagem, gerar texto e imagens e conversar em linguagem natural. Para obter mais informações, consulte O que são modelos de base?.

G

IA generativa

Um subconjunto de modelos de IA que foram treinados em grandes quantidades de dados e que podem usar um simples prompt de texto para criar novos artefatos e conteúdo, como imagens, vídeos, texto e áudio. Para obter mais informações, consulte O que é IA generativa?.

bloqueio geográfico

Veja restrições geográficas.

restrições geográficas (bloqueio geográfico)

Na Amazon CloudFront, uma opção para impedir que usuários em países específicos acessem distribuições de conteúdo. É possível usar uma lista de permissões ou uma lista de bloqueios para especificar países aprovados e banidos. Para obter mais informações, consulte Restringir a distribuição geográfica do seu conteúdo na CloudFront documentação.

Fluxo de trabalho do GitFlow

Uma abordagem na qual ambientes inferiores e superiores usam ramificações diferentes em um repositório de código-fonte. O fluxo de trabalho do Gitflow é considerado legado, e o fluxo de trabalho trunk-based é a abordagem moderna e preferencial.

golden image

Um snapshot de um sistema ou software usado como modelo para implantar novas instâncias desse sistema ou software. Por exemplo, na manufatura, uma golden image pode ser usada para provisionar software em vários dispositivos e ajudar a melhorar a velocidade, a escalabilidade e a produtividade nas operações de fabricação de dispositivos.

estratégia greenfield

A ausência de infraestrutura existente em um novo ambiente. Ao adotar uma estratégia greenfield para uma arquitetura de sistema, é possível selecionar todas as novas tecnologias sem a restrição da compatibilidade com a infraestrutura existente, também conhecida como brownfield. Se estiver expandindo a infraestrutura existente, poderá combinar as estratégias brownfield e greenfield.

barreira de proteção

Uma regra de alto nível que ajuda a governar recursos, políticas e conformidade em todas as unidades organizacionais ()OUs. Barreiras de proteção preventivas impõem políticas para

garantir o alinhamento a padrões de conformidade. Elas são implementadas usando políticas de controle de serviço e limites de permissões do IAM. Barreiras de proteção detectivas detectam violações de políticas e problemas de conformidade e geram alertas para remediação. Eles são implementados usando AWS Config, AWS Security Hub CSPM, Amazon GuardDuty AWS Trusted Advisor, Amazon Inspector e verificações personalizadas AWS Lambda .

H

HA

Veja alta disponibilidade.

migração heterogênea de bancos de dados

Migrar seu banco de dados de origem para um banco de dados de destino que usa um mecanismo de banco de dados diferente (por exemplo, Oracle para Amazon Aurora). A migração heterogênea geralmente faz parte de um esforço de redefinição da arquitetura, e converter o esquema pode ser uma tarefa complexa. O AWS fornece o AWS SCT para ajudar nas conversões de esquemas.

alta disponibilidade (HA)

A capacidade de uma workload operar continuamente, sem intervenção, em caso de desafios ou desastres. Os sistemas AH são projetados para realizar o failover automático, oferecer consistentemente desempenho de alta qualidade e lidar com diferentes cargas e falhas com impacto mínimo no desempenho.

modernização de historiador

Uma abordagem usada para modernizar e atualizar os sistemas de tecnologia operacional (OT) para melhor atender às necessidades do setor de manufatura. Um historiador é um tipo de banco de dados usado para coletar e armazenar dados de várias fontes em uma fábrica.

dados de hold-out

Uma parte dos dados históricos rotulados que são retidos de um conjunto de dados usado para treinar um modelo de machine learning. Você pode usar dados de hold-out para avaliar a performance do modelo comparando as predições do modelo com os dados de retenção.

migração homogênea de bancos de dados

Migrar seu banco de dados de origem para um banco de dados de destino que compartilha o mesmo mecanismo de banco de dados (por exemplo, Microsoft SQL Server para Amazon RDS para SQL Server). A migração homogênea geralmente faz parte de um esforço de redefinição da hospedagem ou da plataforma. É possível usar utilitários de banco de dados nativos para migrar o esquema.

dados quentes

Dados acessados com frequência, como dados em tempo real ou dados translacionais recentes. Esses dados normalmente exigem uma camada ou classe de armazenamento de alto desempenho para fornecer respostas rápidas às consultas.

hotfix

Uma correção urgente para um problema crítico em um ambiente de produção. Devido à sua urgência, um hotfix geralmente é feito fora do fluxo de trabalho normal de DevOps lançamento.

período de hipercuidados

Imediatamente após a substituição, o período em que uma equipe de migração gerencia e monitora as aplicações migradas na nuvem para resolver quaisquer problemas. Normalmente, a duração desse período é de 1 a 4 dias. No final do período de hipercuidados, a equipe de migração normalmente transfere a responsabilidade pelas aplicações para a equipe de operações de nuvem.

eu

IaC

Veja infraestrutura como código.

Política baseada em identidade

Uma política anexada a um ou mais diretores do IAM que define suas permissões no Nuvem AWS ambiente.

aplicação ociosa

Uma aplicação que tem um uso médio de CPU e memória entre 5 e 20% em um período de 90 dias. Em um projeto de migração, é comum retirar essas aplicações ou retê-las on-premises.

IIoT

Veja Internet das Coisas Industrial.

infraestrutura imutável

Um modelo que implanta uma nova infraestrutura para workloads de produção em vez de atualizar, aplicar patches ou modificar a infraestrutura existente. Infraestruturas imutáveis são inerentemente mais consistentes, confiáveis e preditivas do que infraestruturas mutáveis. Para obter mais informações, consulte a prática recomendada Implantar usando infraestrutura imutável no AWS Well-Architected Framework.

VPC de entrada (admissão)

Em uma arquitetura de AWS várias contas, uma VPC que aceita, inspeciona e roteia conexões de rede de fora de um aplicativo. A Arquitetura de Referência de AWS Segurança recomenda configurar sua conta de rede com entrada, saída e inspeção VPCs para proteger a interface bidirecional entre seu aplicativo e a Internet em geral.

migração incremental

Uma estratégia de substituição na qual você migra a aplicação em pequenas partes, em vez de realizar uma única substituição completa. Por exemplo, é possível mover inicialmente apenas alguns microsserviços ou usuários para o novo sistema. Depois de verificar se tudo está funcionando corretamente, mova os microsserviços ou usuários adicionais de forma incremental até poder descomissionar seu sistema herdado. Essa estratégia reduz os riscos associados a migrações de grande porte.

Indústria 4.0

Um termo que foi introduzido por Klaus Schwab em 2016 para se referir à modernização dos processos de manufatura por meio de avanços em conectividade, dados em tempo real, automação, analytics e IA/ML.

infraestrutura

Todos os recursos e ativos contidos no ambiente de uma aplicação.

Infraestrutura como código (IaC)

O processo de provisionamento e gerenciamento da infraestrutura de uma aplicação por meio de um conjunto de arquivos de configuração. A IaC foi projetada para ajudar você a centralizar o gerenciamento da infraestrutura, padronizar recursos e escalar rapidamente para que novos ambientes sejam reproduzíveis, confiáveis e consistentes.

Internet industrial das coisas (IIoT)

O uso de sensores e dispositivos conectados à Internet nos setores industriais, como manufatura, energia, automotivo, saúde, ciências biológicas e agricultura. Para obter mais informações, consulte Criando uma estratégia de transformação digital industrial da Internet das Coisas (IIoT).

VPC de inspeção

Em uma arquitetura de AWS várias contas, uma VPC centralizada que gerencia as inspeções do tráfego de rede entre VPCs (na mesma ou em diferentes Regiões da AWS) a Internet e as redes locais. A Arquitetura de Referência de AWS Segurança recomenda configurar sua conta de rede com entrada, saída e inspeção VPCs para proteger a interface bidirecional entre seu aplicativo e a Internet em geral.

Internet das coisas (IoT)

A rede de objetos físicos conectados com sensores ou processadores incorporados que se comunicam com outros dispositivos e sistemas pela Internet ou por uma rede de comunicação local. Para obter mais informações, consulte O que é IoT?

interpretabilidade

Uma característica de um modelo de machine learning que descreve o grau em que um ser humano pode entender como as previsões do modelo dependem de suas entradas. Para obter mais informações, consulte Interpretabilidade do modelo de aprendizado de máquina com AWS.

IoT

Veja Internet das Coisas.

Biblioteca de informações de TI (ITIL)

Um conjunto de práticas recomendadas para fornecer serviços de TI e alinhar esses serviços a requisitos de negócios. A ITIL fornece a base para o ITSM.

Gerenciamento de serviços de TI (ITSM)

Atividades associadas a design, implementação, gerenciamento e suporte de serviços de TI para uma organização. Para obter informações sobre a integração de operações em nuvem com ferramentas de ITSM, consulte o guia de integração de operações.

ITIL

Veja biblioteca de informações de TI.

ITSM

Veja gerenciamento de serviços de TI.

L

controle de acesso baseado em etiqueta (LBAC)

Uma implementação do controle de acesso obrigatório (MAC) em que os usuários e os dados em si recebem explicitamente um valor de etiqueta de segurança. A interseção entre a etiqueta de segurança do usuário e a etiqueta de segurança dos dados determina quais linhas e colunas podem ser vistas pelo usuário.

zona de pouso

Uma landing zone é um AWS ambiente bem arquitetado, com várias contas, escalável e seguro. Um ponto a partir do qual suas organizações podem iniciar e implantar rapidamente workloads e aplicações com confiança em seu ambiente de segurança e infraestrutura. Para obter mais informações sobre zonas de pouso, consulte Configurar um ambiente da AWS com várias contas seguro e escalável.

grande modelo de linguagem (LLM)

Um modelo de IA de aprendizado profundo pré-treinado em uma grande quantidade de dados. Um LLM pode realizar várias tarefas, como responder a perguntas, resumir documentos, traduzir texto para outros idiomas e completar frases. Para obter mais informações, consulte O que são LLMs.

migração de grande porte

Uma migração de 300 servidores ou mais.

LBAC

Veja controle de acesso baseado em rótulo.

privilégio mínimo

A prática recomendada de segurança de conceder as permissões mínimas necessárias para executar uma tarefa. Para obter mais informações, consulte Aplicar permissões de privilégios mínimos na documentação do IAM.

mover sem alterações (lift-and-shift)

Veja 7 Rs.

sistema little-endian

Um sistema que armazena o byte menos significativo antes. Veja também endianness.

LLM

Veja grande modelo de linguagem.

ambientes inferiores

Veja ambiente.

M

machine learning (ML)

Um tipo de inteligência artificial que usa algoritmos e técnicas para reconhecimento e aprendizado de padrões. O ML analisa e aprende com dados gravados, por exemplo, dados da Internet das Coisas (IoT), para gerar um modelo estatístico baseado em padrões. Para obter mais informações, consulte Machine learning.

ramificação principal

Veja ramificação.

Malware

Software projetado para comprometer a segurança ou a privacidade do computador. O malware pode interromper os sistemas do computador, vazar informações sensíveis ou obter acesso não autorizado. Exemplos de malware incluem vírus, worms, ransomware, cavalos de Troia, spyware e keyloggers.

Serviços gerenciados

Serviços da AWS para o qual AWS opera a camada de infraestrutura, o sistema operacional e as plataformas, e você acessa os endpoints para armazenar e recuperar dados. O Amazon Simple Storage Service (Amazon S3) e o Amazon DynamoDB são exemplos de serviços gerenciados. Eles também são conhecidos como serviços abstraídos.

sistema de execução de manufatura (MES)

Um sistema de software para rastrear, monitorar, documentar e controlar processos de produção que convertem matérias-primas em produtos acabados no chão de fábrica.

MAP

Veja Programa de Aceleração da Migração.

mecanismo

Um processo completo em que você cria uma ferramenta, impulsiona a adoção da ferramenta e, em seguida, inspeciona os resultados para fazer ajustes. Um mecanismo é um ciclo que se reforça e se aprimora à medida que opera. Para obter mais informações, consulte Construindo mecanismos no AWS Well-Architected Framework.

conta de membro

Todos, Contas da AWS exceto a conta de gerenciamento, que fazem parte de uma organização em AWS Organizations. Uma conta só pode ser membro de uma organização de cada vez.

MES

Veja sistema de execução de manufatura.

Transporte de Telemetria de Enfileiramento de Mensagens (MQTT)

Um protocolo de comunicação leve machine-to-machine (M2M), baseado no padrão de publicação/assinatura, para dispositivos de IoT com recursos limitados.

microsserviço

Um serviço pequeno e independente que se comunica de forma bem definida APIs e normalmente é de propriedade de equipes pequenas e independentes. Por exemplo, um sistema de seguradora pode incluir microsserviços que mapeiam as capacidades comerciais, como vendas ou marketing, ou subdomínios, como compras, reclamações ou análises. Os benefícios dos microsserviços incluem agilidade, escalabilidade flexível, fácil implantação, código reutilizável e resiliência. Para obter mais informações, consulte Integração de microsserviços usando serviços sem AWS servidor.

arquitetura de microsserviços

Uma abordagem à criação de aplicações com componentes independentes que executam cada processo de aplicação como um microsserviço. Esses microsserviços se comunicam por meio

de uma interface bem definida usando leveza. APIs Cada microsserviço nessa arquitetura pode ser atualizado, implantado e escalado para atender à demanda por funções específicas de uma aplicação. Para obter mais informações, consulte Implementação de microsserviços em. AWS

Programa de Aceleração da Migração (MAP)

Um AWS programa que fornece suporte de consultoria, treinamento e serviços para ajudar as organizações a criar uma base operacional sólida para migrar para a nuvem e ajudar a compensar o custo inicial das migrações. O MAP inclui uma metodologia de migração para executar migrações legadas de forma metódica e um conjunto de ferramentas para automatizar e acelerar cenários comuns de migração.

migração em escala

O processo de mover a maior parte do portfólio de aplicações para a nuvem em ondas, com mais aplicações sendo movidas em um ritmo mais rápido a cada onda. Essa fase usa as práticas recomendadas e lições aprendidas nas fases anteriores para implementar uma fábrica de migração de equipes, ferramentas e processos para agilizar a migração de workloads por meio de automação e entrega ágeis. Esta é a terceira fase da estratégia de migração para a AWS.

fábrica de migração

Equipes multifuncionais que simplificam a migração de workloads por meio de abordagens automatizadas e ágeis. As equipes da fábrica de migração geralmente incluem operações, analistas e proprietários de negócios, engenheiros de migração, desenvolvedores e DevOps profissionais que trabalham em sprints. Entre 20 e 50% de um portfólio de aplicações corporativas consiste em padrões repetidos que podem ser otimizados por meio de uma abordagem de fábrica. Para obter mais informações, consulte discussão sobre fábricas de migração e o guia do Cloud Migration Factory neste conjunto de conteúdo.

metadados de migração

As informações sobre a aplicação e o servidor necessárias para concluir a migração. Cada padrão de migração exige um conjunto de metadados de migração diferente. Exemplos de metadados de migração incluem a sub-rede, o grupo de segurança e AWS a conta de destino.

padrão de migração

Uma tarefa de migração repetível que detalha a estratégia de migração, o destino da migração e a aplicação ou o serviço de migração usado. Exemplo: rehospede a migração para o Amazon EC2 AWS com o Application Migration Service.

Avaliação de Portfólio para Migração (MPA)

Uma ferramenta on-line que fornece informações para validar o caso de negócios para migrar para a Nuvem AWS. O MPA fornece avaliação detalhada do portfólio (dimensionamento correto do servidor, preços, comparações de TCO, análise de custos de migração), bem como planejamento de migração (análise e coleta de dados de aplicações, agrupamento de aplicações, priorização de migração e planejamento de ondas). A ferramenta MPA (requer login) está disponível gratuitamente para todos os AWS consultores e consultores parceiros da APN.

Avaliação de Preparação para Migração (MRA)

O processo de obter insights sobre o status de prontidão de uma organização para a nuvem, identificar pontos fortes e fracos e criar um plano de ação para fechar as lacunas identificadas, usando o CAF. AWS Para mais informações, consulte o guia de preparação para migração. A MRA é a primeira fase da estratégia de migração para a AWS.

estratégia de migração

A abordagem usada para migrar uma workload para a Nuvem AWS. Para obter mais informações, veja a entrada 7 Rs neste glossário e consulte Mobilize sua organização para acelerar migrações em grande escala.

ML

Veja machine learning.

modernização

Transformar uma aplicação desatualizada (herdada ou monolítica) e sua infraestrutura em um sistema ágil, elástico e altamente disponível na nuvem para reduzir custos, ganhar eficiência e aproveitar as inovações. Para obter mais informações, consulte Strategy for modernizing applications in the Nuvem AWS.

avaliação de preparação para modernização

Uma avaliação que ajuda a determinar a preparação para modernização das aplicações de uma organização. Ela identifica benefícios, riscos e dependências e determina o quão bem a organização pode acomodar o estado futuro dessas aplicações. O resultado da avaliação é um esquema da arquitetura de destino, um roteiro que detalha as fases de desenvolvimento e os marcos do processo de modernização e um plano de ação para abordar as lacunas identificadas. Para obter mais informações, consulte Evaluating modernization readiness for applications in the Nuvem AWS.

aplicações monolíticas (monólitos)

Aplicações que são executadas como um único serviço com processos fortemente acoplados. As aplicações monolíticas apresentam várias desvantagens. Se um recurso da aplicação apresentar um aumento na demanda, toda a arquitetura deverá ser escalada. Adicionar ou melhorar os recursos de uma aplicação monolítica também se torna mais complexo quando a base de código cresce. Para resolver esses problemas, é possível criar uma arquitetura de microsserviços. Para obter mais informações, consulte Decompor monólitos em microsserviços.

MPA

Veja Avaliação do Portfólio para Migração.

MQTT

Veja Transporte de Telemetria de Enfileiramento de Mensagens.

classificação multiclasse

Um processo que ajuda a gerar previsões para várias classes (prevendo um ou mais de dois resultados). Por exemplo, um modelo de ML pode perguntar “Este produto é um livro, um carro ou um telefone?” ou “Qual categoria de produtos é mais interessante para este cliente?”

infraestrutura mutável

Um modelo que atualiza e modifica a infraestrutura existente para workloads de produção. Para melhorar a consistência, confiabilidade e previsibilidade, o AWS Well-Architected Framework recomenda o uso de infraestrutura imutável como uma prática recomendada.

O

OAC

Veja controle de acesso de origem.

OAI

Veja identidade de acesso de origem.

OCM

Veja gerenciamento de alterações organizacionais.

migração offline

Um método de migração no qual a workload de origem é desativada durante o processo de migração. Esse método envolve tempo de inatividade prolongado e geralmente é usado para workloads pequenas e não críticas.

OI

Veja integração de operações.

Ola

Veja acordo de nível operacional.

migração online

Um método de migração no qual a workload de origem é copiada para o sistema de destino sem ser colocada offline. As aplicações conectadas à workload podem continuar funcionando durante a migração. Esse método envolve um tempo de inatividade nulo ou mínimo e normalmente é usado para workloads essenciais para a produção.

OPC-UA

Veja Open Process Communications - Unified Architecture.

Open Process Communications - Unified Architecture (OPC-UA)

Um protocolo de comunicação machine-to-machine (M2M) para automação industrial. O OPC-UA fornece um padrão de interoperabilidade com esquemas de criptografia, autenticação e autorização de dados.

acordo de nível operacional (OLA)

Um acordo que esclarece o que os grupos funcionais de TI prometem oferecer uns aos outros para apoiar um acordo de serviço (SLA).

análise de prontidão operacional (ORR)

Uma lista de verificação de perguntas e práticas recomendadas associadas que ajudam você a entender, avaliar, prevenir ou reduzir o escopo de incidentes e possíveis falhas. Para obter mais informações, consulte Operational Readiness Reviews (ORR) no AWS Well-Architected Framework.

tecnologia operacional (TO)

Sistemas de hardware e software que trabalham com o ambiente físico para controlar operações, equipamentos e infraestrutura industriais. Na manufatura, a integração dos sistemas de

tecnologia da informação (TI) e tecnologia operacional (TO) é o foco principal das transformações da Indústria 4.0.

integração de operações (OI)

O processo de modernização das operações na nuvem, que envolve planejamento de preparação, automação e integração. Para obter mais informações, consulte o guia de integração de operações.

trilha organizacional

Uma trilha criada por ela AWS CloudTrail registra todos os eventos de todos Contas da AWS em uma organização em AWS Organizations. Essa trilha é criada em cada Conta da AWS que faz parte da organização e monitora a atividade em cada conta. Para obter mais informações, consulte Criação de uma trilha para uma organização na CloudTrail documentação.

gerenciamento de alterações organizacionais (OCM)

Uma estrutura para gerenciar grandes transformações de negócios disruptivas de uma perspectiva de pessoas, cultura e liderança. O OCM ajuda as organizações a se prepararem e fazerem a transição para novos sistemas e estratégias, acelerando a adoção de alterações, abordando questões de transição e promovendo mudanças culturais e organizacionais. Na estratégia de AWS migração, essa estrutura é chamada de aceleração de pessoas, devido à velocidade de mudança exigida nos projetos de adoção da nuvem. Para obter mais informações, consulte o guia do OCM.

controle de acesso de origem (OAC)

Em CloudFront, uma opção aprimorada para restringir o acesso para proteger seu conteúdo do Amazon Simple Storage Service (Amazon S3). O OAC oferece suporte a todos os buckets S3 Regiões da AWS, criptografia do lado do servidor com AWS KMS (SSE-KMS) e solicitações dinâmicas ao bucket S3. `PUT` `DELETE`

Identidade do acesso de origem (OAI)

Em CloudFront, uma opção para restringir o acesso para proteger seu conteúdo do Amazon S3. Quando você usa o OAI, CloudFront cria um principal com o qual o Amazon S3 pode se autenticar. Os diretores autenticados podem acessar o conteúdo em um bucket do S3 somente por meio de uma distribuição específica. CloudFront Veja também OAC, que fornece um controle de acesso mais granular e aprimorado.

ORR

Veja análise de prontidão operacional.

OT

Veja tecnologia operacional.

VPC de saída (egresso)

Em uma arquitetura de AWS várias contas, uma VPC que gerencia conexões de rede que são iniciadas de dentro de um aplicativo. A Arquitetura de Referência de AWS Segurança recomenda configurar sua conta de rede com entrada, saída e inspeção VPCs para proteger a interface bidirecional entre seu aplicativo e a Internet em geral.

P

limite de permissões

Uma política de gerenciamento do IAM anexada a entidades principais do IAM para definir as permissões máximas que o usuário ou perfil podem ter. Para obter mais informações, consulte Limites de permissões na documentação do IAM.

Informações de identificação pessoal (PII)

Informações que, quando visualizadas diretamente ou combinadas com outros dados relacionados, podem ser usadas para inferir razoavelmente a identidade de um indivíduo. Exemplos de PII incluem nomes, endereços e informações de contato.

PII

Veja informações de identificação pessoal.

manual

Um conjunto de etapas predefinidas que capturam o trabalho associado às migrações, como a entrega das principais funções operacionais na nuvem. Um manual pode assumir a forma de scripts, runbooks automatizados ou um resumo dos processos ou etapas necessários para operar seu ambiente modernizado.

PLC

Veja controlador lógico programável.

PLM

Veja gerenciamento do ciclo de vida do produto.

política

Um objeto que pode definir permissões (veja política baseada em identidade), especificar condições de acesso (veja políta baseada em recurso) ou definir as permissões máximas para todas as contas em uma organização no AWS Organizations (veja política de controle de serviços).

persistência poliglota

Escolher de forma independente a tecnologia de armazenamento de dados de um microsserviço com base em padrões de acesso a dados e outros requisitos. Se seus microsserviços tiverem a mesma tecnologia de armazenamento de dados, eles poderão enfrentar desafios de implementação ou apresentar baixa performance. Os microsserviços serão implementados com mais facilidade e alcançarão performance e escalabilidade melhores se usarem o armazenamento de dados mais bem adaptado às suas necessidades.

avaliação do portfólio

Um processo de descobrir, analisar e priorizar o portfólio de aplicações para planejar a migração. Para obter mais informações, consulte Avaliar a preparação para a migração.

predicado

Uma condição de consulta que retorna `true` ou `false`, normalmente localizada em uma cláusula `WHERE`.

pushdown de predicados

Uma técnica de otimização de consultas de banco de dados que filtra os dados na consulta antes da transferência. Isso reduz a quantidade de dados que devem ser recuperados e processados do banco de dados relacional e melhora a performance das consultas.

controle preventivo

Um controle de segurança projetado para evitar que um evento ocorra. Esses controles são a primeira linha de defesa para ajudar a evitar acesso não autorizado ou alterações indesejadas em sua rede. Para obter mais informações, consulte Controles preventivos em Como implementar controles de segurança na AWS.

principal (entidade principal)

Uma entidade AWS que pode realizar ações e acessar recursos. Essa entidade geralmente é um usuário raiz para um Conta da AWS, uma função do IAM ou um usuário. Para obter mais

informações, consulte Entidade principal em Termos e conceitos de perfis na documentação do IAM.

Privacidade por design

Uma abordagem em engenharia de sistemas que leva em consideração a privacidade em todo o processo de desenvolvimento.

zonas hospedadas privadas

Um contêiner que contém informações sobre como você deseja que o Amazon Route 53 responda às consultas de DNS para um domínio e seus subdomínios em um ou mais. VPCs Para obter mais informações, consulte Como trabalhar com zonas hospedadas privadas na documentação do Route 53.

controle proativo

Um controle de segurança desenvolvido para evitar a implantação de recursos não conformes. Esses controles verificam os recursos antes de serem provisionados. Se o recurso não estiver em conformidade com o controle, ele não será provisionado. Para obter mais informações, consulte o guia de referência de controles na AWS Control Tower documentação e consulte Controles proativos em Implementação de controles de segurança em AWS.

gerenciamento do ciclo de vida do produto (PLM)

O gerenciamento de dados e processos de um produto em todo o seu ciclo de vida, desde a concepção, o desenvolvimento e o lançamento, passando pelo crescimento e maturidade, até o declínio e a remoção.

ambiente de produção

Veja ambiente.

controlador lógico programável (PLC)

Na manufatura, um computador altamente confiável e adaptável que monitora as máquinas e automatiza os processos de fabricação.

encadeamento de prompts

Uso da saída de um prompt do LLM como entrada para o próximo prompt para gerar respostas melhores. Essa técnica é usada para dividir uma tarefa complexa em subtarefas, ou para refinar ou expandir iterativamente uma resposta preliminar. Isso ajuda a melhorar a precisão e a relevância das respostas de um modelo e permite resultados mais granulares e personalizados.

pseudonimização

O processo de substituir identificadores pessoais em um conjunto de dados por valores de espaço reservado. A pseudonimização pode ajudar a proteger a privacidade pessoal. Os dados pseudonimizados ainda são considerados dados pessoais.

publish/subscribe (pub/sub)

Um padrão que permite comunicações assíncronas entre microsserviços para melhorar a escalabilidade e a capacidade de resposta. Por exemplo, em um MES baseado em microsserviços, um microsserviço pode publicar mensagens de eventos em um canal em que outros microsserviços possam assinar. O sistema pode adicionar novos microsserviços sem alterar o serviço de publicação.

Q

plano de consulta

Uma série de etapas, como instruções, usadas para acessar os dados em um sistema de banco de dados relacional SQL.

regressão de planos de consultas

Quando um otimizador de serviço de banco de dados escolhe um plano menos adequado do que escolhia antes de uma determinada alteração no ambiente de banco de dados ocorrer. Isso pode ser causado por alterações em estatísticas, restrições, configurações do ambiente, associações de parâmetros de consulta e atualizações do mecanismo de banco de dados.

R

Matriz RACI

Veja responsável, aprovador, consultado, informado (RACI).

RAG

Veja geração aumentada via recuperação.

ransomware

Um software mal-intencionado desenvolvido para bloquear o acesso a um sistema ou dados de computador até que um pagamento seja feito.

Matriz RASCI

Veja responsável, aprovador, consultado, informado (RACI).

RCAC

Veja controle de acesso por linha e coluna.

réplica de leitura

Uma cópia de um banco de dados usada somente para leitura. É possível encaminhar consultas para a réplica de leitura e reduzir a carga no banco de dados principal.

Redefinir arquitetura

Veja 7 Rs.

objetivo de ponto de recuperação (RPO).

O máximo período de tempo aceitável desde o último ponto de recuperação de dados. Isso determina o que é considerado uma perda aceitável de dados entre o último ponto de recuperação e a interrupção do serviço.

objetivo de tempo de recuperação (RTO)

O máximo atraso aceitável entre a interrupção e a restauração do serviço.

refatorar

Veja 7 Rs.

Região

Uma coleção de AWS recursos em uma área geográfica. Cada um Região da AWS é isolado e independente dos outros para fornecer tolerância a falhas, estabilidade e resiliência. Para obter informações, consulte Specify which Regiões da AWS your account can use.

regressão

Uma técnica de ML que prevê um valor numérico. Por exemplo, para resolver o problema de “Por qual preço esta casa será vendida?” um modelo de ML pode usar um modelo de regressão linear para prever o preço de venda de uma casa com base em fatos conhecidos sobre a casa (por exemplo, a metragem quadrada).

redefinir a hospedagem

Veja 7 Rs.

versão

Em um processo de implantação, o ato de promover mudanças em um ambiente de produção.

realocar

Veja 7 Rs.

redefinir a plataforma

Veja 7 Rs.

recomprar

Veja 7 Rs.

resiliência

A capacidade de uma aplicação de resistir ou se recuperar de interrupções. Alta disponibilidade e recuperação de desastres são considerações comuns ao planejar a resiliência na Nuvem AWS. Para obter mais informações, consulte Nuvem AWS Resilience.

política baseada em recurso

Uma política associada a um recurso, como um bucket do Amazon S3, um endpoint ou uma chave de criptografia. Esse tipo de política especifica quais entidades principais têm acesso permitido, ações válidas e quaisquer outras condições que devem ser atendidas.

matriz responsável, accountable, consultada, informada (RACI)

Uma matriz que define as funções e responsabilidades de todas as partes envolvidas nas atividades de migração e nas operações de nuvem. O nome da matriz é derivado dos tipos de responsabilidade definidos na matriz: responsável (R), responsabilizável (A), consultado (C) e informado (I). O tipo de suporte (S) é opcional. Se você incluir suporte, a matriz será chamada de matriz RASCI e, se excluir, será chamada de matriz RACI.

controle responsivo

Um controle de segurança desenvolvido para conduzir a remediação de eventos adversos ou desvios em relação à linha de base de segurança. Para obter mais informações, consulte Controles responsivos em Como implementar controles de segurança na AWS.

reter

Veja 7 Rs.

Retirada

Veja 7 Rs.

Geração Aumentada de Recuperação (RAG)

Uma tecnologia de IA generativa em que um LLM faz referência a uma fonte de dados autorizada que está fora de suas fontes de dados de treinamento antes de gerar uma resposta. Por exemplo, um modelo RAG pode realizar uma pesquisa semântica na base de conhecimento ou nos dados personalizados de uma organização. Para obter mais informações, consulte O que é RAG (geração aumentada via recuperação)?.

alternância

O processo de atualizar periodicamente um segredo para dificultar o acesso de um invasor às credenciais.

controle de acesso por linha e coluna (RCAC)

O uso de expressões SQL básicas e flexíveis que tenham regras de acesso definidas. O RCAC consiste em permissões de linha e máscaras de coluna.

RPO

Veja objetivo de ponto de recuperação.

RTO

Veja objetivo de tempo de recuperação.

runbook

Um conjunto de procedimentos manuais ou automatizados necessários para realizar uma tarefa específica. Eles são normalmente criados para agilizar operações ou procedimentos repetitivos com altas taxas de erro.

S

SAML 2.0

Um padrão aberto que muitos provedores de identidade (IdPs) usam. Esse recurso permite o login único federado (SSO), para que os usuários possam fazer login Console de gerenciamento da AWS ou chamar as operações da AWS API sem que você precise criar um usuário no IAM

para todos em sua organização. Para obter mais informações sobre a federação baseada em SAML 2.0, consulte Sobre a federação baseada em SAML 2.0 na documentação do IAM.

SCADA

Veja controle de supervisão e aquisição de dados.

SCP

Veja política de controle de serviço.

secret

Em AWS Secrets Manager, informações confidenciais ou restritas, como uma senha ou credenciais de usuário, que você armazena de forma criptografada. Consiste no valor secreto e em seus metadados. O valor secreto pode ser binário, uma única string ou várias strings. Para obter mais informações, consulte What's in a Secrets Manager secret? na documentação do Secrets Manager.

segurança desde a concepção

Uma abordagem em engenharia de sistemas que leva em consideração a segurança em todo o processo de desenvolvimento.

controle de segurança

Uma barreira de proteção técnica ou administrativa que impede, detecta ou reduz a capacidade de uma ameaça explorar uma vulnerabilidade de segurança. Existem quatro tipos primários de controles de segurança: preventivos, detectivos, responsivos e proativos.

hardening da segurança

O processo de reduzir a superfície de ataque para torná-la mais resistente a ataques. Isso pode incluir ações como remover recursos que não são mais necessários, implementar a prática recomendada de segurança de conceder privilégios mínimos ou desativar recursos desnecessários em arquivos de configuração.

sistema de gerenciamento de eventos e informações de segurança (SIEM)

Ferramentas e serviços que combinam sistemas de gerenciamento de informações de segurança (SIM) e gerenciamento de eventos de segurança (SEM). Um sistema SIEM coleta, monitora e analisa dados de servidores, redes, dispositivos e outras fontes para detectar ameaças e violações de segurança e gerar alertas.

automação de resposta de segurança

Uma ação predefinida e programada projetada para responder ou remediar automaticamente um evento de segurança. Essas automações servem como controles de segurança responsivos ou detectivos que ajudam você a implementar as melhores práticas AWS de segurança. Exemplos de ações de resposta automatizada incluem a modificação de um grupo de segurança da VPC, a aplicação de patches em uma instância do Amazon EC2 ou a alternância de credenciais.

Criptografia do lado do servidor

Criptografia dos dados em seu destino, por AWS service (Serviço da AWS) quem os recebe.

política de controle de serviços (SCP)

Uma política que fornece controle centralizado sobre as permissões de todas as contas em uma organização em AWS Organizations. SCPs defina barreiras ou estabeleça limites nas ações que um administrador pode delegar a usuários ou funções. Você pode usar SCPs como listas de permissão ou listas de negação para especificar quais serviços ou ações são permitidos ou proibidos. Para obter mais informações, consulte Políticas de controle de serviço na AWS Organizations documentação.

service endpoint (endpoint de serviço)

O URL do ponto de entrada para um AWS service (Serviço da AWS). Você pode usar o endpoint para se conectar programaticamente ao serviço de destino. Para obter mais informações, consulte Endpoints do AWS service (Serviço da AWS) na Referência geral da AWS.

acordo de serviço (SLA)

Um acordo que esclarece o que uma equipe de TI promete fornecer aos clientes, como tempo de atividade e performance do serviço.

indicador de nível de serviço (SLI)

Uma avaliação de um aspecto de performance de um serviço, como taxa de erro, disponibilidade ou throughput.

objetivo de nível de serviço (SLO)

Uma métrica alvo que representa a integridade de um serviço, conforme avaliado por um indicador de nível de serviço.

modelo de responsabilidade compartilhada

Um modelo que descreve a responsabilidade com a qual você compartilha AWS pela segurança e conformidade na nuvem. AWS é responsável pela segurança da nuvem, enquanto você é responsável pela segurança na nuvem. Para obter mais informações, consulte o Modelo de responsabilidade compartilhada.

SIEM

Veja sistema de gerenciamento de eventos e informações de segurança.

ponto único de falha (SPOF)

Uma falha em um único componente crítico de uma aplicação que pode interromper o sistema.

SLA

Veja cacordo de serviço.

SLI

Veja indicador de nível de serviço.

SLO

Veja objetivo de nível de serviço.

split-and-seed modelo

Um padrão para escalar e acelerar projetos de modernização. À medida que novos recursos e lançamentos de produtos são definidos, a equipe principal se divide para criar novas equipes de produtos. Isso ajuda a escalar os recursos e os serviços da sua organização, melhora a produtividade do desenvolvedor e possibilita inovações rápidas. Para obter mais informações, consulte Phased approach to modernizing applications in the Nuvem AWS.

SPOF

Veja ponto único de falha.

esquema em estrela

Uma estrutura organizacional de banco de dados que usa uma grande tabela de fatos para armazenar dados transacionais ou medidos e usa uma ou mais tabelas dimensionais menores para armazenar atributos de dados. Essa estrutura foi projetada para ser usada em um data warehouse ou para fins de inteligência comercial.

padrão strangler fig

Uma abordagem à modernização de sistemas monolíticos que consiste em reescrever e substituir incrementalmente a funcionalidade do sistema até que o sistema herdado possa ser desativado. Esse padrão usa a analogia de uma videira que cresce e se torna uma árvore estabelecida e, eventualmente, supera e substitui sua hospedeira. O padrão foi apresentado por Martin Fowler como forma de gerenciar riscos ao reescrever sistemas monolíticos. Para ver um exemplo de como aplicar esse padrão, consulte Modernizar incrementalmente os serviços Web herdados do Microsoft ASP.NET (ASMX) usando contêineres e o Amazon API Gateway.

sub-rede

Um intervalo de endereços IP na VPC. Cada sub-rede fica alocada em uma única zona de disponibilidade.

controle supervisório e aquisição de dados (SCADA)

Na manufatura, um sistema que usa hardware e software para monitorar ativos físicos e operações de produção.

symmetric encryption (criptografia simétrica)

Um algoritmo de criptografia que usa a mesma chave para criptografar e descriptografar dados.

testes sintéticos

Testar um sistema de forma que simule as interações do usuário para detectar possíveis problemas ou monitorar a performance. Você pode usar o Amazon CloudWatch Synthetics para criar esses testes.

prompt do sistema

Uma técnica para fornecer contexto, instruções ou orientações a um LLM a fim de direcionar seu comportamento. Os prompts do sistema ajudam a definir o contexto e a estabelecer regras para interações com os usuários.

T

tags

Pares de valores-chave que atuam como metadados para organizar seus recursos. AWS As tags podem ajudar você a gerenciar, identificar, organizar, pesquisar e filtrar recursos da . Para obter mais informações, consulte Marcar seus recursos do AWS.

variável-alvo

O valor que você está tentando prever no ML supervisionado. Ela também é conhecida como variável de resultado. Por exemplo, em uma configuração de fabricação, a variável-alvo pode ser um defeito do produto.

lista de tarefas

Uma ferramenta usada para monitorar o progresso por meio de um runbook. Uma lista de tarefas contém uma visão geral do runbook e uma lista de tarefas gerais a serem concluídas. Para cada tarefa geral, ela inclui o tempo estimado necessário, o proprietário e o progresso.

ambiente de teste

Veja ambiente.

treinamento

O processo de fornecer dados para que seu modelo de ML aprenda. Os dados de treinamento devem conter a resposta correta. O algoritmo de aprendizado descobre padrões nos dados de treinamento que mapeiam os atributos dos dados de entrada no destino (a resposta que você deseja prever). Ele gera um modelo de ML que captura esses padrões. Você pode usar o modelo de ML para obter previsões de novos dados cujo destino você não conhece.

gateway de trânsito

Um hub de trânsito de rede que você pode usar para interconectar sua rede com VPCs a rede local. Para obter mais informações, consulte O que é um gateway de trânsito na AWS Transit Gateway documentação.

fluxo de trabalho baseado em troncos

Uma abordagem na qual os desenvolvedores criam e testam recursos localmente em uma ramificação de recursos e, em seguida, mesclam essas alterações na ramificação principal. A ramificação principal é então criada para os ambientes de desenvolvimento, pré-produção e produção, sequencialmente.

Acesso confiável

Conceder permissões a um serviço que você especifica para realizar tarefas em sua organização AWS Organizations e em suas contas em seu nome. O serviço confiável cria um perfil vinculado ao serviço em cada conta, quando esse perfil é necessário, para realizar tarefas de gerenciamento para você. Para obter mais informações, consulte Usando AWS Organizations com outros AWS serviços na AWS Organizations documentação.

tuning (ajustar)

Alterar aspectos do processo de treinamento para melhorar a precisão do modelo de ML. Por exemplo, você pode treinar o modelo de ML gerando um conjunto de rótulos, adicionando rótulos e repetindo essas etapas várias vezes em configurações diferentes para otimizar o modelo.

equipe de duas pizzas

Uma pequena DevOps equipe que você pode alimentar com duas pizzas. Uma equipe de duas pizzas garante a melhor oportunidade possível de colaboração no desenvolvimento de software.

U

incerteza

Um conceito que se refere a informações imprecisas, incompletas ou desconhecidas que podem minar a confiabilidade dos modelos preditivos de ML. Há dois tipos de incertezas: a incerteza epistêmica é causada por dados limitados e incompletos, enquanto a incerteza aleatória é causada pelo ruído e pela aleatoriedade inerentes aos dados.

tarefas indiferenciadas

Também conhecido como trabalho pesado, trabalho necessário para criar e operar um aplicativo, mas que não fornece valor direto ao usuário final nem oferece vantagem competitiva. Exemplos de tarefas indiferenciadas incluem aquisição, manutenção e planejamento de capacidade.

ambientes superiores

Veja ambiente.

V

aspiração

Uma operação de manutenção de banco de dados que envolve limpeza após atualizações incrementais para recuperar armazenamento e melhorar a performance.

controle de versões

Processos e ferramentas que rastreiam mudanças, como alterações no código-fonte em um repositório.

emparelhamento da VPC

Uma conexão entre duas VPCs que permite rotear o tráfego usando endereços IP privados. Para ter mais informações, consulte O que é emparelhamento de VPC? na documentação da Amazon VPC.

Vulnerabilidade

Uma falha de software ou hardware que compromete a segurança do sistema.

W

cache quente

Um cache de buffer que contém dados atuais e relevantes que são acessados com frequência. A instância do banco de dados pode ler do cache do buffer, o que é mais rápido do que ler da memória principal ou do disco.

dados mornos

Dados acessados raramente. Ao consultar esse tipo de dados, consultas moderadamente lentas geralmente são aceitáveis.

função de janela

Uma função SQL que executa um cálculo em um grupo de linhas que se relacionam de alguma forma com o registro atual. As funções de janela são úteis para processar tarefas, como calcular uma média móvel ou acessar o valor das linhas com base na posição relativa da linha atual.

workload

Uma coleção de códigos e recursos que geram valor empresarial, como uma aplicação voltada para o cliente ou um processo de backend.

workstreams

Grupos funcionais em um projeto de migração que são responsáveis por um conjunto específico de tarefas. Cada workstream é independente, mas oferece suporte aos outros workstreams do projeto. Por exemplo, o workstream de portfólio é responsável por priorizar aplicações, planejar ondas e coletar metadados de migração. O workstream de portfólio entrega esses ativos ao workstream de migração, que então migra os servidores e as aplicações.

WORM

Veja gravação única e várias leituras.

WQF

Veja AWS Workload Qualification Framework.

gravação única e várias leituras (WORM)

Um modelo de armazenamento que grava dados uma única vez e evita que os dados sejam excluídos ou modificados. Os usuários autorizados podem ler os dados quantas vezes forem necessárias, mas não podem alterá-los. Essa infraestrutura de armazenamento de dados é considerada imutável.

Z

exploração de dia zero

Um ataque, normalmente malware, que tira proveito de uma vulnerabilidade zero-day.

vulnerabilidade de dia zero

Uma falha ou vulnerabilidade não mitigada em um sistema de produção. Os agentes de ameaças podem usar esse tipo de vulnerabilidade para atacar o sistema. Os desenvolvedores frequentemente ficam cientes da vulnerabilidade como resultado do ataque.

prompt zero shot

Fornecer a um LLM instruções para realizar uma tarefa, mas sem exemplos (shots) que possam ajudar a orientá-lo. O LLM deve usar seu conhecimento pré-treinado para lidar com a tarefa. A eficácia dos prompts zero-shot depende da complexidade da tarefa e da qualidade do prompt. Veja também prompts few-shot.

aplicação zumbi

Uma aplicação que tem um uso médio de CPU e memória inferior a 5%. Em um projeto de migração, é comum retirar essas aplicações.

As traduções são geradas por tradução automática. Em caso de conflito entre o conteúdo da tradução e da versão original em inglês, a versão em inglês prevalecerá.