Migração para o Amazon Service OpenSearch

AWS Recomendações

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

AWS Recomendações: Migração para o Amazon Service OpenSearch

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

Introdução 1
Visão geral 1
Benefícios do OpenSearch Service 3
Mais fácil de implantar e gerenciar 3
Econômica 3
Mais escalável e confiável 4
Seguro e compatível 4
Jornada de migração 5
Planejamento 6
Dimensionamento 6
Armazenamento 7
Número de nós e tipos de instância 8
Determinação da estratégia de indexação e contagem de fragmentos 9
Utilização da CPU 9
Tipos de instância 10
Funcionalidade 11
Funcionalidade atual da solução 11
Funcionalidade OpenSearch do Amazon Service 11
Plug-ins empacotados 12
Plugins personalizados 12
Dependências de versão 13
Seleção da versão do mecanismo 13
Atualizando para a versão mais recente do OpenSearch Serviço 13
Estratégia de atualização da versão 13
Verificações pré-atualização 14
KPIs e continuidade dos negócios 14
Performance operacional 16
Performance do processo 16
Transição suave para novos serviços 17
Métricas financeiras 17
Operações e segurança 18
Runbooks e novos processos 18
Sistema de geração de tíquetes e suporte 18
Segurança 19

Treinamento 20
Opções de treinamento 20
Fluxo de dados 21
Ingestão de dados 21
Retenção de dados 23
Abordagens da migração de dados 23
Frameworks de implantação 25
Prova de conceito 27
Definição de critérios de entrada e saída 27
Garantia de financiamento 27
Automatização 28
Testes completos 28
Etapas da PoC 29
Simulação de falhas 30
Implantação 31
Migrações de dados 32
Criar com base em um snapshot 32
Considerações sobre snapshots 33
Criar com base na fonte 34
Reindexação remota 35
Usar o Logstash 36
substituição 37
Sincronização de dados 37
Trocar ou substituir 41
Excelência operacional 42
Conclusão 43
Recursos 44
Colaboradores 45
Histórico do documentos 46
..... xlvii

Migração para o Amazon OpenSearch Service

Amazon Web Services (colaboradores)

Agosto de 2023 (histórico do documento)

Para muitos clientes, migrar implantações autogerenciadas do Elasticsearch ou do OpenSearch para o Amazon OpenSearch Service é um desafio. Os desafios comuns são a avaliação da workload, o planejamento da capacidade e a otimização da arquitetura. Também há perguntas sobre como atender a todos os requisitos de aplicações de analytics operacional de data centers on-premises na Nuvem Amazon Web Services (AWS). Este guia aborda a jornada geral de uma migração para o Amazon OpenSearch Service e fornece as práticas recomendadas que os especialistas da AWS reuniram ao longo do tempo. As instruções passo a passo podem ajudar você a conduzir suas migrações com uma abordagem eficaz e eficiente. Este guia aborda principalmente os domínios provisionados do Amazon OpenSearch Service e não as coletas do Amazon OpenSearch Sem Servidor.

Visão geral

O OpenSearch é um pacote de analytics e pesquisa distribuído e de código aberto usado para um amplo conjunto de casos de uso de analytics operacional, como monitoramento de aplicações em tempo real, analytics de logs, observabilidade de dados e pesquisa de catálogos de aplicações e produtos. O OpenSearch fornece uma resposta de pesquisa de baixa latência. Ele também oferece acesso rápido a grandes volumes de dados com uma ferramenta integrada de visualização de dados de código aberto chamada OpenSearch Dashboards.

O Amazon OpenSearch Service oferece suporte ao analytics interativo de logs, monitoramento de aplicações em tempo real, pesquisa em sites e muito mais. O Amazon OpenSearch Service oferece as versões mais recentes do OpenSearch e é compatível com 19 versões do Elasticsearch (versões de 1.5 a 7.10). Ele também fornece recursos de visualização baseados no OpenSearch Dashboards e no Kibana (versões de 1.5 a 7.10). Atualmente, o Amazon OpenSearch Service tem dezenas de milhares de clientes ativos com centenas de milhares de clusters processando centenas de trilhões de solicitações por mês.

Gerenciar clusters do OpenSearch ou do Elasticsearch on-premises ou na infraestrutura de nuvem é um trabalho altamente complexo, oneroso e tedioso. Para executar esses clusters, você deve provisionar e manter a infraestrutura. Os esforços incluem os seguintes:

- Aquisição e configuração de hardware
- Instalação de software
- Configuração, aplicação de patches e upgrades
- Considerações sobre confiabilidade e disponibilidade
- Considerações sobre performance e escalabilidade
- Considerações de segurança e conformidade, como isolamento de rede, controle de acesso refinado, criptografias e programas de conformidade, como os seguintes:
  - Federal Risk and Authorization Management Program (FedRAMP)
  - Regulamento Geral sobre a Proteção de Dados (General Data Protection Regulation, ou GDPR)
  - Health Insurance Portability and Accountability Act (HIPAA)
  - International Organization for Standardization (ISO)
  - Padrão de segurança de dados do setor de cartão de pagamento (PCI DSS – Payment Card Industry Data Security Standard)
  - Controles do Sistema e da Organização (CSO).

Em comparação, o Amazon OpenSearch Service gerencia essas tarefas para você. Neste guia, você aprenderá as abordagens e as práticas recomendadas para migrar o Elasticsearch ou o OpenSearch on-premises ou autogerenciados para o Amazon OpenSearch Service totalmente gerenciado.

Benefícios da migração para o Amazon OpenSearch Service

O Amazon OpenSearch Service ajuda nas tarefas de implantação e gerenciamento contínuo. É econômico e fornece escalabilidade, o que melhora a confiabilidade. Ele também oferece segurança e ajuda a atender às suas necessidades de conformidade.

Mais fácil de implantar e gerenciar

É mais fácil implantar um cluster do OpenSearch usando o Amazon OpenSearch Service do que implantar um cluster sozinho. O Amazon OpenSearch Service ajuda a gerenciar tarefas como provisionamento de hardware, aplicação de patches e instalação de software, recuperação de falhas, backups e monitoramento. Você não precisa ter uma equipe dedicada de especialistas em OpenSearch para gerenciar seus clusters.

Um cluster do OpenSearch no Amazon OpenSearch Service também é chamado de domínio. O Amazon OpenSearch Service fornece monitoramento da integridade do domínio por meio do serviço do Amazon CloudWatch. Você pode configurar alertas para ser notificado sobre quaisquer alterações na integridade de seus domínios. O AWS Support fornece suporte técnico individual de engenheiros experientes. Clientes com desafios operacionais ou dúvidas técnicas podem entrar em contato com o AWS Support e receber suporte personalizado com tempos de resposta confiáveis.

Econômica

O Amazon OpenSearch Service é econômico. Ele fornece uma gama completa de recursos avançados sem cobrar taxas adicionais de licenciamento. Você pode usar recursos como segurança de nível corporativo, alertas em tempo real, pesquisa entre clusters, gerenciamento automatizado de índices e detecção de anomalias sem custo adicional. Não há cobranças pelas transferências de dados entre as zonas de disponibilidade, e os snapshots por hora são fornecidos sem custo adicional.

Com o UltraWarm, você pode executar analytics interativo em até três petabytes de dados de log e, ao mesmo tempo, reduzir o custo por GB em até 90% em comparação com a camada de armazenamento quente. Além disso, o Amazon OpenSearch Service disponibiliza instâncias reservadas que oferecem descontos significativos em comparação com as instâncias sob demanda padrão. Para obter mais informações, consulte Orientado pelos custos.

Mais escalável e confiável

Com o Amazon OpenSearch Service, você pode armazenar petabytes de dados em um único domínio. Você pode consultar dados em vários domínios e analisar todos os seus dados em uma única interface do OpenSearch Dashboards. O Amazon OpenSearch Service foi projetado para ser altamente confiável, usando implantações de várias zonas de disponibilidade (multi-AZ) para que você possa replicar dados entre até três zonas de disponibilidade na mesma região da AWS. Não há tempo de inatividade quando você faz atualizações e upgrades de software ou escala seu ambiente.

Com o recurso multi-AZ com modo de espera, os domínios do OpenSearch Service são resilientes a possíveis falhas de infraestrutura, como uma falha de nó ou zona de disponibilidade. Isso permite 99,99% de disponibilidade e performance consistente para workloads essenciais para os negócios. Com o multi-AZ com modo de espera, os clusters são resilientes a falhas de infraestrutura, como falhas de hardware ou de rede. Essa opção oferece maior confiabilidade e o benefício adicional de simplificar a configuração e o gerenciamento do cluster, aplicando as práticas recomendadas e reduzindo a complexidade.

Seguro e compatível

O Amazon OpenSearch Service cuida de todos os patches de segurança. Ele também oferece isolamento de rede por meio de uma nuvem privada virtual (VPC), controle de acesso refinado e suporte multilocatário ao OpenSearch Dashboards. Você pode criptografar seus dados em repouso e em trânsito. Para ajudar você a atender aos requisitos regulamentares e específicos do setor, o Amazon OpenSearch Service está qualificado para a HIPAA e está em conformidade com os seguintes padrões:

- FedRAMP
- RGPD
- PCI DSS
- ISO
- SOC

Para obter mais informações, consulte a documentação do Amazon OpenSearch Service.

Jornada de migração

Dependendo da sua implantação atual, migrar para um Amazon OpenSearch Service pode ser um procedimento básico ou complexo com várias etapas. Nas seções a seguir, você explorará as abordagens de migração e as principais considerações em cada etapa do processo. Isso inclui as práticas recomendadas baseadas em nossa experiência em ajudar muitos clientes da AWS a migrar de ferramentas existentes para o Amazon OpenSearch Service. Esta seção também analisa o que constitui uma estratégia de migração eficaz.

Uma jornada de migração padrão envolve cinco etapas:

1. Planejamento
2. Prova de conceito (PoC)
3. Implantação
4. Migrações de dados
5. substituição

Você pode estar migrando de um cluster autogerenciado do Elasticsearch ou do OpenSearch ou pode estar migrando de outra tecnologia para o Amazon OpenSearch Service. Na maioria dos casos, as etapas permanecem as mesmas. O tempo gasto em cada etapa vai variar de acordo com a complexidade do seu ambiente.

A jornada de migração começa com uma cuidadosa atividade de planejamento, seguida por um exercício de PoC para garantir que o ambiente de destino atende às suas metas de custo, segurança, performance e migração. A atividade de PoC é seguida pela implantação do ambiente de destino e pela migração dos dados para ele. Depois de confirmar que seus dados estão sincronizados entre o ambiente atual e o novo, você pode fazer a substituição para o novo ambiente. Depois da substituição, você opera o ambiente seguindo as práticas recomendadas operacionais. As seções a seguir analisam cada etapa em detalhes.

Etágio 1: planejamento

A migração começa com o planejamento do ambiente de destino que você criará para atender às suas necessidades. O planejamento envolve a análise de um conjunto de áreas de foco, cada uma das quais exigirá uma análise cuidadosa:

- Dimensionamento
- Funcionalidade
- Dependências de versão
- Indicadores-chave de desempenho (KPIs) e continuidade dos negócios
- Operações e segurança
- Treinamento
- Fluxo de dados
- Frameworks de implantação

Essas áreas de foco ajudarão você a tomar decisões que formarão a estratégia de migração. Elas também ajudam você a atingir suas metas de migração ao reduzir a complexidade e os custos da migração.

Durante a etapa de planejamento, também é fundamental avaliar seu ambiente atual e identificar os pontos problemáticos que você deseja abordar como parte dessa migração. Esses pontos problemáticos podem estar relacionados à performance, segurança, confiabilidade, velocidade de entrega, custo ou facilidade das operações. Ao analisar as áreas de foco, considere quais melhorias você pode fazer como parte da migração.

Dimensionamento

O dimensionamento ajuda você a determinar o tipo de instância, a quantidade de nós de dados e os requisitos de armazenamento corretos para seu ambiente de destino. Recomendamos que você dimensione primeiro pelo armazenamento e depois por CPUs. Se você já estiver usando o Elasticsearch or OpenSearch, o tamanho geralmente permanecerá o mesmo. No entanto, você precisa identificar o tipo de instância equivalente ao seu ambiente atual. Para ajudar a determinar o tamanho certo, recomendamos o uso das diretrizes a seguir.

Armazenamento

O dimensionamento do cluster começa com a definição dos requisitos de armazenamento. Identifique o armazenamento bruto de que você precisa para seu cluster. Isso é determinado pela avaliação dos dados gerados pelo sistema de origem (por exemplo, servidores gerando logs ou tamanho bruto do catálogo de produtos). Depois de identificar a quantidade de dados brutos que você tem, use a fórmula a seguir para calcular os requisitos de armazenamento. Depois, você pode usar o resultado como ponto de partida para sua PoC.

```
storage needed = (daily source data in bytes × 1.45) (number_of_replicas + 1) × number of days retained
```

A fórmula leva em consideração o seguinte:

- O tamanho de um índice no disco varia, mas normalmente é 10% maior do que os dados de origem.
- A sobrecarga do sistema operacional de 5% é reservada pelo Linux para recuperação do sistema e para proteção contra problemas de desfragmentação de disco.
- OpenSearch reserva 20% do espaço de armazenamento de cada instância para mesclagens de segmentos, registros e outras operações internas.
- Recomendamos manter 10% de armazenamento adicional para ajudar a minimizar o impacto da falha nos nós e das interrupções nas zonas de disponibilidade.

Combinadas, essas reservas e despesas gerais exigem 45% de espaço adicional com base nos dados brutos reais na origem. É por isso que você multiplica os dados de origem por 1,45. Em seguida, multiplique pelo número de cópias dos dados (por exemplo, uma primária mais o número de réplicas que você usará). A contagem de réplicas depende de seus requisitos de resiliência e throughput. Para um caso de uso comum, você começa com um primário e uma réplica. Por fim, multiplique pelo número de dias pelos quais você deseja reter dados em uma camada de armazenamento quente.

O Amazon OpenSearch Service oferece níveis de armazenamento quente, quente e frio. O nível de armazenamento quente usa UltraWarm armazenamento. UltraWarm fornece uma maneira econômica de armazenar grandes quantidades de dados somente para leitura no Amazon Service. OpenSearch Os nós de dados padrão usam o armazenamento quente, que assume a forma de armazenamentos de instâncias ou volumes do Amazon Elastic Block Store (Amazon EBS) anexados a cada nó. O armazenamento dinâmico fornece o desempenho mais rápido possível para indexar

e pesquisar novos dados. UltraWarm os nós usam o Amazon Simple Storage Service (Amazon S3) como armazenamento e uma solução sofisticada de armazenamento em cache para melhorar o desempenho. Para índices nos quais você não está escrevendo ativamente ou consultando com menos frequência e não têm os mesmos requisitos de desempenho, UltraWarm oferece custos significativamente mais baixos por GiB de dados. Para obter mais informações sobre UltraWarm, consulte a documentação da AWS.

Ao criar um domínio OpenSearch de serviço e usar armazenamento dinâmico, talvez seja necessário definir o tamanho do volume do EBS. Depende do tipo de instância escolhido para os nós de dados. Você pode usar a mesma fórmula de requisitos de armazenamento para determinar o tamanho do volume das instâncias baseadas no Amazon EBS. Recomendamos usar volumes gp3 para famílias de instâncias T3, R5, R6G, M5, M5g, C5 e C6g de última geração. Usando os volumes gp3 do Amazon EBS, você pode provisionar a performance independentemente da capacidade de armazenamento. Os volumes gp3 do Amazon EBS também oferecem melhor desempenho básico, a um custo 9,6% menor por GB do que os volumes gp2 existentes em serviço. OpenSearch Com o gp3, você também obtém um armazenamento mais denso nas famílias de instâncias R5, R6g, M5 e M6g, o que pode ajudar a otimizar ainda mais seus custos. Você pode criar volumes do EBS até a cota máxima permitida. Para obter mais informações sobre cotas, consulte Cotas do Amazon OpenSearch Service.

Para nós de dados que têm unidades NVM Express (NVMe), como instâncias i3 e r6gd, o tamanho do volume é fixo, portanto, os volumes do EBS não são uma opção.

Número de nós e tipos de instância

O número de nós é baseado no número de nós CPUs necessários para operar sua carga de trabalho. O número de CPUs é baseado na contagem de fragmentos. Um índice em OpenSearch é composto por vários fragmentos. Quando você cria um índice, você especifica o número de fragmentos do índice. Portanto, você precisa fazer o seguinte:

1. Calcule a contagem total de fragmentos que você pretende armazenar no domínio.
2. Determine a CPU.
3. Encontre o tipo e a contagem de nós mais econômicos que oferecem o número CPUs e o armazenamento necessários.

Geralmente, este é um ponto de partida. Execute testes para determinar se o tamanho estimado está atendendo aos seus requisitos funcionais e não funcionais.

Determinação da estratégia de indexação e contagem de fragmentos

Depois de conhecer os requisitos de armazenamento, você pode decidir de quantos índices precisa e identificar a contagem de fragmentos de cada um. Geralmente, os casos de uso de pesquisa têm um ou alguns índices, cada um representando uma entidade pesquisável ou um catálogo. Para casos de uso de analytics de logs, um índice pode representar um arquivo de logs diários ou semanais. Depois de decidir a quantidade de índices, comece com a seguinte orientação de escala e determine a contagem de fragmentos apropriada:

- Casos de uso de pesquisa: 10 a 30 GB/fragmento
- Casos de uso de analytics de logs: 50 GB/fragmento

Você pode dividir o volume total de dados em um único índice pelo tamanho do fragmento que você almeja em seu caso de uso. Isso lhe dará o número de fragmentos do índice. Identificar o número total de fragmentos ajudará você a encontrar os tipos de instância adequados à sua workload. Esses fragmentos não devem ser muito grandes nem muito numerosos. Fragmentos grandes podem dificultar OpenSearch a recuperação de falhas, mas como cada fragmento usa uma certa quantidade de CPU e memória, ter muitos fragmentos pequenos pode causar problemas e erros de desempenho. out-of-memory Além disso, o desequilíbrio na alocação de fragmentos aos nós de dados pode levar à distorção. Quando tiver índices com vários fragmentos, tente fazer a contagem de fragmentos em um múltiplo par da contagem de nós de dados. Isso ajuda a garantir que os fragmentos sejam distribuídos uniformemente entre os nós de dados e evita os nós quentes. Por exemplo, se você tiver 12 fragmentos primários, sua contagem de nós de dados deverá ser 2, 3, 4, 6 ou 12. No entanto, a contagem de fragmentos é secundária ao tamanho do fragmento. Se você tiver 5 GiB de dados, ainda deverá usar um único fragmento. Equilibrar uniformemente o número de fragmentos de réplicas em toda a zona de disponibilidade também ajuda a melhorar a resiliência.

Utilização da CPU

A próxima etapa é identificar quantos CPUs você precisa para sua carga de trabalho. Recomendamos começar com uma contagem de CPU 1,5 vez maior que a de seus fragmentos ativos. Um fragmento ativo é qualquer fragmento de um índice que está recebendo gravações substanciais. Use a contagem primária de fragmentos para determinar fragmentos ativos para índices que estão recebendo solicitações substanciais de leitura ou gravação. Para analytics de logs, somente o índice atual geralmente está ativo. Para casos de uso de pesquisa, todos os fragmentos primários serão considerados fragmentos ativos. Embora recomendemos 1,5 CPU por fragmento

ativo, isso depende muito da workload. Certifique-se de testar e monitorar a utilização da CPU e escalar adequadamente.

Uma prática recomendada para manter a utilização da CPU é garantir que o domínio do OpenSearch serviço tenha recursos suficientes para realizar suas tarefas. Um cluster com alta utilização consistente da CPU pode degradar a estabilidade do cluster. Quando seu cluster estiver sobrecarregado, o OpenSearch Serviço bloqueará as solicitações recebidas, o que resulta em rejeições de solicitações. Isso é para proteger o domínio contra falhas. As diretrizes gerais sobre o uso da CPU serão cerca de 60% em média, e 80% no máximo de utilização da CPU. Picos ocasionais de 100% ainda são aceitáveis e podem não exigir escalabilidade ou reconfiguração.

Tipos de instância

O Amazon OpenSearch Service oferece a opção de vários tipos de instância. Você pode escolher os tipos de instância de acordo com seu caso de uso. O Amazon OpenSearch Service oferece suporte às famílias de instâncias R, C, M, T e I. Você escolhe uma família de instâncias com base na workload: otimizada para memória, otimizada para computação ou mista. Depois de identificar uma família de instâncias, escolha o tipo de instância de última geração. Geralmente, recomendamos o Graviton e as gerações posteriores porque eles foram criados para oferecer melhor performance com custos mais baixos em comparação com as instâncias da geração anterior.

Com base em vários testes realizados para analytics de logs e casos de uso de pesquisa, recomendamos o seguinte:

- Para casos de uso de analytics de logs, uma diretriz geral é começar com a família R de instâncias Graviton para nós de dados. Recomendamos que você execute testes, estabeleça avaliações comparativas para seus requisitos e identifique o tamanho de instância apropriado para sua workload.
- Para casos de uso de pesquisa, recomendamos usar instâncias Graviton da família R e C para nós de dados, porque os casos de uso de pesquisa exigem mais CPU em comparação com os casos de uso de analytics de logs. Para workloads menores, você pode usar instâncias Graviton da família M para pesquisa e logs. As instâncias da família I oferecem NVMe drives e são usadas por clientes com requisitos de pesquisa de indexação rápida e baixa latência.

O cluster é composto por nós de dados e nós gerenciadores de clusters. Embora os nós principais dedicados não processem solicitações de pesquisa e consulta, seu tamanho está amplamente correlacionado ao tamanho da instância e ao número de instâncias, índices e fragmentos que podem

gerenciar. A documentação da AWS fornece uma matriz que recomenda o tipo mínimo de instância dedicada do gerenciador de clusters.

A AWS oferece uso geral (m6g), otimizado para computação (C6g) e otimizado para memória (R6g e R6gd) para o Amazon OpenSearch Service versão 7.9 ou posterior, com tecnologia de processadores AWS Graviton2. Essas instâncias são criadas usando chip personalizado projetado pela Amazon. São inovações de hardware e software projetadas pela Amazon que permitem a entrega de serviços de nuvem eficientes, flexíveis e seguros com multilocação isolada, rede privada e armazenamento local rápido.

A família de instâncias Graviton2 reduz a latência de indexação em até 50% e melhora o desempenho das consultas em até 30% em comparação com as instâncias baseadas em Intel da geração anterior disponíveis em OpenSearch serviço (M5, C5, R5).

Funcionalidade

A área de foco da funcionalidade ajuda a garantir que você não perca nenhuma funcionalidade ao migrar para um ambiente de destino da Amazon OpenSearch Service. Recomendamos prestar muita atenção aos seguintes aspectos:

- Funcionalidade atual da solução
- Funcionalidade OpenSearch do Amazon Service
- Plug-ins empacotados

Funcionalidade atual da solução

Recomendamos que você analise sua solução atual e determine os recursos e plug-ins APIs que você usa na pilha de tecnologia atual (por exemplo, Elasticsearch ou outra solução). OpenSearch Determine qual funcionalidade é essencial para sua empresa, o que pode ser modificado e o que pode ser descartado durante a migração.

Funcionalidade OpenSearch do Amazon Service

Para garantir que a funcionalidade necessária esteja disponível após a migração, recomendamos que você realize uma análise da OpenSearch versão mais recente suportada pelo Amazon OpenSearch Service, incluindo os recursos que ele oferece e os plug-ins que estão disponíveis no Amazon OpenSearch Service. Você quer confirmar se a plataforma de destino é compatível com os

recursos de que você precisa (por exemplo, gerenciamento do estado do índice, que automatiza a transferência dos índices, ou recursos de machine learning, como detecção de anomalias). Mapeie a funcionalidade existente da sua solução atual para recursos no Amazon OpenSearch Service que fornecem capacidade equivalente para que você possa continuar a oferecer suporte às suas cargas de trabalho.

Para obter mais informações sobre a funcionalidade disponível em cada versão compatível do Elasticsearch ou do OpenSearch software, consulte a documentação do Amazon OpenSearch Service.

Plug-ins empacotados

O Amazon OpenSearch Service oferece suporte a vários plug-ins que fazem parte do OpenSearch projeto de código aberto. Se você estiver usando algum plug-in licenciado da suíte Elasticsearch que faça parte do X-Pack ou não, talvez queira determinar um plug-in equivalente ou recurso nativo nas ofertas. OpenSearch Talvez você também queira registrar isso como um ponto a ser comprovado na etapa de PoC.

OpenSearch tem vários plug-ins que fornecem recursos de nível corporativo equivalentes aos plug-ins licenciados. Para determinar o plug-in e a versão corretos para o ambiente de destino, revise a lista de plug-ins por versões da documentação do OpenSearch serviço. Embora o Amazon OpenSearch Service ofereça suporte a vários OpenSearch plug-ins prontos para uso, você pode estar usando um OpenSearch plug-in de código aberto que atualmente não está disponível no Amazon OpenSearch Service. Para solicitar a adição do plug-in ao roteiro futuro do Amazon OpenSearch Service, entre em contato com a AWS.

Plugins personalizados

No momento da elaboração deste guia, não havia suporte para plug-ins personalizados. Portanto, você precisará considerar formas alternativas de fornecer a função e a experiência personalizadas do plug-in. Se sua solução usa plug-ins personalizados, analise a funcionalidade para determinar se você pode portar os plug-ins personalizados para o ambiente de destino usando plug-ins compatíveis com o Amazon OpenSearch Service ou recursos nativos contidos nele OpenSearch. Recomendamos testar e comprovar todas as opções de plug-ins durante a etapa de PoC. A migração é um bom momento para avaliar a funcionalidade atual da solução e determinar se ela é essencial para seus negócios.

Dependências de versão

A área de foco das dependências de versão ajuda você a criar um roteiro de sua jornada de migração por meio de várias versões para alcançar a versão mais recente do Amazon OpenSearch Service. Considere os seguintes pontos principais:

- Seleção da versão do mecanismo
- Atualização para a versão mais recente
- Estratégia de atualização da versão
- Verificações pré-atualização

Seleção da versão do mecanismo

É muito importante considerar cuidadosamente as dependências de versão. O Amazon OpenSearch Service oferece suporte a várias versões do Elasticsearch e a todas as principais versões OpenSearch do mecanismo. (No entanto, a versão mais recente do OpenSearch pode levar algumas semanas para ser suportada no Amazon OpenSearch Service a partir da data de lançamento.) Recomendamos que você analise os recursos suportados pela versão do mecanismo na documentação do Amazon OpenSearch Service para identificar a versão correta para suas necessidades. Ao escolher a mesma versão principal (e a secundária mais próxima), você pode usar a abordagem de restauração de snapshots para migrar. Geralmente, esta é a abordagem mais direta.

Atualizando para a versão mais recente do OpenSearch Serviço

Embora você possa operar uma versão anterior do Amazon OpenSearch Service, é altamente recomendável atualizar para a versão mais recente disponível. Isso ajuda você a aproveitar as melhorias de performance, a confiabilidade, a redução de custos e os muitos novos recursos disponíveis nas versões mais recentes do mecanismo. A migração é uma boa oportunidade para reduzir a dívida técnica que pode ser decorrente da execução de versões anteriores do software.

Estratégia de atualização da versão

Se você decidir que deseja atualizar para a versão mais recente do software durante a migração, determine as etapas e uma estratégia de atualização. A documentação OpenSearch do Amazon Service fornece informações sobre caminhos de upgrade. É importante entender as mudanças

significativas entre as diferentes versões. Em alguns casos, as alterações significativas podem exigir que você planeje ajustes na modelagem e no projeto do índice.

> **Note**
>
> Observação: a funcionalidade Vários tipos de mapeamento está disponível somente nas versões 5.x e anteriores do Elasticsearch. Os índices criados nas versões 6.x e posteriores são compatíveis somente com um único tipo de mapeamento para cada índice. Se você estiver usando vários tipos de mapeamento, recomendamos remodelar esses dados em vários índices.

No caso de uma migração urgente, considere uma opção básica em que você executa uma migração de versão equivalente (por exemplo, 5.x para 5.x) e, em seguida, atualize a versão do OpenSearch Serviço em uma data posterior. OpenSearch O serviço oferece atualizações no local para domínios que executam as versões 5.1 (se compatível) ou posterior e 1.0 ou posterior do Elasticsearch. OpenSearch Faça um teste para verificar se seus índices são compatíveis com atualizações no local ao executar a versão 5.x do Elasticsearch. Isso significa que você poderá migrar para a versão equivalente e realizar uma atualização no local depois de fazer as alterações necessárias para tornar seus índices e outras funcionalidades compatíveis com a versão mais recente. Revise a documentação de atualização de domínio com cuidado.

Verificações pré-atualização

A funcionalidade OpenSearch de atualização do Amazon Service pode realizar verificações de pré-atualização examinando o ambiente para determinar problemas que podem bloquear a atualização. A atualização não prosseguirá para a próxima etapa, a menos que essas verificações tenham êxito.

KPIs e continuidade dos negócios

É essencial que, durante a migração, você estabeleça suas metas de negócios e indicadores-chave de desempenho (KPIs) para medir o sucesso. É importante determinar suas metas no início do processo de migração e estabelecer uma linha de base para seu sistema atual para que você possa determinar melhorias mensuráveis. As metas comuns nas jornadas dos clientes incluem as seguintes:

- Melhorar a agilidade operacional.

Nessa meta, você pode avaliar e comparar sua implantação existente com o ambiente de destino usando as seguintes métricas:

- Tempo médio para o provisionamento do cluster.
- Tempo para a distribuição da implantação para uma nova região geográfica
- Tempo médio para a configuração da segurança do cluster
- Tempo médio para escalar seu ambiente (como adicionar nós e armazenamento)
- Tempo médio para detectar consultas com baixa performance e tempo médio para corrigi-las
- Tempo médio para atualizar a versão do software

- Reduzir o custo total de propriedade (TCO).

  Para calcular seu TCO atual, você pode usar as seguintes métricas:

  - Número de horas da equipe para criar e operar a solução (desenvolvimento DevOps, monitoramento, escalabilidade, backup, restauração)
  - Custo da licença associado ao software existente
  - Custos do data center (aquisição e atualização de hardware, eletricidade, resfriamento, espaço, racks, equipamentos de rede)
  - Horas da equipe para configurar a solução (instalações de software, rede)
  - Custo das auditorias de conformidade (HIPAA, PCI DSS, SOC, ISO, RGPD, FedRAMP)
  - Custo da configuração da segurança (criptografia em repouso e em trânsito, configuração de autenticação e autorização, controle de acesso refinado)
  - Custo de reter um grande volume de dados warm e frios
  - Custo da configuração da alta disponibilidade em todas as zonas de disponibilidade
  - Custo do provisionamento excessivo para evitar a aquisição frequente de hardware ou o manuseio de cargas de pico

  Essa lista não é exaustiva.

- Monitore o tempo de atividade e outros contratos de nível de serviço (). SLAs SLAs que você pode medir e melhorar migrando para o novo ambiente incluem o seguinte:

  - Tempo de atividade total (dados históricos de tempo de atividade da implantação existente em comparação com o SLA de 99,9% fornecido pelo Amazon Service) OpenSearch
  - Recuperação de falhas (objetivo do ponto de recuperação e do tempo de recuperação)
  - Tempo de resposta associado a várias funções (por exemplo, pesquisa e indexação)

- Número de usuários simultâneos
- Tempo de replicação entre diferentes geografias e clusters.

Ao migrar para o Amazon OpenSearch Service, use um processo iterativo para verificar se você está cumprindo ou excedendo essas metas KPIs e se está alcançando os resultados desejados.

Performance operacional

Uma área importante a ser analisada em sua solução atual são as métricas de performance. Estabeleça uma avaliação comparativa e determine as melhorias que você espera alcançar em seu ambiente de destino. Isso inclui seu SLA de tempo de atividade e requisitos de latência. Ele ajudará você a estabelecer e, na maioria dos casos, a melhorar seus níveis de serviço atuais. Normalmente, os clientes analisam os seguintes indicadores de nível de serviço:

- Leituras e gravação por segundo
- Latência de leitura e gravação
- Percentual de tempo de atividade

Quando você arquiteta o seu SLAs, é importante entender completamente o Acordo de Nível OpenSearch de Serviço e Serviço da Amazon.

Performance do processo

Para estabelecer metas de continuidade dos negócios, é importante avaliar a performance atual do seu processo. Identifique e analise os runbooks existentes ou os procedimentos operacionais padrão (SOPs) da plataforma atual e determine as áreas em que sua equipe passa a maior parte do tempo. A migração é uma boa oportunidade para trabalhar na melhoria dessas áreas para que sua equipe possa se concentrar em inovar, criar funcionalidades de negócios e melhorar a experiência do cliente. Você pode identificar os pontos problemáticos do seu ambiente existente revisando os dados históricos de suporte ou os tíquetes de problemas para determinar o tempo gasto pela equipe de suporte e desenvolvimento resolvendo esses problemas. O registro das métricas abaixo pode ajudar você a avaliar as melhorias fornecidas pelo seu ambiente de destino:

- Tempo médio até a falha (MTTF) (tempo de atividade)
- Tempo médio entre falhas (MTBF)
- Tempo médio para detectar (MTTD) uma falha

- Tempo médio de resolução (MTTR)
- Número de tíquetes de suporte recebidos

Transição suave para novos serviços

Para garantir a continuidade dos negócios de seus serviços, é importante planejar cuidadosamente uma transição sem interrupções. A migração é um bom momento para modernizar sua aplicação e os serviços associados à sua plataforma de pesquisa ou analytics de logs. No entanto, você deseja planejar uma estratégia de substituição cuidadosa que não afete seus serviços existentes. A seção de estratégia de substituição neste documento fornece informações sobre como planejar uma substituição perfeita para o ambiente de destino.

Métricas financeiras

Pode haver muitos motivos para migrar para o Amazon OpenSearch Service, mas o custo geralmente é um fator importante. Entenda o custo total de propriedade (TCO) do ambiente existente para que você possa avaliar a economia de custos obtida ao migrar para o serviço gerenciado. Você pode começar com a lista de métricas que se encontram na meta Reduzir o custo total de propriedade. A AWS publicou um estudo de avaliação comparativa de valor da nuvem que pode ajudar as equipes a criar um caso comercial de migração para a Nuvem AWS. Embora o estudo não seja específico do Amazon OpenSearch Service, ele abrange as principais áreas de valor que são comuns na maioria das migrações para a nuvem, incluindo a migração para o Amazon OpenSearch Service.

Na maioria dos casos, o Amazon OpenSearch Service oferece menor TCO. Ao calcular o TCO, é fundamental incorporar o custo da equipe. Analisar o tempo e o custo que seus engenheiros gastam para manter o ambiente atual é um fator importante. Muitos clientes comparam somente o custo da infraestrutura de armazenamento, computação e rede com o custo do serviço gerenciado. No entanto, isso pode não fornecer um custo total de propriedade preciso. O Amazon OpenSearch Service fornece eficiência operacional à sua equipe gerenciando tarefas que, de outra forma, precisariam ser executadas por seus engenheiros. Isso inclui as seguintes tarefas:

- Escalabilidade de um cluster adicionando ou removendo nós
- Aplicação de patches
- Atualização no local
- Realização de backups
- Configuração de ferramentas de monitoramento para registrar logs e métricas

Essas atividades são automatizadas pelo serviço, e a AWS oferece uma equipe de suporte em nível de produção. Isso significa que sua equipe pode se concentrar em atividades que agregam valor direto ao seu negócio.

Operações e segurança

Quando você migra para o Amazon OpenSearch Service, suas atividades operacionais mudam. Você não será mais responsável por provisionar nós, adicionar armazenamento, instalar e aplicar patches no sistema operacional, configurar e manter a alta disponibilidade, a escalabilidade e outras atividades de baixo nível. Em vez disso, você pode focar sua atenção na criação de seus casos de uso e novas experiências de usuário.

O Amazon OpenSearch Service oferece recursos de registro, monitoramento e solução de problemas com os quais você precisará se familiarizar para otimizar seus processos operacionais.

Runbooks e novos processos

Durante a fase de planejamento, identifique os processos existentes que precisarão ser modificados ou eliminados. Em seguida, você pode adicionar novos processos operacionais para os quais talvez não tivesse largura de banda no passado.

Embora o Amazon OpenSearch Service elimine o trabalho pesado indiferenciado, você ainda precisará garantir que seu aplicativo seja projetado e monitorado para oferecer o melhor desempenho. Você precisará configurar o monitoramento e os alertas do seu domínio para estar totalmente ciente de quaisquer problemas de integridade causados por fatores internos ou externos. Você precisará agendar e iniciar as atualizações para as versões mais recentes.

Todas essas atividades operacionais exigirão a criação de runbooks e a modificação dos runbooks existentes. Para monitorar a infraestrutura e analisar métricas operacionais no Amazon OpenSearch Service, é fundamental manter os runbooks. Os runbooks garantem que você opere de forma consistente de acordo com seus requisitos regulatórios e de conformidade. Se você não usa runbooks, é um bom momento para considerar o uso deles. Crie processos para executar periodicamente etapas pré-planejadas para garantir que os processos de remediação, como recuperação de falhas de aplicações e falhas inesperadas, sejam totalmente automatizados.

Sistema de geração de tíquetes e suporte

Para registrar incidentes associados às suas implantações, recomendamos planejar e operar um sistema de geração de tíquetes (talvez você já esteja fazendo isso). Talvez você precise treinar

sua equipe de suporte sobre como criar tíquetes de suporte com o AWS Support. Recomendamos simplificar o processo de escalonamento durante a triagem de tíquetes.

A seção Excelência operacional, mais adiante neste guia, fornecerá links para várias práticas recomendadas e áreas que talvez você precise considerar em seus runbooks e desenvolver processos.

Segurança

Na AWS, a segurança é nossa maior prioridade. O Amazon OpenSearch Service fornece segurança em várias camadas. O serviço cuida de todos os patches de segurança e oferece isolamento de rede por meio da VPC, do controle de acesso refinado e do suporte multilocatário. Seus dados são criptografados em repouso usando chaves que você cria e controla por meio do AWS Key Management Service (AWS KMS). O recurso de node-to-node criptografia fornece Transport Layer Security (TLS) para todas as comunicações entre instâncias em um domínio. O Amazon OpenSearch Service também está qualificado para a HIPAA e está em conformidade com os padrões PCI DSS, SOC, ISO e FedRAMP para ajudar você a atender aos requisitos regulamentares ou específicos do setor.

Durante a etapa de planejamento, identifique as pessoas e os processos que interagem com o domínio, escolha uma topologia de rede e planeje a autenticação e autorização para cada entidade principal. Dependendo dos requisitos de segurança e conformidade organizacionais, você pode usar vários recursos de segurança para criar um ambiente que atenda às necessidades da sua empresa. Além disso, considere os seguintes fatores:

- VPC — Você pode configurar o Amazon OpenSearch Service em uma nuvem privada virtual (VPC) na AWS. Esta é a configuração recomendada. Não recomendamos criar um domínio com um endpoint público. Planeje criar a arquitetura de rede necessária para permitir que seus usuários e aplicações clientes acessem o ambiente de destino.
- Autenticação — O Amazon OpenSearch Service oferece suporte a várias formas de autenticar um usuário ou cliente de software. Ele oferece suporte à autenticação Amazon Cognito ou SAML com seu provedor de identidade existente para acessar painéis. OpenSearch Também oferece integração com identidades do IAM e autenticação HTTP básica usando um banco de dados interno de usuários. Você deve planejar configurar e testar uma opção apropriada para autenticação. Para obter mais informações, consulte a documentação OpenSearch de segurança do serviço.

- Autorização: recomendamos seguir o princípio de privilégio mínimo ao configurar o acesso ao serviço. O Amazon OpenSearch Service fornece controle de acesso refinado para ajudá-lo a configurar o acesso nos níveis de documento, linha e coluna.

Familiarize-se com os recursos de segurança e teste-os durante o estágio de PoC.

Treinamento

Quando você inicia sua jornada de migração para a AWS, suas equipes de desenvolvimento de software, operações, suporte e segurança precisam estar equipadas com o conhecimento do Amazon OpenSearch Service. Considere todas as equipes que interagem com sua solução. Quando você está migrando de um Elasticsearch ou de um OpenSearch ambiente, a maior parte do conhecimento pode ser transferida. Ofereça treinamento às seguintes equipes:

- Equipe de desenvolvimento de software — Eduque sua equipe de desenvolvimento de software sobre os recursos APIs e recursos, como mecanismos para configurar a ingestão de dados.
- Equipe de operações — Treine sua equipe de operações sobre como interagir com os domínios do Amazon OpenSearch Service, monitorar métricas operacionais e acessar registros usando a Amazon CloudWatch. Os membros da equipe devem aprender a configurar alarmes automatizados para avisar quando os domínios de OpenSearch serviço precisam de atenção. Se você estiver migrando de um conjunto de ferramentas existente que usa localmente, como o Splunk, identifique as opções de monitoramento no Amazon OpenSearch Service que podem fornecer visibilidade semelhante às suas cargas de trabalho.
- Equipe de suporte — instrua sua equipe de suporte sobre como implementar runbooks que envolvam recursos de OpenSearch serviço. Talvez você queira atualizar os runbooks e os procedimentos de gerenciamento de eventos para usar os serviços do AWS Support.
- Equipe de segurança — eduque sua equipe de segurança sobre como configurar um controle de acesso refinado e como se integrar aos provedores de identidade existentes (). IDPs

Opções de treinamento

O AWS Training and Certification oferece treinamento digital e presencial para iniciantes e profissionais sobre as habilidades de nuvem necessárias para criar e operar soluções na AWS. O conteúdo é criado por especialistas da AWS e é atualizado regularmente. Você tem várias opções de treinamento.

Você pode trabalhar com sua equipe de contas da AWS para ajudar você a identificar um recurso apropriado. A seguir estão alguns dos recursos que você pode usar para ajudar a aprimorar suas equipes no Amazon OpenSearch Service:

- Dias de imersão: os arquitetos de soluções da AWS podem oferecer dias de imersão, que são workshops práticos personalizados para abordar casos de uso, padrões comuns de implementação e itens de roteiro que podem estar especificamente relacionados a casos de uso.
- Workshops práticos: as equipes podem acompanhar workshops de autoatendimento criados por especialistas da AWS.
- Whitepapers e guias: os whitepapers da AWS são uma ótima maneira de expandir seu conhecimento sobre a nuvem. De autoria da AWS e da comunidade AWS, eles fornecem conteúdo detalhado que geralmente aborda situações específicas de clientes.
- Publicações no blog: escritas por especialistas e clientes da AWS, essas postagens discutem os últimos anúncios, as práticas recomendadas, as soluções, os recursos de serviços, os casos de uso de clientes e outros tópicos.
- Melhores práticas — Participe de palestras on-line ou em conferências, ou de sessões conduzidas por especialistas da AWS que ajudam você a entender as melhores práticas do Amazon OpenSearch Service.
- AWS Professional Services: a equipe de AWS Professional Services pode fornecer as práticas recomendadas e aconselhamento prescritivo. A equipe oferece um programa de treinamento para ajudar os profissionais de TI a entender e realizar migrações com êxito.

Fluxo de dados

A área de foco do fluxo de dados inclui as três seguintes áreas:

- Ingestão de dados
- Retenção de dados
- Abordagem da migração de dados

Ingestão de dados

A ingestão de dados se concentra em como colocar dados em seu domínio do Amazon OpenSearch Service. Uma compreensão completa das fontes e formatos de dados é fundamental ao escolher a estrutura de ingestão certa para. OpenSearch

Há muitas maneiras diferentes de criar ou modernizar seu projeto de ingestão. Há muitas ferramentas de código aberto para criar um pipeline de ingestão autogerenciado. OpenSearch O serviço oferece suporte à integração com Fluentd, Logstash ou Data Prepper. OpenSearch Essas ferramentas são conhecidas pela maioria dos desenvolvedores de soluções de analytics de logs. Você pode implantar essas ferramentas em uma instância do Amazon EC2, no Amazon Elastic Kubernetes Service (Amazon EKS) ou on-premises. Tanto o Logstash quanto o Fluentd oferecem suporte aos domínios do OpenSearch Amazon Service como destino de saída. No entanto, isso exigirá que você mantenha, aplique patches, teste e mantenha as versões do software Fluentd ou Logstash atualizadas.

Para reduzir sua sobrecarga operacional, você pode usar um dos serviços AWS gerenciados que oferecem suporte à integração com o Amazon OpenSearch Service. Por exemplo, o Amazon OpenSearch Ingestion é um coletor de dados totalmente gerenciado e sem servidor que fornece dados de log, métricas e rastreamento em tempo real para os domínios do Amazon Service. OpenSearch Com a OpenSearch Ingestão, você não precisa mais usar soluções de terceiros, como Logstash ou Jaeger, para ingerir dados em seus domínios de serviço. OpenSearch Você configura seus produtores de dados para enviar dados para o OpenSearch Ingestion. Em seguida, ele entrega automaticamente os dados para o domínio ou coleção que você especificar. Você também pode configurar a OpenSearch ingestão para transformar seus dados antes de entregá-los.

Outra opção é o Amazon Data Firehose, que é um serviço totalmente gerenciado que ajuda a criar um pipeline de ingestão sem servidor. O Firehose fornece uma maneira segura de ingerir, transformar e entregar dados de streaming para os domínios do Amazon OpenSearch Service. Ele pode escalar automaticamente para corresponder ao throughput de seus dados, e não requer administração contínua. O Firehose também pode transformar registros recebidos usando AWS Lambda, compactando e agrupando os dados em lotes antes de carregá-los em seu domínio de serviço. OpenSearch

Com um serviço gerenciado, você pode desativar seu pipeline de ingestão de dados existente ou aumentar sua configuração atual para reduzir a sobrecarga operacional.

O planejamento da migração é um bom momento para avaliar se seu pipeline de ingestão atual atende às necessidades dos casos de uso atuais e futuros. Se você estiver migrando de um Elasticsearch ou OpenSearch cluster autogerenciado, seu pipeline de ingestão deve suportar a troca dos endpoints do cluster atual para o domínio do Amazon OpenSearch Service com o mínimo de atualizações na biblioteca do cliente.

Retenção de dados

Ao planejar a ingestão e o armazenamento de dados, certifique-se de planejar e concordar com a retenção de dados. Para casos de uso de analytics de logs, é fundamental que você tenha as políticas certas criadas em seu domínio para descontinuar os dados históricos. Ao migrar de uma arquitetura existente baseada em VM on-premises e na nuvem, você pode estar usando um tipo específico de instância para todos os seus nós de dados. Os nós de dados têm o mesmo perfil de armazenamento, memória e CPU. A maioria dos clientes configura o armazenamento de alto throughput para atender às suas necessidades de indexação de alta velocidade. Essa arquitetura de perfil de armazenamento singular é chamada de arquitetura apenas de nó quente, ou simplemente hot-only. A arquitetura hot-only associa armazenamento com computação, o que implica que você precisará adicionar nós de computação se sua necessidade de armazenamento aumentar.

Para dissociar o armazenamento da computação, o Amazon OpenSearch Service oferece o nível UltraWarm de armazenamento. UltraWarm fornece uma maneira econômica de armazenar dados somente para leitura no Amazon OpenSearch Service, fornecendo nós que podem acomodar um volume maior de dados do que os nós de dados tradicionais.

Durante o planejamento, decida os requisitos de processamento e retenção de dados. Para reduzir o custo de sua solução existente, aproveite o UltraWarm nível. Identifique o requisito de retenção de seus dados. Depois, crie políticas de gerenciamento do estado do índice para mover os dados de quente para warm ou para excluir os dados automaticamente do domínio quando não forem necessários. Isso também ajuda a garantir que seu domínio não fique sem espaço de armazenamento.

Abordagens da migração de dados

Durante a etapa de planejamento, é fundamental que você decida sobre uma abordagem específica de migração de dados. Sua abordagem de migração de dados determina como você move os dados que estão em seu armazenamento de dados atual para o armazenamento de destino sem nenhuma lacuna. Os detalhes processuais dessas abordagens são abordados na seção Etapa 4: migração de dados, que é quando você implementa sua abordagem.

Esta seção aborda diferentes formas e padrões que você pode usar para migrar um Elasticsearch ou cluster OpenSearch para o Amazon Service. OpenSearch Ao escolher um padrão, considere a seguinte lista de fatores (não exaustiva):

- Se você deseja copiar dados de um cluster autogerenciado existente ou se está recriando da fonte de dados original (arquivos de log, banco de dados do catálogo de produtos)

- Compatibilidade de versão do Elasticsearch ou OpenSearch cluster de origem e do domínio Amazon OpenSearch Service de destino
- Aplicativos e serviços dependentes do Elasticsearch ou do cluster OpenSearch
- A janela disponível para a migração
- O volume de dados indexados em seu ambiente existente

Criar de um snapshot

Os snapshots são a forma mais popular de migrar de um cluster autogerenciado do Elasticsearch para o Amazon Service. OpenSearch Os snapshots fornecem uma forma de fazer backup de seus dados OpenSearch ou do Elasticsearch usando um serviço de armazenamento durável, como o Amazon S3. Com essa abordagem, você captura um instantâneo do seu Elasticsearch ou OpenSearch ambiente atual e o restaura no ambiente de destino do Amazon OpenSearch Service. Depois de restaurar o snapshot, você pode direcionar sua aplicação para o novo ambiente. Esta é uma solução mais rápida nas seguintes situações:

- Sua origem e destino são compatíveis.
- O cluster existente contém um grande volume de dados indexados, o que pode levar tempo para ser reindexado.
- Seus dados de origem não estão disponíveis para reindexação.

Para considerações adicionais, consulte Considerações sobre snapshots na seção Etapa 4: migração de dados.

Criar da origem

Essa abordagem implica que você não moverá dados do seu Elasticsearch ou OpenSearch cluster atual. Em vez disso, você recarrega os dados diretamente do seu log ou da fonte do catálogo de produtos para o domínio de destino do Amazon OpenSearch Service. Isso geralmente é feito com pequenas alterações nos pipelines de ingestão de dados existentes. No caso de uso da análise de registros, a criação a partir da fonte também pode exigir o recarregamento dos registros históricos de suas fontes para o novo ambiente OpenSearch de serviço. Para casos de uso de pesquisa, pode ser necessário que você recarregue seu catálogo completo de produtos e conteúdo no novo domínio do Amazon OpenSearch Service. Essa abordagem funciona bem nos seguintes cenários:

- Suas versões do ambiente de origem e de destino não são compatíveis com a restauração de snapshots.

- Você deseja alterar seu modelo de dados no ambiente de destino como parte da migração.
- Você quer ir para a versão mais recente do Amazon OpenSearch Service para evitar atualizações contínuas e quer resolver as alterações mais importantes de uma só vez. Isso pode ser uma boa ideia se você estiver gerenciando por conta própria uma versão relativamente mais antiga (5.x ou anterior) do Elasticsearch.
- Talvez você queira alterar sua estratégia de indexação. Por exemplo, em vez de fazer o roll over todos os dias, você pode fazê-lo todos os meses no novo ambiente.

Para obter informações sobre as opções de criação da origem, consulte 2. Criando da origem na seção Etapa 4: migração de dados.

Reindexe remotamente a partir de um Elasticsearch ou ambiente existente OpenSearch

Essa abordagem usa a API de reindexação remota do Amazon OpenSearch Service. Usando a reindexação remota, você pode copiar dados diretamente do seu Elasticsearch ou cluster existente no local ou na nuvem OpenSearch para o seu domínio do Amazon Service. OpenSearch Você pode criar uma automação que possa manter os dados sincronizados entre os dois locais do ambiente até que você passe para o ambiente de destino.

Usar ferramentas de migração de dados de código aberto

Há várias ferramentas de código aberto disponíveis para migrar dados do seu ambiente Elasticsearch existente para o ambiente de destino da Amazon. OpenSearch Um exemplo é o utilitário Logstash. Você pode usar o utilitário Logstash para extrair dados de um Elasticsearch ou OpenSearch cluster e copiá-los para o domínio do Amazon Service. OpenSearch

Recomendamos que você avalie todas as suas opções e opte por aquela com a qual se sinta mais confortável. Para garantir que a abordagem selecionada seja à prova de falhas, teste todas as suas ferramentas e automação durante o estágio de PoC. Para obter detalhes e step-by-step orientações sobre como implementar essas abordagens, consulte a seção Etapa 4 — Migração de dados.

Frameworks de implantação

Muitas equipes modernas usam integração contínua e entrega contínua (CI/CD) practices and pipelines to automate the deployment of their solutions and infrastructure. If your team already uses CI/CDpipelines). Você deve ser capaz de incorporar o Amazon OpenSearch Service em seu ambiente. Se você estiver implantando manualmente em sua configuração atual, considere criar

pipelines para automatizar o trabalho repetível, reduzir a sobrecarga operacional e reduzir os erros humanos.

Você pode implantar o Amazon OpenSearch Service usando uma variedade de estruturas de infraestrutura como código (IaC), incluindo Terraform by HashiCorp, Chef e Puppet. O Terraform oferece um OpenSearch módulo que você pode usar para criar domínios do Amazon OpenSearch Service. Em muitos casos, você pode usar seu pipeline de implantação de infraestrutura existente e direcionar o módulo do mecanismo de pesquisa para o módulo Amazon OpenSearch Service.

Se você está pensando em criar pipelines do zero ou se quiser usar os serviços nativos da AWS, a AWS fornece várias opções CI/CD de ferramentas e serviços. Incluindo o seguinte:

- AWS CodePipeline
- AWS CodeBuild
- AWS Cloud Development Kits (AWS CDK)
- AWS CloudFormation
- AWS CodeDeploy

Você pode usar esses serviços para automatizar a criação, o teste e a implantação da infraestrutura. Implantar seus pipelines usando qualquer um desses serviços nativos da nuvem tem muitas vantagens, incluindo as seguintes:

- Lançamentos de produtos totalmente automatizados end-to-end (criação, teste, implantação)
- Implantação em vários ambientes (desenvolvimento, teste, pré-produção, produção)
- Integração com outros serviços da AWS
- A capacidade de modernizar seus pipelines de implantação para automatizar implantações do OpenSearch Amazon Service em vários ambientes

Etapa 2: prova de conceito

Ao realizar uma migração, é fundamental comprovar se a solução do estado final funcionará conforme necessário. É altamente recomendável realizar um exercício de prova de conceito (PoC). Esta seção se concentra nos vários aspectos a serem considerados ao executar uma PoC:

- Definição de critérios de entrada e saída
- Garantia de financiamento
- Automatização
- Testes completos
- Etapas da PoC
- Simulação de falhas

Definição de critérios de entrada e saída

Ter critérios claros de entrada e saída é fundamental para um exercício de PoC com êxito. Ao definir seus critérios de entrada, considere o seguinte:

- Definição de caso de uso
- Acesso a ambientes
- Familiaridade com os diversos serviços
- Requisitos de treinamentos associados

Da mesma forma, defina os critérios de saída que você pode usar para avaliar o resultado da PoC, incluindo o seguinte:

- Funcionalidade
- Requisitos de performance
- PoC de implementações de segurança

Garantia de financiamento

Com base na definição dos critérios da PoC, garanta o financiamento para a PoC. Certifique-se de ter realizado o dimensionamento correto e considerado todos os custos associados. Se você estiver

migrando de on-premises para a AWS, inclua o custo associado à migração de seus frameworks on-premises para a Nuvem AWS. Se você já for um cliente da AWS, trabalhe com seu gerente da conta da AWS para entender se você se qualifica para receber créditos que podem ser usados para a migração para o Amazon OpenSearch Service.

Automatização

Identifique onde a automação pode ser feita e planeje uma trilha dedicada para automatizar e limitar o tempo dos testes. A implantação e os testes automatizados ajudam você a repetir o ciclo, testar e validar em um ritmo rápido e sem erros introduzidos por humanos.

Ao limitar o tempo de um teste, você pode garantir a entrega dentro do prazo e passar para outras atividades se surgirem desafios. Por exemplo, se seus testes de performance estiverem demorando mais do que o tempo estimado, você poderá pausar essa atividade. Você pode então passar para outras atividades de teste e validação enquanto seus desenvolvedores corrigem os problemas. Você pode voltar aos testes deperformance depois que os problemas forem resolvidos. Faça uma avaliação comparativa da performance de sua solução existente e crie testes de performance automatizados que possam validar o efeito de suas alterações de configuração durante o PoC.

Testes completos

Teste todas as partes da pilha certificando-se de realizar as validações necessárias para as diferentes camadas, como pipelines de ingestão e mecanismos de consulta, que se integram ao seu domínio do Amazon OpenSearch Service. Isso o ajudará a validar a implementação completa da solução.

Camada de apresentação

Na camada de apresentação, certifique-se de executar um exercício de PoC que inclua as seguintes atividades:

- Autenticação: valide os mecanismos planejados para autenticar seus usuários.
- Autorização: identifique os mecanismos de autorização que você deseja seguir e confirme se eles estão funcionando conforme o esperado.
- Consulta: quais são os casos de uso mais comuns que você encontrará na produção? Quais são alguns cenários extremos que são críticos para sua empresa? Identifique esses padrões e valide-os durante a PoC.

- Renderização: os dados estão sendo renderizados de forma precisa e adequada para vários usuários em todos os casos de uso? Para casos de uso de analytics de logs, talvez você queira criar e testar o painel no OpenSearch Dashboards ou no Kibana, dependendo da versão de destino, para confirmar se ele atende aos seus requisitos.

Camada de ingestão

Na camada de ingestão, certifique-se de avaliar vários componentes, como coleta, armazenamento em buffer, agregação e armazenamento:

- Coleta: para casos de uso de analytics de logs, valide se todos os dados que você está registrando em log estão sendo coletados. Para casos de uso de pesquisa, identifique as fontes que alimentam os dados e realize validações sobre a integridade e a exatidão dos dados para garantir que a fase de coleta tenha sido executada com sucesso.
- Buffer: se você tiver um pico no tráfego, talvez queira ter certeza de que está armazenando em buffer os dados que estão sendo ingeridos. Existem vários formas de criar um projeto de armazenamento em buffer. Por exemplo, é possível coletar dados no Amazon Data Firehose ou usar o armazenamento do Amazon S3 como um buffer.
- Agregação: valide qualquer agregação de dados, como o uso em massa da API, que você realiza durante a ingestão.
- Armazenamento: valide se o armazenamento é capaz de lidar de forma ideal com a ingestão que você está realizando.

Etapas da PoC

Recomendamos que você use as etapas a seguir para implementar sua PoC e validar o resultado. Não tenha medo de iterar por essas fases da PoC e ajustar o plano dela, mesmo que você tenha investido tempo no planejamento prévio.

- Teste funcional e teste de carga: certifique-se de que todos os níveis estejam sendo testados por completo. Simule falhas em todas as partes da pilha. Por exemplo, se você tiver um cluster com dois nós grandes e um deles falhar, o outro nó deverá absorver todo o tráfego do seu cluster. Nesse cenário, ter um número maior de nós menores pode resultar em uma recuperação mais eficiente de uma falha no nó. Teste suas workloads em cargas de pico e superiores para garantir que a performance não seja afetada nesses cenários. Durante o teste, levante as questões

com antecedência para que quaisquer possíveis problemas sejam avaliados por várias partes interessadas no momento certo.

- Verificação dos KPIs e ajuste: durante a PoC, certifique-se de atender aos KPIs e resultados comerciais definidos nos critérios de saída da PoC. Ajuste as configurações de forma que elas atendam aos KPIs.
- Automatização e implantação: a automação e o monitoramento são os outros aspectos principais nos quais se concentrar durante o teste de PoC. Refine suas etapas de automação e valide-as junto com o monitoramento detalhado para fornecer a todas as partes interessadas informações suficientes para avaliar com confiança os resultados da PoC. Documente todas as etapas e crie um runbook que você possa reutilizar para a migração de produção.

Simulação de falhas

É altamente recomendável que você simule um cenário de falha e valide se seu projeto oferece a resiliência e a tolerância a falhas necessárias para atender aos requisitos do usuário. Talvez você queira simular uma falha em um nó de dados para ver se seu cluster tem recursos suficientes para lidar com a recuperação sem problemas. Para verificar se seu domínio pode ficar sobrecarregado com a ingestão de grandes volumes, você pode testar as configurações de buffer simulando uma expansão repentina de logs de algumas de suas origens. Valide se seu projeto não excede nenhuma cota ao escalar para uma implantação de produção. Para obter mais informações, consulte a documentação do Amazon OpenSearch Service sobre cotas de serviços.

Etapa 3: implantação

Quando chegar à etapa de implantação, você terá concluído sua PoC e terá uma boa ideia de como implantar seu ambiente de destino na produção. Lembre-se dos seguintes fatores:

- Validação da automação: durante a implantação, execute a automação que você criou durante a PoC e confirme se ela está funcionando conforme o esperado. Além disso, valide se sua automação de CI/CD está funcionando conforme o esperado ao fazer alterações no código de configuração.
- Verificação da segurança: é fundamental verificar se todas as suas configurações de segurança estão funcionando conforme o esperado e se seus dados estão seguros. Confirme se a solução foi avaliada de acordo com os padrões de segurança da sua empresa, como a integração do provedor de identidades, e se seus principais usuários podem fazer login e acessar dados que estão autorizados a acessar.
- Monitoramento: certifique-se de ter testado suas configurações de monitoramento e configurado os alertas recomendados. Monitore métricas importantes, como CPU, memória, discos, JVMs e alocações de fragmentos. Para fornecer insights sobre a integridade do seu domínio do Amazon OpenSearch Service e das integrações associadas, você pode criar um painel no Amazon CloudWatch. Você pode verificar se sua equipe de suporte operacional tem acesso ao painel. A seção Excelência operacional fornece links para dicas úteis para configurar um domínio do OpenSearch Service resiliente e de alta performance.
- Testes de alarmes: certifique-se de testar todos os seus alarmes. Se você estiver usando o Amazon CloudWatch ou um plug-in de alerta, verifique se todas as integrações, como o Amazon Simple Notification Service (Amazon SNS) ou o Slack, funcionam conforme o esperado. Simule alertas para validar se eles foram entregues corretamente no canal de destino. Confirme se o texto do alerta fornece informações úteis. Por exemplo, o alerta pode fornecer um link para o runbook associado para que sua equipe de suporte implemente um processo de remediação associado.

Etapa 4: migração de dados

Agora que seu ambiente de destino está pronto, você pode implementar a estratégia de migração de dados que escolheu durante a etapa de planejamento.

Esta seção aborda as etapas de implementação dos quatro padrões diferentes:

- Criação com base em um snapshot
- Criação com base na fonte
- Reindexação remota
- Uso do Logstash

1. Criação com base em um snapshot

Ao usar a abordagem de restauração de instantâneos, você copia dados do Elasticsearch ou cluster de origem OpenSearch para o domínio de destino do Amazon Service. OpenSearch

Em termos gerais, o processo de restauração de snapshots consiste nas seguintes etapas:

1. Obter um snapshot dos dados necessários (índices) do cluster existente e fazer upload do snapshot para um bucket do S3.
2. Crie um domínio do Amazon OpenSearch Service.
3. Dê permissões ao Amazon OpenSearch Service para acessar o bucket e conceda à sua conta de usuário permissões para trabalhar com snapshots. Criar um repositório de snapshots e apontá-lo para o seu bucket.
4. Restaure o snapshot no domínio do Amazon OpenSearch Service.
5. Direcione seus aplicativos clientes para o domínio do Amazon OpenSearch Service.
6. Criar políticas do Index State Management (ISM) para configurar a retenção (opcional).

Os snapshots são incrementais. Portanto, um snapshot pode ser executado e restaurado de forma incremental. Usando snapshots, você pode extrair dados em massa como arquivos em um sistema de armazenamento (por exemplo, o Amazon S3). Em seguida, você pode carregar esses arquivos no ambiente de destino usando a operação da API `_restore`. Isso elimina a necessidade de reindexação, que é demorada, e também reduz o tráfego na rede.

Considerações sobre snapshots

Ao usar a abordagem de restauração de snapshots, considere o seguinte:

- Você não pode pesquisar ou reindexar enquanto um índice está sendo restaurado. No entanto, você pode pesquisar e reindexar um índice enquanto o snapshot está sendo obtido.
- O Elasticsearch ou as OpenSearch versões de origem e destino devem ser compatíveis. Um snapshot de um índice que foi criado em:
  - 5.x pode ser restaurado para 6.x
  - 2.x pode ser restaurado para 5.x
  - 1.x pode ser restaurado para 2.x
- Como essa é uma point-in-time restauração do Elasticsearch ou do OpenSearch snapshot, as alterações subsequentes no cluster de origem não serão replicadas para o domínio de destino do Amazon Service. OpenSearch Você pode interromper a ingestão dos dados no Elasticsearch ou no OpenSearch cluster de origem até que a restauração seja concluída, ou você pode repetir o processo de restauração do snapshot algumas vezes. Como o snapshot é incremental, somente as alterações serão copiadas e restauradas no ambiente de destino em menos tempo do que na primeira restauração. Depois que a restauração for concluída com sucesso, você direciona os aplicativos de ingestão para o domínio do Amazon OpenSearch Service.
- A captura de um snapshot inclui, por padrão, um snapshot do estado do cluster e de todos os índices. Ao migrar do Elasticsearch, talvez seja necessário criar políticas de ciclo de vida de índice equivalentes no ambiente de destino usando o recurso ISM em. OpenSearch O Elasticsearch Index Lifecycle Management (ILM) não é compatível com o Amazon Service. OpenSearch
- Você não pode restaurar um snapshot para uma versão anterior do Elasticsearch ou. OpenSearch Por exemplo, você não pode restaurar um snapshot da versão 7.10 para a 7.9. Da mesma forma, você não pode restaurar snapshots do Elasticsearch 7.11 ou posterior para um domínio do Amazon Service. OpenSearch Se você migrou seu ambiente autogerenciado do Elasticsearch para a versão 7.11 ou posterior, você pode usar o Logstash para carregar dados do cluster Elasticsearch e gravá-los no domínio. OpenSearch
- Você exporta um snapshot para um local de armazenamento designado chamado repositório. Elasticsearch ou OpenSearch cria vários arquivos no repositório. Você não pode modificar nem excluir esses arquivos. Isso pode criar inconsistências ou fazer com que o processo de restauração falhe.

2. Criação com base na fonte

Conforme descrito anteriormente, criar a partir da fonte é a abordagem em que você não migra dados do Elasticsearch ou do ambiente atual. OpenSearch Em vez disso, você cria índices no domínio de destino diretamente do seu log, da fonte de dados do catálogo de produtos ou da fonte do conteúdo.

Duas opções estão disponíveis para a criação com base na fonte. A opção escolhida depende do tipo de dados dos seus dados:

- Usando o AWS Database Migration Service — Se a fonte dos seus dados for um sistema de gerenciamento de banco de dados relacional (RDBMS) e a fonte for suportada pelo AWS Database Migration Service (AWS DMS), você poderá usar o AWS DMS para copiar dados da sua fonte de dados para o domínio de destino do Amazon Service. OpenSearch O AWS DMS é compatível com opções de carga total e captura de dados de alteração (CDC). Na opção de carregamento completo, a tarefa do AWS DMS copia todos os dados da tabela do banco de dados de origem para um OpenSearch índice de destino. Você pode usar o mapeamento padrão ou fornecer configurações de mapeamento personalizadas. Na opção CDC, o AWS DMS primeiro faz uma cópia completa dos registros da tabela de origem em um índice de destino OpenSearch . Em seguida, ele captura os dados alterados (atualizações e inserções) e os copia para o OpenSearch índice. Para obter mais informações, consulte as postagens do blog Introducing Amazon Elasticsearch Service as a target in AWS Database Migration Service e Scale Amazon Elasticsearch Service for AWS Database Migration Service migrations.
- Criação com base na fonte do documento: se sua fonte de dados não for um RDBMS ou não for compatível com o AWS DMS, talvez seja necessário criar uma solução personalizada usando ferramentas de código aberto ou uma combinação de ferramentas de código aberto e serviços da AWS. Você deve converter seus dados de origem em documentos JSON antes que eles possam ser carregados OpenSearch. Se você já tem pipelines configurados da sua fonte para o seu Elasticsearch ou OpenSearch ambiente atual, você pode direcionar esses pipelines de dados OpenSearch com as alterações apropriadas nas bibliotecas de clientes e (se necessário) mudanças no modelo de dados nos índices no domínio do Amazon Service. OpenSearch Ao criar índices com base na fonte, leve em conta o seguinte:
  - A localização dos documentos: os documentos já podem estar disponíveis na Nuvem AWS, em armazenamento de objetos, como o Amazon S3, ou podem estar armazenados em um local de armazenamento on-premises, como um sistema de arquivos.
  - O formato dos documentos — Os documentos já podem estar no formato JSON, prontos para serem ingeridos no domínio do Amazon OpenSearch Service, ou talvez precisem ser limpos,

processados e formatados em JSON antes de serem ingeridos no domínio do Amazon Service. OpenSearch

Criar com base na fonte envolve as seguintes etapas de alto nível:

1. Defina o mapeamento e as configurações do índice no domínio do Amazon OpenSearch Service.
2. Extrair dados da fonte do documento e copiá-los em um local de armazenamento de objetos, como o Amazon S3. Você pode usar uma ferramenta de código aberto (por exemplo, o Logstash), um cliente de serviço da AWS (por exemplo, o agente do Amazon Kinesis), uma ferramenta comercial de terceiros ou um programa personalizado.
3. Configure uma ferramenta de código aberto (por exemplo, Logstash ou Fluent Bit) ou um serviço nativo da AWS (por exemplo, AWS Lambda ou AWS DMS) para converter dados em documentos JSON e carregá-los periodicamente ou continuamente do armazenamento de objetos para o domínio do Amazon Service. OpenSearch

Para obter mais informações, consulte Carregamento de dados de streaming no Amazon OpenSearch Service.

3. Reindexação remota

Nesse caso, os índices do Elasticsearch ou OpenSearch cluster autogerenciado de origem são migrados para o domínio do OpenSearch Amazon Service usando a operação de API de reindexação de documentos. Você pode usar a operação da API de reindexação de documentos para criar um índice a partir de um Elasticsearch ou índice existente. OpenSearch O índice existente pode estar no mesmo cluster em que você executa a operação de reindexação, ou pode estar em um cluster remoto. O Amazon OpenSearch Service oferece suporte ao uso da operação de API de reindexação de documentos com clusters remotos. Você pode reindexar de um índice em um Elasticsearch autogerenciado para um índice no Amazon Service. OpenSearch

A reindexação remota oferece suporte ao Elasticsearch 1.5 e posterior para o cluster remoto do Elasticsearch e ao OpenSearch Amazon Service 6.7 e posterior para o domínio local. Para obter mais informações, consulte a publicação do blog Migrate data into Amazon ES using remote reindex. A postagem do blog se refere ao Amazon Elasticsearch, mas a orientação se aplica igualmente aos domínios do Amazon OpenSearch Service.

4. Uso do Logstash

O Logstash é uma ferramenta de processamento de dados de código aberto que pode coletar dados da fonte, realizar transformação ou filtragem e enviar dados para um ou mais destinos. Para gravar dados no domínio do Amazon OpenSearch Service, o Logstash fornece os seguintes plug-ins:

- logstash-input-elasticsearch
- logstash-input-opensearch
- logstash-output-opensearch

Para obter mais informações, consulte Carregamento de dados no Amazon OpenSearch Service com o Logstash e a postagem do OpenSearch blog Apresentando o logstash-input-opensearch plug-in para. OpenSearch

Etapa 5: substituição

Esta etapa discute várias abordagens que você pode empregar para migrar do seu Elasticsearch ou OpenSearch ambiente atual para o domínio de destino do Amazon Service. OpenSearch A substituição pode ser realizada em duas etapas:

- Estabeleça um mecanismo de sincronização de dados para manter o ambiente de destino sincronizado com a origem.
- Execute a troca do ambiente atual para o ambiente de destino com ou sem tempo de inatividade.

Sincronização de dados

Para qualquer sistema que receba dados contínuos, a migração de dados pode exigir que você pare de receber novos dados durante a migração e execute a migração em uma janela de manutenção (com possível tempo de inatividade). Se você não puder arcar com o tempo de inatividade, poderá capturar as alterações depois de iniciar a migração. Você repete as alterações no destino para mantê-lo atualizado e sincronizado com a origem até realizar a substituição. As seções a seguir discutem várias maneiras de manter a origem e o destino sincronizados.

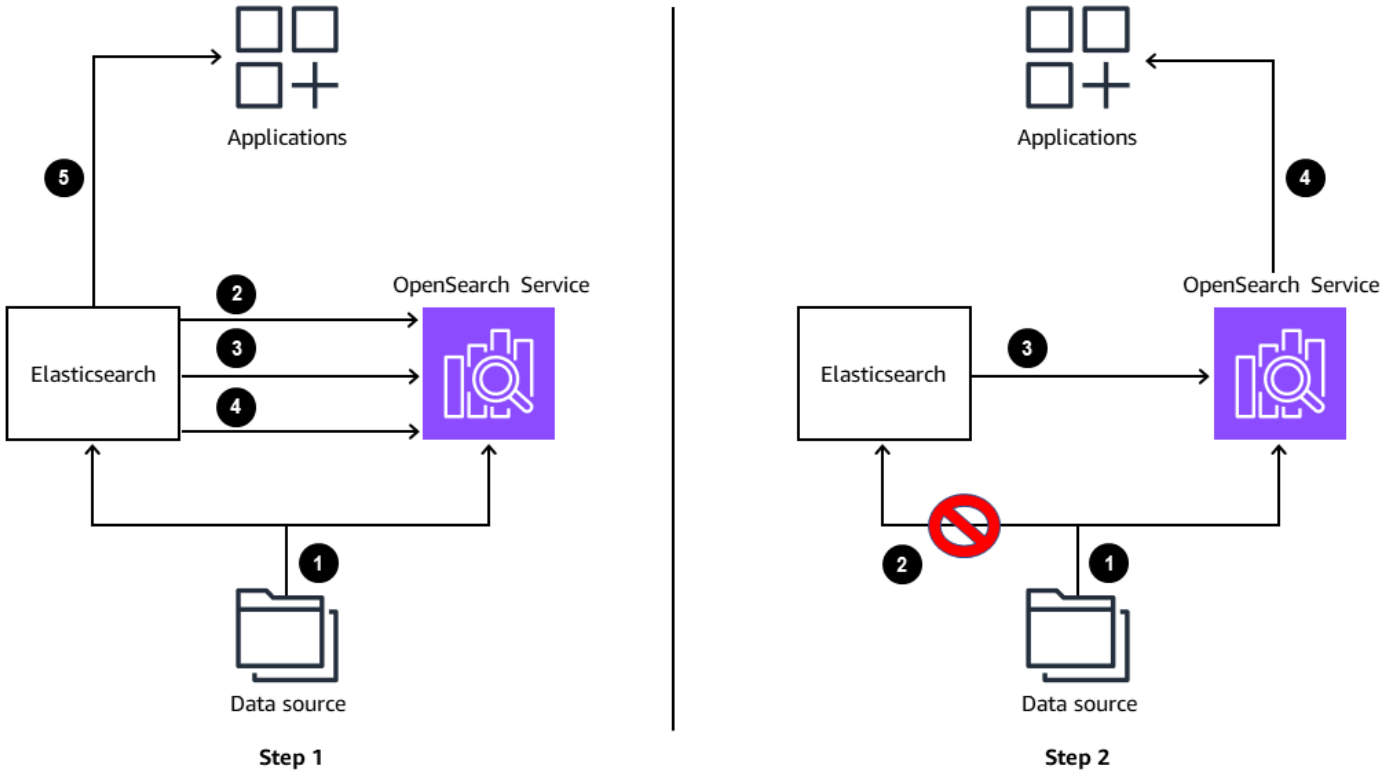
Workloads de analytics de logs

Para workloads de analytics de logs, você pode realizar uma sincronização de atualização das seguintes formas:

- Você pode executar dois ambientes lado a lado até que o período de retenção e a ingestão nos ambientes atual e de destino sejam concluídos. Em algum momento, você decide substituir e direcionar suas aplicações para o novo ambiente. Às vezes, você pode ingerir novos dados das fontes de registro ou documento para o cluster existente e os ambientes de OpenSearch serviço de destino. Você pode então preencher os dados mais antigos no ambiente de destino copiando-os do ambiente atual. Em todos os casos, você deve garantir que seus dados não tenham lacunas que afetem seus usuários.
- Antes da migração de dados, você pode decidir pausar sua ingestão no ambiente atual. No entanto, essa abordagem significa que seus usuários talvez não consigam pesquisar os dados mais recentes ou alterados do seu ambiente atual até que a migração de dados seja concluída. Após a conclusão da migração de dados, você pode direcionar sua ingestão de dados para o ambiente de destino e transferir suas aplicações e clientes para o ambiente de destino. Isso significa que nenhum dado novo estará disponível até que a migração seja concluída. No entanto,

o sistema permanecerá disponível para pesquisa. Você deve ter os meios para manter os logs e dados de origem em sua fonte até que o novo ambiente esteja disponível.

- Você pode continuar a usar o mecanismo atual de analytics de logs até que sua primeira leva de dados seja migrada. Depois, você preenche os dados restantes que foram produzidos desde o início da migração da primeira leva. Supondo que os dados restantes sejam muito menores do que os da primeira leva, você pode pausar a ingestão enquanto os dados restantes estão sincronizados, pois a sincronização pode levar apenas alguns minutos ou pode levar algumas horas. Você também pode realizar algumas etapas usando essa abordagem até que a janela de sincronização fique pequena o suficiente para pausar a ingestão da origem para o ambiente de destino e migrar para o ambiente de destino sem afetar seus usuários. O diagrama a seguir mostra o uso de snapshot incremental e restauração para atualizar ou sincronizar dados.



Etapa 1

1. Os dados fluem da fonte por meio do pipeline de ingestão de dados para o ambiente atual do Elasticsearch e o domínio do Amazon OpenSearch Service.
2. A primeira passagem leva mais tempo para ser transferida do Elasticsearch para o domínio Amazon OpenSearch Service.
3. A primeira atualização ou sincronização leva menos tempo.

4. A segunda atualização ou sincronização é a que demora menos tempo.
5. Os dados continuam a fluir do Elasticsearch para as aplicações.

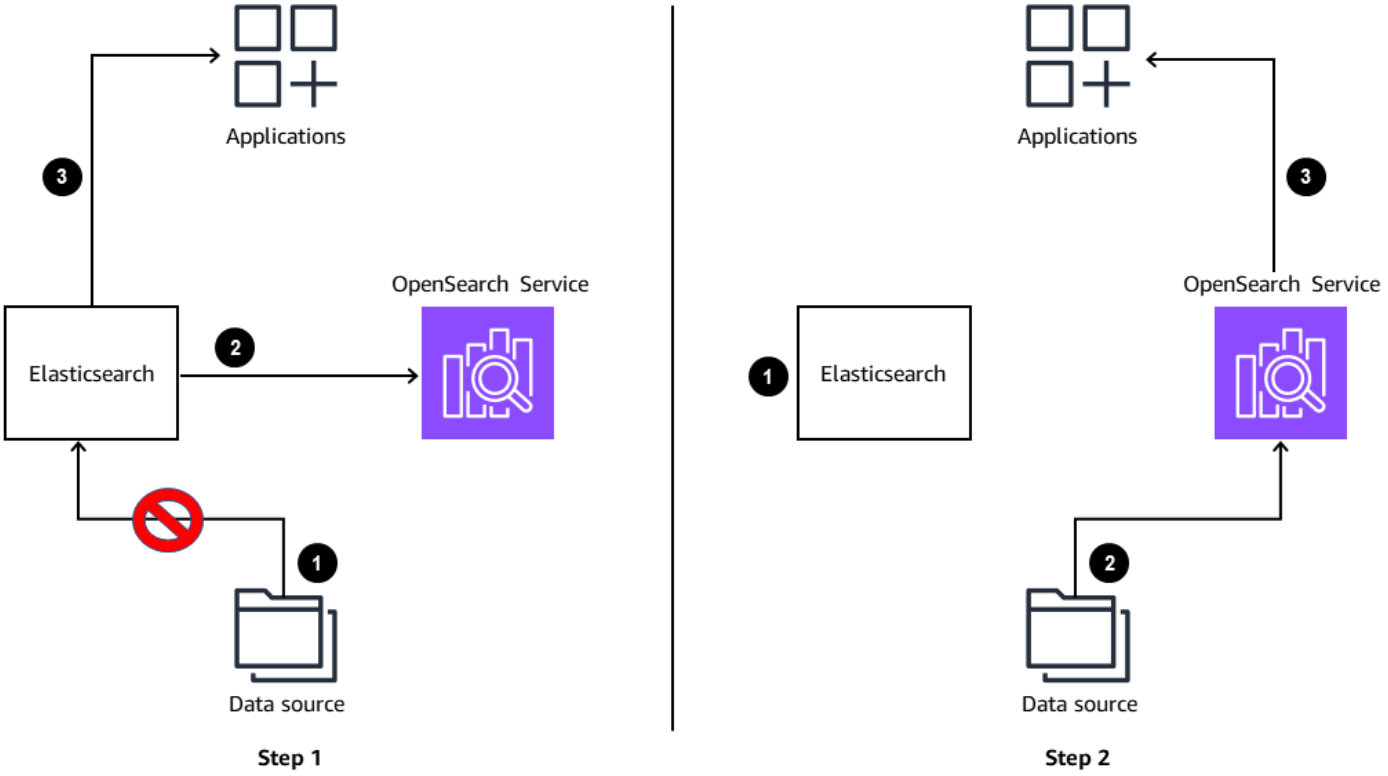
Etapa 2

1. Os dados fluem da fonte através do pipeline de ingestão de dados para o domínio do OpenSearch Serviço.
2. A ingestão no ambiente atual do Elasticsearch foi interrompida.
3. A atualização ou a sincronização final é a mais rápida.
4. Os dados fluem do OpenSearch Serviço para os aplicativos.

Workloads de pesquisa

Nas três abordagens analisadas anteriormente, você deve garantir que todos os dados no seu destino estejam atualizados antes de realizar a substituição. Para workloads de pesquisa, você pode considerar as seguintes sugestões para atualização ou sincronização:

- Para workloads de pesquisa, normalmente você pausa a ingestão da origem para o ambiente atual. Você copia todos os seus dados do ambiente atual para o ambiente de destino e implementa um mecanismo de captura de dados de alteração (CDC) que pode determinar quais dados foram alterados desde o início da migração. Em seguida, você copia os dados alterados para o OpenSearch ambiente da Amazon. Na maioria dos casos, os pipelines de ingestão de dados da aplicação de pesquisa já têm um mecanismo CDC incorporado e, geralmente, é uma questão de direcionar seu pipeline para o novo ambiente após a migração dos dados do ambiente atual. O diagrama a seguir mostra a criação de um índice inteiramente da fonte para casos de uso de pesquisa.



Etapa 1

1. A ingestão no ambiente atual do Elasticsearch está pausada.
2. Os dados são copiados do ElasticSearch domínio do OpenSearch Serviço.
3. Os dados continuam fluindo ElasticSearch para os aplicativos.

Etapa 2

1. O ambiente do Elasticsearch não está mais conectado à fonte de dados ou às aplicações.
2. Os dados de captura de dados de alteração (CDC) são ingeridos no pipeline e fluem para o domínio do OpenSearch Serviço.
3. Os dados fluem do domínio do OpenSearch Serviço para os aplicativos.

- Algumas cargas de trabalho de pesquisa exigem o carregamento somente de dados completos do banco de dados de origem ou da fonte de dados para o novo ambiente OpenSearch de serviço. Após a conclusão do carregamento, as aplicações cliente poderão ser migradas para o novo ambiente. Esta é a maneira mais simples de realizar a migração para workloads de pesquisa.

Trocar ou substituir

A etapa final da jornada de migração é trocar, ou substituir, para o novo ambiente. É uma das fases críticas. Neste momento, está tudo pronto para ser lançado. Você tem os dados sincronizados e atualizados, tem monitoramento e alertas configurados, seus runbooks estão atualizados e está tudo pronto para migrar para o novo ambiente. Você deve garantir que sua ingestão esteja fluindo normalmente e que as métricas do seu novo ambiente estejam íntegras. Durante esse estágio, você planeja e executa a transferência das conexões do cliente do seu Elasticsearch ou OpenSearch cluster existente para o novo domínio do Amazon OpenSearch Service. Fique atento a quaisquer alterações na biblioteca do cliente que possam ser necessárias. Neste ponto, você deve ter testado todas as funcionalidades do seu cliente com o Amazon OpenSearch Service em seus ambientes inferiores para verificar a compatibilidade e o desempenho.

Se você tiver uma aplicação cliente que precise apontar para o novo ambiente, atualize a entrada de DNS do antigo para o novo ambiente. Em seguida, monitore de perto o comportamento da aplicação para garantir que seus usuários tenham a experiência certa.

Geralmente, se você seguiu as diretrizes deste documento, você terá uma transição segura. No entanto, recomendamos que você mantenha seu ambiente de origem atualizado para que ele possa atuar como um substituto caso você encontre algum problema com o novo ambiente. Alguns clientes da AWS continuam operando os dois ambientes por algumas semanas após a troca antes de desativar o ambiente antigo. Recomendamos que você escolha uma estratégia que se alinhe aos seus requisitos de continuidade de negócios.

Etapa 6: excelência operacional

A documentação do Amazon OpenSearch Service tem uma seção dedicada às melhores práticas operacionais. Os tópicos úteis incluem:

- Monitoramento e alertas
- Estratégia de fragmentos
- Estabilidade
- Desempenho
- Segurança
- Otimização de custos
- Dimensionamento de domínios do Amazon OpenSearch Service
- Escala de petabytes no Amazon Service OpenSearch
- Nodes mestres dedicados no Amazon OpenSearch Service
- CloudWatch Alarmes recomendados para o Amazon Service OpenSearch

Recomendamos que você siga as orientações fornecidas na documentação para operar seu ambiente recém-migrado.

Conclusão

O Amazon OpenSearch Service elimina o trabalho pesado e indiferenciado que é necessário para desenvolver e operar clusters autogerenciados do Elasticsearch ou do OpenSearch. Se você está considerando uma migração para o Amazon OpenSearch Service, é possível usar o processo descrito neste guia para planejar e escolher uma estratégia de migração que seja eficaz no seu caso.

As migrações podem ser tão básicas quanto obter um snapshot de um cluster autogerenciado e restaurá-lo no domínio do Amazon OpenSearch Service, ou podem ser tão complexas quanto testar todas as funcionalidades e integrações existentes. Este guia fornece informações que podem ser usadas pelas equipes do projeto de migração para garantir que elas tenham abordado todos os aspectos de uma migração e para criar uma estratégia de implementação robusta.

A documentação do Amazon OpenSearch Service tem uma seção dedicada às práticas operacionais recomendadas. Recomendamos que você siga as orientações fornecidas na documentação para operar seu ambiente recém-migrado.

Recursos

- Criação de snapshots de índices no Amazon OpenSearch Service
- Use Amazon S3 to Store a Single Amazon OpenSearch Service Index (publicação do blog)
- Elasticsearch snapshot and restore (documentação do Elasticsearch)
- S3 Repository Plugin (documentação do Elasticsearch)
- Elasticsearch Repository Settings: Recommended S3 Permissions (documentação do Elasticsearch)
- Elasticsearch Client Settings (documentação do Elasticsearch)

Colaboradores

Colaboradores

Os colaboradores deste documento incluem:

- Muhammad Ali, arquiteto principal de soluções OpenSearch
- Gene Alpert, especialista sênior em gerenciamento de contas técnicas, Analytics
- Jon Handler, arquiteto sênior principal de soluções
- Prashant Agrawal, arquiteto sênior de soluções especialista em OpenSearch
- Ina Felsheim, gerente sênior de marketing de produto
- Sung-il Kim, arquiteto sênior de soluções de analytics
- Hajer Bouafif, arquiteto de soluções OpenSearch
- Kevin Fallis, arquiteto principal de soluções especialista em OpenSearch
- Muthu Pitchaimani, arquiteto sênior de soluções especialista em OpenSearch
- Kunal Kusoorkar, gerente, arquiteto de soluções OpenSearch
- Imtiaz Sayed, líder técnico principal de arquitetura de soluções de analytics
- Soujanya Konka, arquiteta sênior de soluções
- Marc Clark, gerente, especialista em OpenSearch
- Bob Taylor, especialista sênior em OpenSearch
- Aneesh Chandra PN, arquiteto principal de soluções de analytics, Saúde e ciências biológicas

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
Publicação inicial	—	28 de agosto de 2023

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.