



Arquivamento de dados no Amazon RDS for MySQL, Amazon RDS for MariaDB e Aurora My - Compatível SQL

# AWS Orientação prescritiva



# AWS Orientação prescritiva: Arquivamento de dados no Amazon RDS for MySQL, Amazon RDS for MariaDB e Aurora My - Compatível SQL

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

As marcas comerciais e imagens comerciais 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
Resultados direcionados .....	2
Arquivar a partir de tabelas particionadas .....	4
Arquivar a partir de tabelas não particionadas .....	6
Mova dados para o Amazon S3 .....	8
Use o SELECT INTO OUTFILE S3 .....	8
Use o Data Pipeline .....	9
Use AWS Glue .....	9
Acessando dados arquivados .....	13
Classe de armazenamento padrão .....	13
Classes de armazenamento do S3 Glacier .....	15
Práticas recomendadas .....	16
Limpeza .....	18
Recursos .....	19
Apêndice I .....	20
Apêndice II .....	22
Histórico do documento .....	25
Glossário .....	26
# .....	26
A .....	27
B .....	30
C .....	32
D .....	35
E .....	40
F .....	42
G .....	43
H .....	44
I .....	45
L .....	48
M .....	49
O .....	53
P .....	56
Q .....	59

---

R .....	59
S .....	62
T .....	66
U .....	67
V .....	68
W .....	68
Z .....	69
.....	lxx

# Arquivamento de dados no Amazon RDS para MySQL, Amazon RDS para MariaDB e Aurora, compatível com MySQL

Shyam Sunder Rakhecha, Abhishek Karmakar, Oliver Francis e Saumya Singh Amazon Web Services (AWS)

Abril de 2023([histórico do documento](#))

A necessidade de arquivar dados históricos pode resultar de diferentes casos de uso. Seu aplicativo pode ter sido projetado sem a capacidade de arquivamento, e o crescimento de seus negócios ao longo do tempo pode resultar em grandes quantidades de dados históricos. Isso inevitavelmente leva à degradação do desempenho. Você também pode reter dados históricos devido aos requisitos de conformidade em sua organização.

Este guia discute como arquivar seus dados históricos no Amazon Simple Storage Service (Amazon S3) com impacto mínimo em seu aplicativo e recuperar informações arquivadas quando necessário.

## Visão geral

Este guia aborda diferentes abordagens para arquivar dados históricos de tabelas grandes no Amazon Relational Database Service (Amazon RDS) para MySQL, no Amazon RDS para MariaDB e na edição compatível com o Amazon Aurora MySQL na nuvem da Amazon Web Services (AWS). Neste guia, você aprenderá a arquivar dados de tabelas particionadas e dados que não estão particionados e residem em tabelas grandes. Você pode implementar as abordagens apresentadas no guia para reduzir o tamanho de seus dados ativos e, ao mesmo tempo, manter dados históricos importantes para análise posterior.

Arquivar os dados da tabela regularmente resulta em um conjunto menor de dados ativos em suas tabelas, o que leva a leituras e gravações mais rápidas e melhora o desempenho do seu aplicativo. O arquivamento regular de dados se enquadra nos pilares de excelência operacional e eficiência de desempenho do [Estrutura bem arquitetada](#). Quando você move dados antigos para o Amazon Simple Storage Service (Amazon S3) e limpa seus dados arquivados em sua instância do Amazon RDS ou em um cluster compatível com o Aurora MySQL, você pode economizar nos custos de armazenamento. Isso se encaixa no pilar de otimização de custos e ajuda a evitar custos desnecessários na AWS.

# Resultados de negócios direcionados

Este guia se concentra nos seguintes resultados comerciais:

- Experiência aprimorada do usuário
- Requisitos de conformidade de dados atendidos
- Custos de armazenamento reduzidos
- Dados organizados

## Experiência aprimorada do usuário

Bancos de dados que retêm dados históricos podem ter um desempenho lento devido a tabelas e índices grandes. Ao arquivar seus dados históricos, você reduz suas tabelas e índices. Isso tem um impacto positivo direto nas operações de API voltadas para o cliente que interagem com seu banco de dados.

## Requisitos de conformidade de dados atendidos

Setores como serviços financeiros, organizações do setor público e saúde têm requisitos rigorosos de arquivamento. Ao arquivar dados de aplicativos que residem em seu banco de dados Amazon RDS para MySQL, Amazon RDS para MariaDB ou Aurora MySQL no Amazon S3, você pode atender aos requisitos de conformidade regulamentada, incluindo os seguintes:

- Padrão de segurança de dados do setor de cartão de pagamento (PCI DSS – Payment Card Industry Data Security Standard)
- Lei de Portabilidade e Responsabilidade de Seguros de Saúde (HIPAA) e Lei de Tecnologia da Informação em Saúde Econômica e Clínica (HITECH)
- Federal Risk and Authorization Management Program (FedRAMP)
- General Data Protection Regulation (GDPR – Regulamento geral de proteção de dados)
- Padrões federais de processamento de informações (FIPS) 140-2
- Instituto Nacional de Padrões e Tecnologia (NIST) 800—171

## Custos de armazenamento reduzidos

Manter os dados no Amazon RDS aumenta o custo de armazenamento e exige maior IOPS. Se você comparar o custo de armazenamento por GB por mês para [Amazon RDS para MySQL Multi-AZ](#)

---

[GP2](#) com o de [Amazon S3 Glacier](#) naus-east-1 Na região da AWS, o custo de armazenamento do S3 Glacier é cerca de 57 vezes menor do que o do Amazon RDS.

### Dados organizados

É bom manter dados informativos que serão acessados com frequência pelo aplicativo no banco de dados. No entanto, os aplicativos geram uma grande quantidade de dados que não são necessários com muita frequência ou se tornam obsoletos. Esses registros podem ser arquivados e mantidos no local, o que é econômico e não afeta o desempenho dos aplicativos.

## Arquivamento de dados em tabelas particionadas

Suporte ao MySQL [particionamento](#) para o mecanismo de armazenamento InnoDB, e você pode usar esse recurso para particionar tabelas grandes. As partições dentro da tabela são armazenadas como tabelas físicas separadas, embora o SQL que opera na tabela particionada leia a tabela inteira. Isso lhe dá a liberdade de remover partições desnecessárias da tabela sem executar `row-by-row` `exclude`, para que você possa arquivar linhas históricas em seu banco de dados.

Considere o código de exemplo a seguir. `TABLE orders` existe dentro do `orderprocessingschema`. Seus dados históricos estão presentes na partição `historical`, que contém dados pertencentes a 2021 e anteriores. Na mesma tabela, os dados ativos em nível de aplicativo estão presentes nas partições ativas de cada mês de 2022. Para arquivar os dados na partição `historical`, você pode criar um arquivo `table orders_2021_and_older` com a mesma estrutura no `archiveeschema`. Você pode então usar o MySQL [PARTIÇÃO DE TROCA](#) para mover a partição `historical` naquela mesa. Observe que a tabela de arquivamento não está particionada. Após o arquivamento, você pode verificar seus dados e movê-los para o Amazon S3.

```
CREATE TABLE orders (
  orderid bigint NOT NULL AUTO_INCREMENT,
  customerid bigint DEFAULT NULL,
  .....
  .....
  order_date date NOT NULL,
  PRIMARY KEY (`orderid`, `order_date`))
PARTITION BY RANGE (TO_DAYS(order_date)) (
  PARTITION pstart VALUES LESS THAN (0),
  PARTITION phistorical VALUES LESS THAN (TO_DAYS('2022-01-01')),
  PARTITION p2022JAN VALUES LESS THAN (TO_DAYS('2022-02-01')),
  PARTITION p2022FEB VALUES LESS THAN (TO_DAYS('2022-03-01')),
  PARTITION p2022MAR VALUES LESS THAN (TO_DAYS('2022-04-01')),
  PARTITION p2022APR VALUES LESS THAN (TO_DAYS('2022-05-01')),
  PARTITION p2022MAY VALUES LESS THAN (TO_DAYS('2022-06-01')),
  PARTITION p2022JUN VALUES LESS THAN (TO_DAYS('2022-07-01')),
  PARTITION p2022JUL VALUES LESS THAN (TO_DAYS('2022-08-01')),
  PARTITION p2022AUG VALUES LESS THAN (TO_DAYS('2022-09-01')),
  PARTITION p2022SEP VALUES LESS THAN (TO_DAYS('2022-10-01')),
  PARTITION p2022OCT VALUES LESS THAN (TO_DAYS('2022-11-01')),
  PARTITION p2022NOV VALUES LESS THAN (TO_DAYS('2022-12-01')),
  PARTITION p2022DEC VALUES LESS THAN (TO_DAYS('2023-01-01')),
  PARTITION pfuture VALUES LESS THAN MAXVALUE
```

```
);  
CREATE TABLE orders_2021_and_older (  
    orderid bigint NOT NULL AUTO_INCREMENT,  
    customerid bigint DEFAULT NULL,  
    .....  
    .....  
    order_date date NOT NULL,  
    PRIMARY KEY (`orderid`, `order_date`));  
mysql> alter table orderprocessing.orders exchange partition phistorical with table  
archive.orders_2021_and_older;  
Query OK, 0 rows affected (0.33 sec)
```

Quando você usa o `EXCHANGE PARTITION` recurso para arquivar dados históricos, recomendamos as seguintes práticas recomendadas:

- Crie um esquema separado para armazenar dados arquivados em seu aplicativo. Esse esquema conterá tabelas de arquivamento que armazenarão os dados arquivados. Uma tabela de arquivamento em seu esquema de arquivamento deve ter a mesma estrutura da sua tabela dinâmica, incluindo seus índices e chave primária. No entanto, a tabela de arquivamento de destino não pode ser uma tabela particionada. A troca de partições entre duas tabelas particionadas não é permitida no MySQL.
- Siga uma convenção de nomenclatura para sua tabela de arquivamento que o ajude a identificar os dados históricos armazenados nela. Isso é útil quando você executa tarefas de auditoria ou trabalhos de design que movem esses dados para o Amazon S3.
- Execute o `EXCHANGE PARTITION` declaração de linguagem de definição de dados (DDL) em uma janela de tempo de inatividade quando não há tráfego entrando em seu gravador compatível com o Aurora MySQL, Amazon RDS para MySQL ou Amazon RDS para instâncias do MariaDB.

Talvez seja possível executar `EXCHANGE PARTITION` durante janelas de baixo tráfego em seu aplicativo ou microsserviço. No entanto, não deve haver gravações e nenhuma ou muito poucas seleções na tabela particionada. As consultas selecionadas existentes e de longa duração podem fazer com que `EXCHANGE PARTITION` DDL aguarde, causando contenções de recursos em seu banco de dados. Crie scripts que verificam que todas essas condições foram atendidas antes da execução `EXCHANGE PARTITION` em seu sistema.

Se o design do seu aplicativo pode suportar dados particionados e você atualmente tem uma tabela não particionada, considere mover seus dados para tabelas particionadas para suportar o arquivamento de seus dados. Para obter mais informações, consulte [a documentação do MySQL](#).

## Arquivamento de dados de tabelas não particionadas

Em tabelas de banco de dados em que o particionamento não é possível, você pode usar o Percona Toolkit [pt-archiver](#) ferramenta para arquivar os dados da sua tabela em outra tabela em seu banco de dados MySQL.

A ferramenta pt-archiver é usada para arquivar os registros de tabelas grandes em outras tabelas ou arquivos. É uma ferramenta de leitura/gravação, o que significa que ela exclui dados da tabela de origem após arquivá-los, para que você não precise gerenciar a exclusão dos dados de origem separadamente. O objetivo principal desse script é arquivar dados antigos da tabela sem afetar a carga de consulta existente de processamento de transações on-line (OLTP) (consulte o Apêndice I) e inserir os dados em outra tabela no mesmo servidor ou em um servidor diferente.

Você pode baixar o [Kit de ferramentas Percona e instalar](#) em sua máquina local ou na instância do Amazon Elastic Compute Cloud (Amazon EC2) de onde você está se conectando ao banco de dados. Para executar a ferramenta pt-archiver, use a seguinte sintaxe.

```
pt-archiver --source h=<HOST>,D=<DATABASE>,t=<TABLE>,u=<USER>,p=<PASSWORD> --dest  
h=<HOST>,D=<DATABASE>,t=<TABLE> --where "'1=1'" --statistics
```

Substitua o HOST, DATABASE, TABLE, e USER com os detalhes e credenciais do banco de dados de origem e destino.

Você também pode usar [AWS Batch](#) para criar e agendar esse trabalho para suas tabelas.

Ao usar a ferramenta pt-archiver para arquivar os dados da sua tabela, considere o seguinte:

- Ter uma chave primária na tabela de origem melhorará o desempenho dessa ferramenta. Se a tabela não tiver uma chave primária, você pode [criar um índice em uma coluna exclusiva](#), o que ajudará o pt-archiver a percorrer todas as linhas da tabela e arquivá-las.
- Por padrão, pt-archiver exclui os dados depois de arquivar a tabela. Antes de executá-lo no servidor de produção, certifique-se de testar suas tarefas de arquivamento com `--dry-run`. Como alternativa, você pode usar o `--no-delete` opção.
- A ferramenta pt-archiver ajusta sua taxa de arquivamento com base na carga do sistema (consulte o Apêndice II). Com cargas maiores, você pode esperar um desempenho de arquivamento mais lento.

Depois de executar o `pt-archiver`, seus dados arquivados devem estar na tabela correspondente no esquema de arquivamento. A partir daí, você pode movê-lo para o Amazon S3.

# Movendo dados de tabelas arquivadas para o Amazon S3

O Amazon Simple Storage Service (Amazon S3) é o destino natural para dados arquivados. Ele fornece 99,999999999% de durabilidade e é mais barato do que o armazenamento de banco de dados.

Além disso, o Amazon S3 tem classes de armazenamento integradas que são precificadas com base no padrão de recuperação. Você tem a opção de fazer a transição dos objetos S3 descarregados para um nível de armazenamento mais barato com base na frequência de recuperação dos dados. Para obter mais informações sobre classes de armazenamento e preços, consulte a documentação do [Amazon S3](#).

Para aplicativos que usam frotas de Minhas SQL instâncias, transferir para o Amazon S3 significaria economizar dinheiro em dados que atendam aos seguintes critérios:

- Deve ser arquivado do banco de dados para aumentar a eficiência
- Não é necessário imediatamente ou é pouco necessário para qualquer processo de negócios
- Deve ser mantido por um longo prazo devido aos requisitos de auditoria

Você pode arquivar Meus SQL dados das seguintes formas:

- Exporte dados usando `SELECT INTO OUTFILE S3` o.
- Exporte dados usando o AWS Data Pipeline.
- Exporte dados usando o AWS Glue.

## Exportar dados usando o `SELECT INTO OUTFILE S3`

Para copiar dados do banco de dados Aurora My SQL -Compatible diretamente para o Amazon S3, você pode usar a declaração `SELECT INTO OUTFILE S3`. Essa SQL instrução pode ser executada na tabela e o número necessário de linhas pode ser descarregado como arquivos de valores (CSV) separados por vírgula com um tamanho máximo de 6 GB. Quando o limite de 6 GB é ultrapassado, vários arquivos.csv são criados.

Para que a exportação funcione, configure o seguinte:

- AWS Funções e políticas do Identity and Access Management (IAM)
- Permissões de banco de dados concedidas ao usuário para executar o comando
- O local do Amazon S3 para descarga

Para obter mais informações, consulte a documentação do [Amazon Aurora](#).

Nota: Para fazer a transição dos objetos do S3 para níveis de armazenamento econômicos, recomendamos configurar as regras do ciclo de vida do Amazon [S3 no bucket do S3](#) para o qual os dados estão sendo exportados.

## Exportar dados usando o AWS Data Pipeline

AWS Data Pipeline é um serviço web para fluxos de trabalho orientados por dados. Você pode usar o Data Pipeline para automatizar a movimentação de dados do Amazon RDS for My SQL para o Amazon S3. O Data Pipeline executa uma série de ações ou tarefas em série e exporta os dados para o Amazon S3 como um arquivo.csv.

O pipeline pode ser criado de uma das seguintes formas:

- No console, selecione o modelo Cópia completa de RDS Minha SQL tabela para o S3 para a fonte e forneça valores para os parâmetros.
- AWSCLIEm, defina os objetos do pipeline de dados comoJSON, crie o pipeline, faça o upload e ative.

A vantagem de usar o Data Pipeline é que o trabalho de exportação pode ser agendado e monitorado. Além disso, para mesas maiores, você pode usar a Amazon EMR em vez da AmazonEC2. Para obter mais informações e exemplos, consulte a [documentação do AWS Data Pipeline](#).

Nota: Para fazer a transição dos objetos do S3 para níveis de armazenamento econômicos, recomendamos configurar as regras do ciclo de vida do Amazon [S3 no bucket do S3](#) em que a exportação de dados está sendo feita.

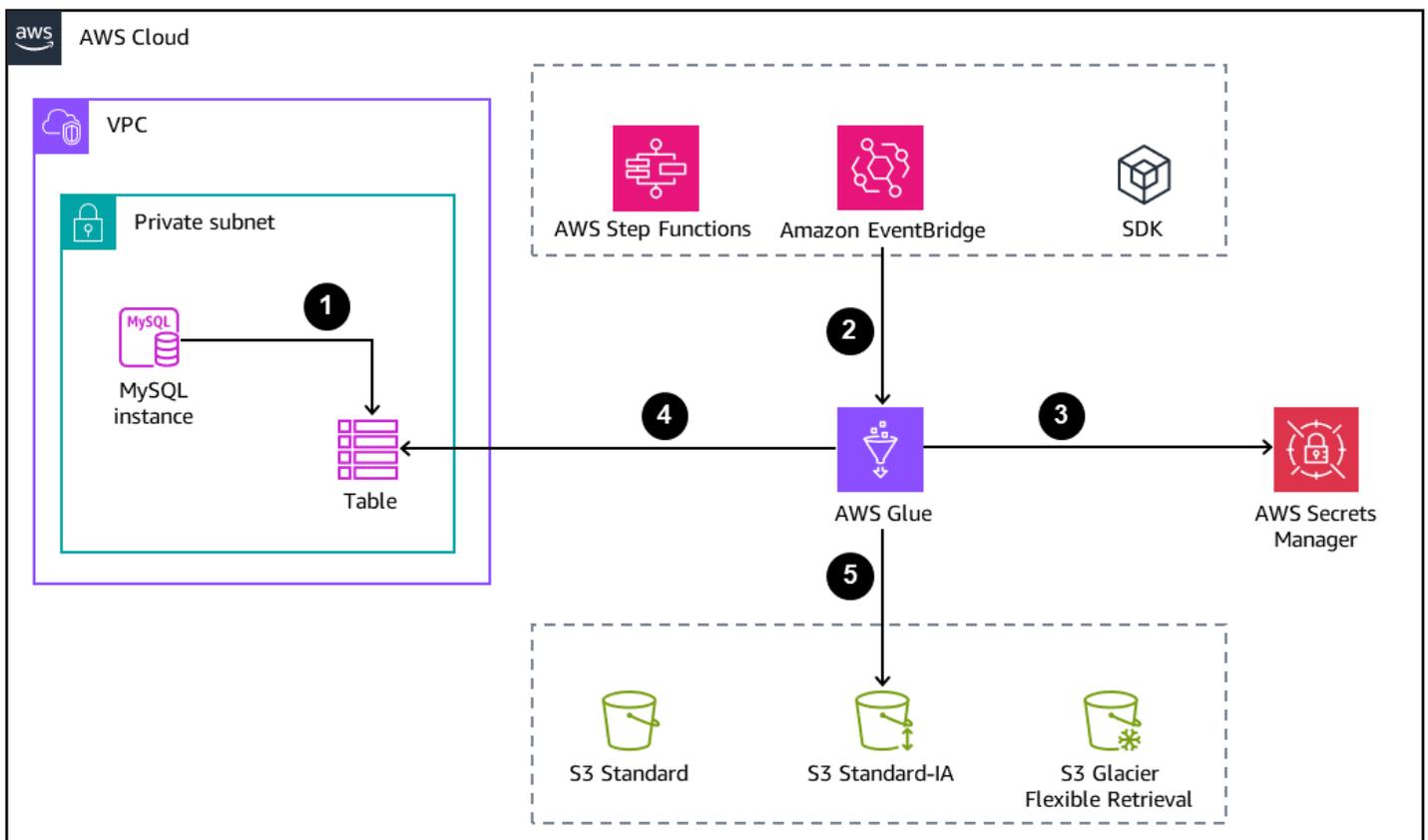
## Exportar dados usando o AWS Glue

Você pode arquivar Meus SQL dados no Amazon S3 usando o AWS Glue, que é um serviço analítico sem servidor para cenários de big data. AWS Glue é desenvolvido com o Apache Spark, uma

estrutura de computação em cluster distribuída amplamente usada que oferece suporte a muitas fontes de banco de dados.

O descarregamento de dados arquivados do banco de dados para o Amazon S3 pode ser realizado com algumas linhas de código em uma AWS tarefa do Glue. A maior vantagem que o AWS Glue oferece é a escalabilidade horizontal e um pay-as-you-go modelo, fornecendo eficiência operacional e otimização de custos.

O diagrama a seguir mostra uma arquitetura básica para arquivamento de banco de dados.



1. Meu SQL banco de dados cria o arquivo ou a tabela de backup para ser descarregada no Amazon S3.
2. Uma tarefa do AWS Glue é iniciada por uma das seguintes abordagens:
  - De forma síncrona como uma etapa dentro de uma máquina de estado [AWS Step Functions](#)
  - [De forma assíncrona por um evento da Amazon EventBridge](#)
  - Por meio de uma solicitação manual usando AWS CLI ou [AWSSDK](#)
3. As credenciais do banco de dados são recuperadas do [AWS Secrets Manager](#).

4. A tarefa AWS Glue usa uma conexão Java Database Connectivity (JDBC) para acessar o banco de dados e ler a tabela.
5. AWSGlue grava os dados no Amazon S3 no formato Parquet, que é um formato de dados aberto, colunar e que economiza espaço.

## Configurando o AWS Glue Job

Para funcionar conforme o esperado, o trabalho AWS Glue requer os seguintes componentes e configurações:

- [AWSConexões Glue](#) — Esse é um objeto do AWS Glue Data Catalog que você anexa à tarefa para acessar o banco de dados. Um trabalho pode ter várias conexões para fazer chamadas para vários bancos de dados. As conexões contêm as credenciais do banco de dados armazenadas com segurança.
- [GlueContext](#) — Este é um invólucro personalizado sobre [SparkContext](#) a GlueContext classe que fornece API operações de ordem superior para interagir com o Amazon S3 e fontes de banco de dados. Ele permite a integração com o Data Catalog. Também elimina a necessidade de depender de drivers para conexão com o banco de dados, que é feita na conexão Glue. Além disso, a GlueContext classe fornece maneiras de lidar com API as operações do Amazon S3, o que não é possível com a classe original SparkContext .
- IAMpolíticas e funções — Como o AWS Glue interage com outros AWS serviços, você deve configurar funções apropriadas com o mínimo de privilégio necessário. Os serviços que exigem permissões apropriadas para interagir com o AWS Glue incluem o seguinte:
  - Amazon S3
  - AWS Secrets Manager
  - AWSServiço de gerenciamento de chaves (AWSKMS)

## Melhores práticas

- Para ler tabelas inteiras que têm um grande número de linhas a serem descarregadas, recomendamos usar o endpoint da réplica de leitura para aumentar a taxa de transferência de leitura sem prejudicar o desempenho da instância principal do gravador.
- Para obter eficiência no número de nós usados para processar a tarefa, ative o [escalonamento automático](#) no AWS Glue 3.0.

- Se o bucket do S3 fizer parte da arquitetura do data lake, recomendamos descarregar os dados organizando-os em partições físicas. O esquema de partição deve ser baseado nos padrões de acesso. O particionamento com base em valores de data é uma das práticas mais recomendadas.
- Salvar os dados em formatos abertos, como Parquet ou Optimized Row Columnar (ORC), ajuda a disponibilizar os dados para outros serviços analíticos, como Amazon Athena e Amazon Redshift.
- Para que os dados descarregados sejam otimizados para leitura por outros serviços distribuídos, o número de arquivos de saída deve ser controlado. Quase sempre é benéfico ter um número menor de arquivos maiores em vez de um grande número de arquivos pequenos. O Spark tem arquivos de configuração e métodos integrados para controlar a geração de arquivos parciais.
- Os dados arquivados, por definição, são conjuntos de dados frequentemente acessados. Para obter eficiência de custo para armazenamento, a classe Amazon S3 deve ser transferida para níveis mais baratos. Isso pode ser feito usando duas abordagens:
  - [Transição síncrona do nível durante o descarregamento — Se você souber de antemão que os dados descarregados devem ser transferidos como parte do processo, você pode usar o mecanismo GlueContext transition\\_s3\\_path no mesmo trabalho do Glue que grava os dados no Amazon S3.](#) AWS
  - Transição assíncrona usando o ciclo de vida do [S3 — Configure as regras do ciclo](#) de vida do S3 com os parâmetros apropriados para a transição e expiração da classe de armazenamento do Amazon S3. Depois que isso for configurado no bucket, ele persistirá para sempre.
- Crie e configure uma sub-rede com um [intervalo de endereços IP suficiente](#) na nuvem privada virtual (VPC) em que o banco de dados está implantado. Isso evitará falhas no trabalho do AWS Glue causadas por um número insuficiente de endereços de rede quando um grande número de unidades de processamento de dados (DPUs) estiver configurado.

# Acessando dados arquivados no Amazon S3

O Amazon S3 fornece várias ferramentas para ler o conteúdo dos dados. No entanto, dependendo da classe de armazenamento, algumas etapas de pré-processamento podem ser necessárias. Esta seção inclui o seguinte:

- Leitura de objetos arquivados do S3 com a classe de armazenamento Standard, usando o AWS Glue ou o Amazon S3 Select
- Lendo um objeto S3 arquivado com as classes de armazenamento do S3 Glacier usando o Amazon S3 Glacier Select ou o S3 Batch Operations
- Práticas recomendadas

## Lendo objetos arquivados do S3 com a classe de armazenamento Standard

Você pode ler objetos do S3 que estão arquivados com a classe de armazenamento Amazon S3 usando o AWS Glue ou o Amazon S3 Select.

Usando o AWS Glue

Os dados transferidos do MySQL para o Amazon S3 mantêm a mesma rigidez estrutural e consistência típicas de um sistema de gerenciamento de banco de dados relacional (RDBMS).

[AWS Glue Crawler](#) rastreia objetos do S3, deduz os tipos de dados e cria metadados de tabela como uma tabela externa DDL. Ao configurar o trabalho do crawler, use o Amazon S3 como origem e especifique a localização do prefixo S3 em que todos os arquivos de dados são criados. Na configuração, inclua o seguinte:

- Opções de execução do crawler
- Preferência opcional de prefixo de tabela
- Banco de dados de destino para criar a tabela
- Funções do IAM com as permissões necessárias

Depois de invocar a tarefa, ela examinará os dados para inferir o esquema e preservá-lo em [Catálogo de dados do AWS Glue](#) como [Tabelas AWS Glue](#). As tabelas do AWS Glue são essencialmente

tabelas externas que podem ser consultadas com instruções SQL, como uma tabela de banco de dados normal, usando serviços analíticos como [Amazon Athena](#), [Amazon Redshift Spectrum](#) e o Apache Hive em [Amazon EMR](#). Para obter mais informações sobre o rastreador, consulte o [Documentação do AWS Glue](#).

Para arquivos.csv com um cabeçalho de coluna especificado, os nomes das colunas da tabela resultante refletirão os mesmos nomes de campo. O tipo de dados é inferido com base nos valores no objeto de dados.

Para arquivos Parquet, o esquema é preservado nos próprios dados e a tabela resultante refletirá os mesmos nomes de campo e tipo de dados.

Como alternativa, você pode executar uma DDL manualmente no Athena para criar a definição da tabela com os nomes de coluna e o tipo de dados necessários. Isso cria a definição da tabela no Catálogo de Dados. Para obter mais informações sobre a criação de tabelas do Athena, consulte o [Documentação do Amazon Athena](#).

Nota: Se a linha do cabeçalho estiver ausente do arquivo CSV, o rastreador criará o nome do campo como genérico c\_0, c\_1, c\_2,...

### Usando o Amazon S3 Select

Você pode usar o Amazon S3 Select para ler os objetos do S3 programaticamente usando expressões SQL. A operação da API pode ser invocada usando o comando `AWS CLI select-object-content` usando um SDK como o Boto3 e invocando a operação `select_object_content` do Python.

As operações da API suportam instruções SQL como parâmetros e só podem ler arquivos do tipo JSON e Parquet. As saídas podem ser redirecionadas como arquivos de saída.

Essas operações são invocadas para cada objeto do S3. Para vários arquivos, execute as operações recursivamente.

Para obter mais informações sobre como executar as operações usando o AWS CLI, consulte o [Documentação da AWS CLI](#). Para obter mais informações sobre como executar o S3 Select usando o SDK do Python Boto3, consulte [Documentação do Boto3](#).

## Lendo objetos arquivados do S3 com as classes de armazenamento do S3 Glacier

As classes do Amazon S3 Glacier são classes especiais de armazenamento com preços baratos, mas com alto tempo de recuperação. Diferentemente dos objetos S3 Standard, os objetos do S3 Glacier não podem ser lidos como tabelas do AWS Glue. Para disponibilizar os dados para consultas analíticas ou relatórios, primeiro você restaura os objetos do S3 Glacier. A restauração é um processo assíncrono que acontece ao longo do tempo e tem um período de retenção. Depois que os objetos forem restaurados, eles poderão ser copiados para um local diferente como objetos do S3 Standard. Além do período de retenção, os objetos restaurados retornam ao Amazon S3 Glacier.

### Usando o Amazon S3 Glacier Select

Semelhante ao uso do Amazon S3 Select com o S3 Standard, você pode consultar objetos do S3 Glacier para buscar um subconjunto de dados. Isso permite o acesso programático aos dados sem a necessidade de qualquer pré-processamento, como restauração de objetos ou outros serviços analíticos da AWS. Para ver um exemplo, consulte [osintaxe de solicitação](#) por usar o `Boto3initiate_job` operação para ler dados do S3 Glacier Select.

### Usando operações em lote do S3

O S3 Batch Operations permite operações em lote em grande escala no Amazon S3 na ordem de bilhões de objetos contendo exabytes de dados. O Amazon S3 monitora o progresso, envia notificações e armazena um relatório de conclusão detalhado de todas as ações, fornecendo uma experiência totalmente gerenciada, auditável e sem servidor.

O S3 Batch Operations suporta o [Restaurar](#) operação, que inicia a restauração de objetos do S3 para os seguintes níveis de armazenamento:

- Objetos arquivados nas classes de armazenamento S3 Glacier Flexible Retrieval ou S3 Glacier Deep Archive
- Objetos arquivados por meio da classe de armazenamento do S3 Intelligent-Tiering nos níveis Acesso de arquivamento ou Acesso de arquivamento profundo

A operação em lote pode ser invocada tanto programaticamente quanto no console do Amazon S3. Para entrada, é necessário um arquivo de manifesto .csv que contenha os objetos da lista a serem restaurados.

Você pode usar um [Inventário do Amazon S3](#) relatório como entrada para o trabalho em lote. O relatório de inventário é configurado para um bucket e pode ser limitado a objetos sob prefixos específicos. É um relatório automatizado e é gerado semanalmente ou diariamente no formato CSV, ORC ou Parquet.

Para obter mais informações sobre como configurar um relatório de inventário, consulte [Documentação do Amazon S3](#). Para obter informações sobre como usar o Boto3 para criar um trabalho de operações em lote do S3, consulte o [Documentação do Boto3](#).

## Práticas recomendadas

Recomendamos as seguintes práticas recomendadas para acessar dados arquivados:

- O Amazon S3 Select e o Amazon S3 Glacier Select são adequados para casos em que um subconjunto de dados é a saída usando expressões SQL básicas. Os aplicativos que precisam consultar o Amazon S3 programaticamente para derivar conjuntos de dados selecionados mostram uma melhoria drástica no desempenho — em muitos casos, uma melhoria de até 400%. Para obter mais informações, consulte a postagem do blog [S3 Select e S3 Glacier Select — Recuperando subconjuntos de objetos](#).
- O S3 Select e o S3 Glacier Select suportam a leitura de dados somente nos formatos CSV, JSON e Parquet. Para outros formatos abertos, como ORC ou Avro, você pode usar o Amazon Athena ou o Amazon Redshift.
- Para grandes conjuntos de dados arquivísticos, recomendamos a criação de tabelas do AWS Glue sobre os dados para que possam ser lidas usando mecanismos de consulta como o Athena e o Amazon Redshift. Tanto o Athena quanto o Amazon Redshift oferecem escalabilidade horizontal do desempenho da consulta. Eles também usam um pay-per-query modelo, que é econômico em um cenário de consulta única. Além disso, o Amazon Redshift tem mecanismos do Advanced Query Accelerator (AQUA) sob o capô, que aceleram o desempenho de leitura sem nenhum custo adicional.
- Os dados arquivados descarregados regularmente no Amazon S3 não devem ser armazenados como uma pilha de lixo. Em vez disso, ela deve ser salva como uma nova partição. Uma partição de data separará os dados em dimensões de data (por exemplo, `year=<value>/month=<value>/day=<value>`). Isso é extremamente benéfico em duas situações:
  - Se as tabelas do AWS Glue forem criadas pelos rastreadores do AWS Glue, essas partições funcionarão como pseudocolunas. Isso melhora o desempenho de leitura ao restringir os dados digitalizados às partições na consulta de intervalo.

- Isso ajuda em uma operação de restauração do S3 Glacier quando você está restaurando somente um subconjunto do objeto como S3 Standard.
- Os rastreadores do AWS Glue mostram grande valor quando os dados arquivados salvos no Amazon S3 são particionados fisicamente. Sempre que esses dados são descarregados como uma nova partição de prefixo, o rastreador verifica somente a nova partição e atualiza os metadados dessa partição. Se o esquema da tabela mudar, essas alterações serão capturadas nos metadados em nível de partição.

## Limpeza da tabela de arquivamento

A etapa final do processo de arquivamento é limpar as tabelas em seu esquema de arquivamento. Você pode fazer isso depois de confirmar que seus dados de arquivamento estão arquivados com segurança no Amazon S3. Para evitar qualquer impacto em seu aplicativo, recomendamos eliminar as tabelas do esquema de arquivamento durante um tempo de inatividade programado ou uma janela de manutenção ou durante uma janela de tráfego muito baixo em seu aplicativo. Essas tabelas não são consultadas ativamente pelo seu aplicativo e não devem ser motivo de alarme por seu impacto nas transações em andamento. Ainda assim, é uma prática recomendada executar DDLs durante um tempo de inatividade.

Depois que o armazenamento de grandes tabelas de esquema de arquivamento é liberado, o Amazon Aurora usa [redimensionamento dinâmico](#) para ajudar você a economizar nos custos de armazenamento.

# Recursos

## Documentação

- [Amazon Aurora](#)
- [AWS Data Pipeline](#)
- [AWS Glue](#)
- [Amazon S3](#)
- [Consultando o inventário do Amazon S3 com o Amazon Athena](#)
- [Configurando o inventário do Amazon S3](#)
- [Execução de operações em lote em grande escala em objetos do Amazon S3](#)
- [Documentação do Boto3](#)
- [Documentação de particionamento do MySQL](#)

## Definição de preço

- [Preços do Amazon RDS para MySQL Multi-AZ GP2](#)
- [Preços do Amazon S3 Glacier](#)

## Ferramentas

- [Kit de ferramentas Percona](#)
- [pt-archiver](#)
- [sysbench](#)

## Postagens no blog

- [S3 Select e S3 Glacier Select — Recuperando subconjuntos de objetos](#)
- [Arquive e elimine dados do Amazon RDS para PostgreSQL e do Amazon Aurora com compatibilidade com PostgreSQL usando pg\\_partman e Amazon S3](#)

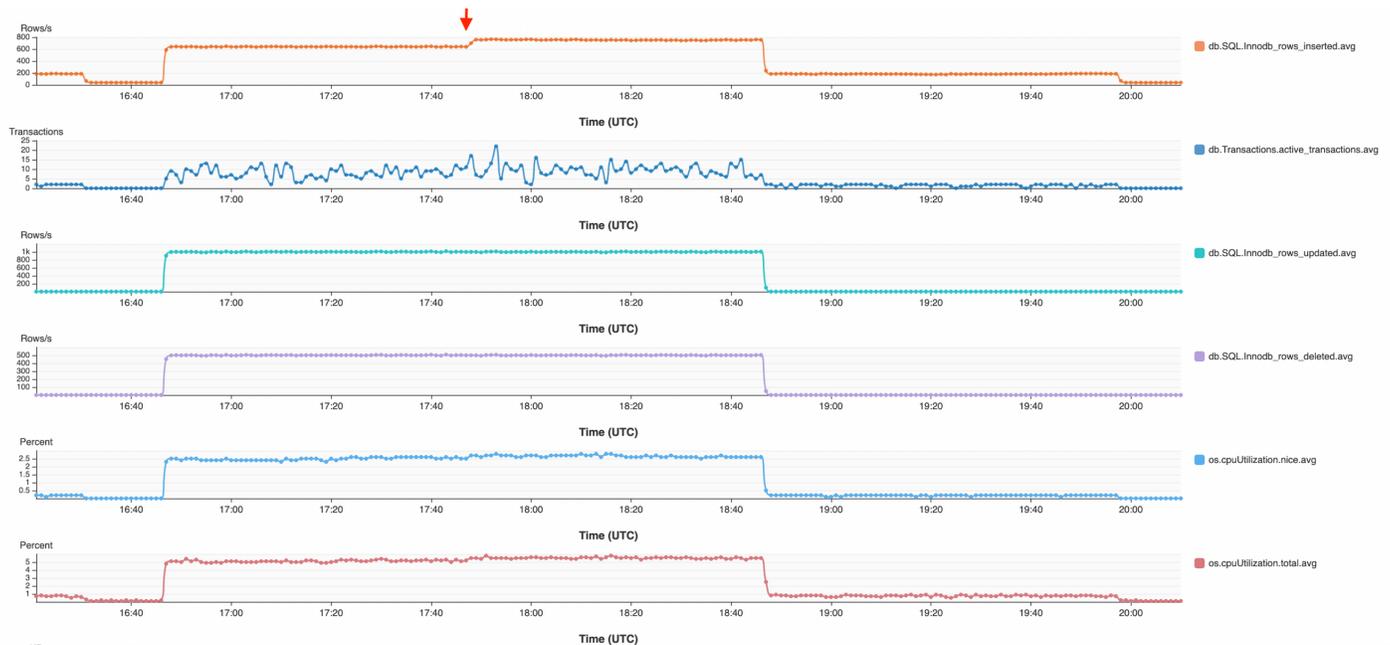
# Apêndice I

Todos os testes são realizados em uma instância do Amazon RDS para MySQL em execução `nodb.r6g.8xlarge` classe de instância.

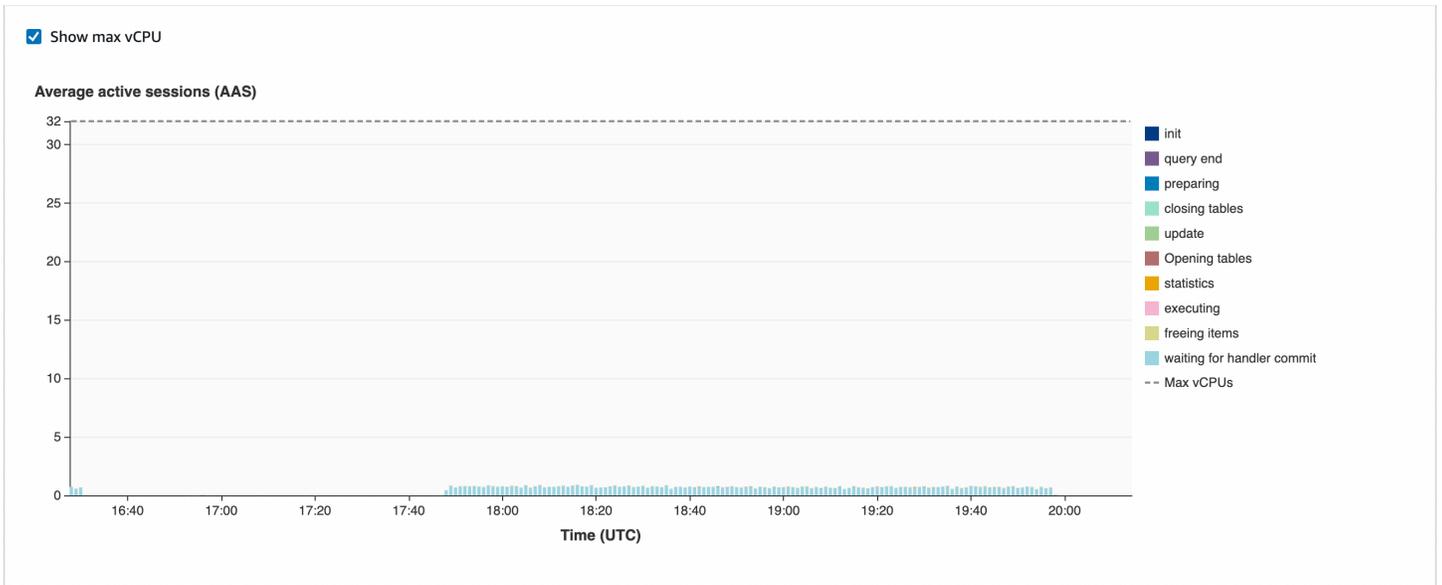
O seguintes [sysbench](#) comandos foram usados para preparar e executar a carga no banco de dados.

```
sysbench oltp_read_write --db-driver=mysql --mysql-db=<DATABASE> --mysql-user=<USER> --mysql-password=<PASSWORD> --mysql-host=<ENDPOINT> --tables=500 --table-size=2000000 --threads=500 prepare
sysbench oltp_read_write --db-driver=mysql --mysql-db=employees --mysql-user=admin --mysql-password=qwertyuiop --mysql-host=mysql8.cbbhujzeoxed.us-east-1.rds.amazonaws.com --tables=500 --rate=500 --time=7200 run
```

No gráfico a seguir, uma carga de trabalho OLTP estava em execução e o processo `pt-archiver` começou onde a seta está marcada.



Não há nenhuma mudança significativa na utilização da CPU com o `pt-archiver` em execução em paralelo, o que infere que o `pt-archiver` não afeta as consultas OLTP durante a execução.



## Apêndice II

Esta seção fornece os resultados do benchmarking para ferramentas pt-archiver em diferentes cenários. O [sysbench](#) ferramenta é usada neste teste para colocar carga no banco de dados. Todos os testes são realizados na instância do Amazon RDS para MySQL em execução `emdb.r6g.8xlarge` classe de instância.

Os seguintes comandos sysbench foram usados para preparar e executar a carga no banco de dados:

```
sysbench oltp_read_write --db-driver=mysql --mysql-db=<DATABASE> --mysql-user=<USER> --mysql-password=<PASSWORD> --mysql-host=<ENDPOINT> --tables=1000 --table-size=2000000 --threads=500 prepare
sysbench oltp_read_write --db-driver=mysql --mysql-db=<DATABASE> --mysql-user=<USER> --mysql-password=<PASSWORD> --mysql-host=<ENDPOINT> --tables=1000 --rate=500 --threads=500 run
```

Arquivamento de uma tabela sem chave primária e somente um índice (sem carga no banco de dados)

```
Started at 2022-11-07T05:29:12, ended at 2022-11-07T06:03:31
Action      Count          Time           Pct
commit     600050         1715.3582     83.31
select     300025         166.5470      8.09
inserting  300024         165.4025      8.03
other      0              11.6644       0.57
```

Foram necessários cerca de 34 minutos para arquivar 300.024 linhas. Essa tabela tinha 2 milhões de linhas, mas a ferramenta arquivou somente as linhas com dados exclusivos para a coluna indexada.

Arquivamento de uma tabela que tem uma chave primária (sem carga no banco de dados)

```
Started at 2022-11-16T08:53:49, ended at 2022-11-16T12:38:18
Action      Count          Time           Pct
commit     4000000        11065.9534    82.16
select     2000000        1278.1854     9.49
inserting  1999999        1050.4961     7.80
other      0              74.1519       0.55
```

Demorou cerca de 3 horas, 44 minutos e 29 segundos para arquivar 1.999.999 linhas.

O gráfico a seguir mostra que o pt-archiver consome muito pouca CPU e recursos quando executado sozinho, sem qualquer carga existente no sistema.

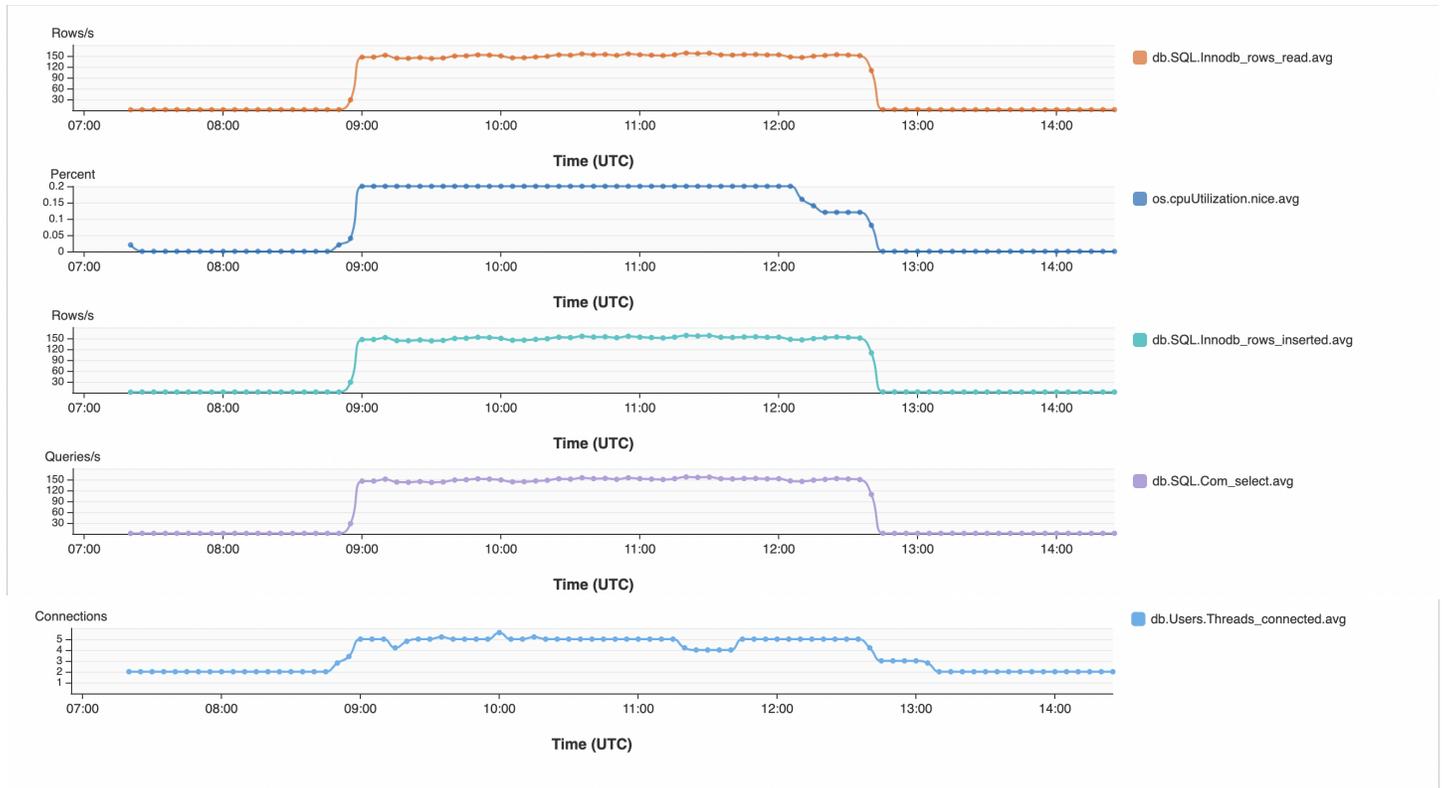


Tabela de arquivamento que tem uma chave primária (com carga no banco de dados)

```

Started at 2022-11-16T17:37:07, ended at 2022-11-17T03:20:43
Action      Count      Time      Pct
commit     4000000   19688.8362  56.23
inserting  1999999   13933.4418  39.79
select     2000000   1305.1770   3.73
other      0         89.1787    0.25
    
```

Demorou cerca de 9 horas, 43 minutos e 36 segundos para arquivar 1999999 linhas.

O gráfico a seguir mostra que, durante o teste, a utilização da CPU foi de até 15% devido à carga aplicada pelo sysbench. Após a conclusão do carregamento, o pt-archiver continuou a funcionar, consumindo o mínimo de CPU, conforme o esperado para concluir o arquivamento.



Como é evidente nos gráficos, o pt-archiver não arquiva agressivamente quando há uma carga em seu banco de dados.

## Histórico do documentos

A tabela a seguir descreve mudanças significativas neste guia. Se você quiser ser notificado sobre futuras atualizações, você pode assinar um [Feed RSS](#).

Alteração	Descrição	Data
<a href="#">Publicação inicial</a>	—	8 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: migre seu banco de dados Oracle local para a edição compatível com o Amazon Aurora 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: Migre seu banco de dados Oracle local para o Amazon Relational Database Service (Amazon RDS) for Oracle no. Nuvem AWS
- Recomprar (drop and shop): mude para um produto diferente, normalmente migrando de uma licença tradicional para um modelo SaaS. Exemplo: migre seu sistema de gerenciamento de relacionamento com o cliente (CRM) para a 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: Migre seu banco de dados Oracle local para o Oracle em uma instância do EC2 no. 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 local para um serviço em 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 atributos](#).

serviços abstratos

Veja os [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 no qual 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.

função agregada

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 as [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 aplicativos

Uma abordagem de segurança que permite o uso somente de aplicativos aprovados 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 as AIOps são usadas na estratégia de migração para a AWS, 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.

## Availability Zone (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 ruim

Um [bot](#) destinado a perturbar ou causar danos a indivíduos ou organizações.

### BCP

Veja o [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 do aplicativo em um ambiente (azul) e a nova versão do aplicativo no outro ambiente (verde). Essa estratégia ajuda você a reverter rapidamente com o mínimo de impacto.

## bot

Um aplicativo de software que executa tarefas automatizadas pela Internet e simula a atividade ou interação humana. Alguns bots são úteis ou benéficos, como rastreadores da Web que indexam informações na Internet. Alguns outros bots, conhecidos como bots ruins, têm como objetivo perturbar ou causar danos a indivíduos ou organizações.

## botnet

Redes de [bots](#) infectadas por [malware](#) e sob o controle de uma única parte, conhecidas como pastor de bots ou operador de bots. As redes de bots 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 em vidro quebrado

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 [Implementar procedimentos de quebra de vidro na orientação do Well-Architected AWS](#) .

## 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 containerizados 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

Consulte [Estrutura de adoção da AWS nuvem](#).

## implantação canária

O lançamento lento e incremental de uma versão para usuários finais. Quando estiver confiante, você implanta a nova versão e substituirá a versão atual em sua totalidade.

## CCoE

Veja o [Centro de Excelência em Nuvem](#).

## CDC

Veja [a captura de dados de alterações](#).

## 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 a [integração e a 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 da Nuvem (CCoE)

Uma equipe multidisciplinar que impulsiona os esforços de adoção da nuvem em toda a organização, incluindo o desenvolvimento de práticas recomendadas 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 [postagens do CCoE 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 geralmente está conectada à tecnologia de [computação de ponta](#).

## 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 quando migram para o Nuvem AWS:

- Projeto: executar alguns projetos relacionados à nuvem para fins de prova de conceito e aprendizado
- Fundação: realizar investimentos fundamentais para escalar sua adoção da nuvem (por exemplo, criar uma zona de pouso, definir um CCoE, 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

Consulte o [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 GitHub ou AWS CodeCommit. 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 da [IA](#) que usa aprendizado de máquina para analisar e extrair informações de formatos visuais, como imagens e vídeos digitais. Por exemplo, AWS Panorama oferece dispositivos que adicionam CV às redes de câmeras locais, e a Amazon SageMaker fornece algoritmos de processamento de imagem para CV.

## desvio de configuração

Para uma carga de trabalho, uma alteração de configuração em relação ao estado esperado. Isso pode fazer com que a carga de trabalho 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. O CI/CD é comumente descrito como um pipeline. O CI/CD pode ajudar você 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.

#### malha de dados

Uma estrutura arquitetônica 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 que oferece suporte à inteligência comercial, como análises. 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

Consulte a [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 o [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 são comumente 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 [desastre](#). Para obter mais informações, consulte [Recuperação de desastres de cargas de trabalho em AWS: Recuperação na nuvem no AWS Well-Architected Framework](#).

## DML

Consulte [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 a [recuperação de desastres](#).

## detecção de deriva

Rastreando desvios de uma configuração básica. 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 o [mapeamento do fluxo de valor do desenvolvimento](#).

## E

### EDA

Veja a [análise exploratória de dados](#).

### computação de borda

A tecnologia que aumenta o poder computacional de dispositivos inteligentes nas bordas de uma rede de IoT. Quando comparada à [computação em nuvem](#), a computação de ponta pode reduzir a latência da comunicação e melhorar o tempo de resposta.

### 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 o [endpoint do 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 corporativos (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).

## environment (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 pipeline de CI/CD, 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

Consulte [planejamento de recursos corporativos](#).

## 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](#). Ele armazena dados quantitativos sobre operações comerciais. Normalmente, uma tabela de fatos contém dois tipos de colunas: aquelas que contêm medidas e aquelas que contêm uma chave externa para uma tabela de dimensões.

### falham rapidamente

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

### limite de isolamento de 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 [Limites de isolamento de AWS falhas](#).

### ramificação de recursos

Veja a [filial](#).

### 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.

## FGAC

Veja o [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 por meio da [captura de dados alterados](#) para migrar dados no menor tempo possível, em vez de usar uma abordagem em fases. O objetivo é reduzir ao mínimo o tempo de inatividade.

## G

### bloqueio geográfico

Veja as [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 baseado em troncos](#) é a abordagem moderna e preferida.

### 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 gerenciar recursos, políticas e conformidade em todas as unidades organizacionais (UOs). 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, Amazon GuardDuty AWS Trusted Advisor, Amazon Inspector e verificações personalizadas AWS Lambda .

# H

## HA

Veja a [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.

## 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 típico de uma DevOps versão.

## 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.

## I

## IaC

Veja a [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 a [Internet das Coisas industrial](#).

### infraestrutura imutável

Um modelo que implanta uma nova infraestrutura para cargas de trabalho de produção em vez de atualizar, corrigir ou modificar a infraestrutura existente. [Infraestruturas imutáveis são inerentemente mais consistentes, confiáveis e previsíveis do que infraestruturas mutáveis](#). Para obter mais informações, consulte as melhores práticas de [implantação usando infraestrutura imutável](#) no Well-Architected AWS 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 segurança da AWS](#) recomenda configurar sua conta de rede com VPCs de entrada, saída e inspeção para proteger a interface bidirecional entre a aplicação 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 fabricação por meio de avanços em conectividade, dados em tempo real, automação, análise 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 das Coisas Industrial (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 [Construir uma estratégia de transformação digital para a Internet das Coisas Industrial \(IIoT\)](#).

### 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 segurança da AWS](#) recomenda configurar sua conta de rede com VPCs de entrada, saída e inspeção para proteger a interface bidirecional entre a aplicação 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 machine learning com a AWS](#).

### IoT

Consulte [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

Consulte [a biblioteca de informações](#) de TI.

## ITSM

Veja o [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](#).

### migração de grande porte

Uma migração de 300 servidores ou mais.

### LBAC

Veja controle de [acesso baseado em etiquetas](#).

### 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](#).

## ambientes inferiores

Veja o [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 a [filial](#).

## malware

Software projetado para comprometer a segurança ou a privacidade do computador. O malware pode interromper os sistemas do computador, vazar informações confidenciais ou obter acesso não autorizado. Exemplos de malware incluem vírus, worms, ransomware, cavalos de Tróia, 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 abstratos.

## 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

Consulte [Migration Acceleration Program](#).

## mecanismo

Um processo completo no qual 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-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 o [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 por meio de APIs bem definidas e normalmente pertence a equipes pequenas e autônomas. 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 APIs leves. 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: rehoste 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 carga de trabalho para o. Nuvem AWS Para obter mais informações, consulte a entrada de [7 Rs](#) neste glossário e consulte [Mobilize sua organização para acelerar migrações em grande escala](#).

### ML

Veja o [aprendizado de máquina](#).

### 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 [Estratégia para modernizar aplicativos no 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 [Avaliação da prontidão para modernização de aplicativos no. 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](#).

## MAPA

Consulte [Avaliação do portfólio de migração](#).

## MQTT

Consulte 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 cargas de trabalho 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 o [controle de acesso de origem](#).

### CARVALHO

Veja a [identidade de acesso de origem](#).

### OCM

Veja o [gerenciamento de mudanças 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 a [integração de operações](#).

## OLA

Veja o [contrato em 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

Consulte [Comunicação de processo aberto — Arquitetura unificada](#).

### Comunicação de processo aberto — Arquitetura unificada (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 melhores práticas 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 Well-Architected AWS Framework.

### tecnologia operacional (OT)

Sistemas de hardware e software que funcionam com o ambiente físico para controlar operações, equipamentos e infraestrutura industriais. Na manufatura, a integração dos sistemas OT e de tecnologia da informação (TI) 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 todas as 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.

## OU

Veja a [análise de prontidão operacional](#).

## NÃO

Veja a [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 segurança da AWS](#)

recomenda configurar sua conta de rede com VPCs de entrada, saída e inspeção para proteger a interface bidirecional entre a aplicação 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

Consulte [controlador lógico programável](#).

### AMEIXA

Veja o gerenciamento [do ciclo de vida do produto](#).

### política

Um objeto que pode definir permissões (consulte a [política baseada em identidade](#)), especificar as condições de acesso (consulte a [política baseada em recursos](#)) ou definir as permissões máximas para todas as contas em uma organização em AWS Organizations (consulte a 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. Para obter mais informações, consulte [Habilitar a persistência de dados em microsserviços](#).

## 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 `WHERE` cláusula.

## pressão de predicados

Uma técnica de otimização de consulta 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 o desempenho 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 engenharia.

## zonas hospedadas privadas

Um contêiner que armazena informações sobre como você quer que o Amazon Route 53 responda a consultas ao DNS para um domínio e seus subdomínios dentro de uma 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](#) projetado para impedir a implantação de recursos não compatíveis. Esses controles examinam 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 o design, desenvolvimento e lançamento, passando pelo crescimento e maturidade, até o declínio e a remoção.

## ambiente de produção

Veja o [ambiente](#).

## controlador lógico programável (PLC)

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

## 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.

## publicar/assinar (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 no qual outros microsserviços possam se inscrever. 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, responsável, consultado, informado \(RACI\)](#).

### 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, responsável, consultado, informado \(RACI\)](#).

### RCAC

Veja o [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.

### rearquiteta

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 mais informações, consulte [Especificar o que Regiões da AWS sua conta pode usar](#).

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 um aplicativo de resistir ou se recuperar de interrupções. [Alta disponibilidade e recuperação de desastres](#) são considerações comuns ao planejar a resiliência no. Nuvem AWS Para obter mais informações, consulte [Nuvem AWS Resiliência](#).

## 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](#).

## aposentar-se

Veja [7 Rs](#).

## rotação

O processo de atualizar periodicamente um [segredo](#) para dificultar o acesso das credenciais por um invasor.

## 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 o [objetivo do ponto de recuperação](#).

## RTO

Veja o [objetivo do 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 AWS Management Console 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 a [política de controle de serviços](#).

### secret

Em AWS Secrets Manager, informações confidenciais ou restritas, como uma senha ou credenciais de usuário, que você armazena de forma criptografada. Ele 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 [O que há em um segredo do Secrets Manager?](#) na documentação do Secrets Manager.

### 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 principais de controles de segurança: preventivos, detectivos, responsivos e proativos.](#)

## fortalecimento 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 correção de uma instância do Amazon EC2 ou a rotação 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 no AWS Organizations. As SCPs definem barreiras de proteção ou estabelecem limites para as ações que um administrador pode delegar a usuários ou perfis. É possível usar SCPs como listas de permissão ou 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 medida de um aspecto de desempenho de um serviço, como taxa de erro, disponibilidade ou taxa de transferência.

## objetivo de nível de serviço (SLO)

Uma métrica alvo que representa a integridade de um serviço, conforme medida 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 [informações de segurança e sistema de gerenciamento de eventos](#).

## ponto único de falha (SPOF)

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

## SLA

Veja o contrato [de nível de serviço](#).

## ESGUIO

Veja o indicador [de nível de serviço](#).

## SLO

Veja o objetivo do [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 [Abordagem em fases para modernizar aplicativos no](#). Nuvem AWS

## CUSPE

Veja [um único ponto de falha](#).

## esquema de estrelas

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 uso 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. Uma sub-rede deve residir em uma única zona de disponibilidade.

## controle de supervisão 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 o desempenho. Você pode usar o [Amazon CloudWatch Synthetics](#) para criar esses testes.

# 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. 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 o [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 pode ser usado para interconectar as VPCs e as redes on-premises. 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. Para obter mais informações, consulte o guia [Como quantificar a incerteza em sistemas de aprendizado profundo](#).

### 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 o [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 de VPC

Uma conexão entre duas VPCs que permite rotear 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 back-end.

## 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.

## MINHOCA

Veja [escrever uma vez, ler muitas](#).

## WQF

Consulte o [AWS Workload Qualification Framework](#).

## escreva uma vez, leia muitas (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, geralmente malware, que tira proveito de uma vulnerabilidade de [dia zero](#).

### 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.

### 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.